World Psychiatry

OFFICIAL JOURNAL OF THE WORLD PSYCHIATRIC ASSOCIATION (WPA)

Volume 24, Number 3 *October 2025*

WILEY NO. 1 OF PSYCHIATRIC JOURNALS FOR IMPACT FACTOR ISSN 1723-8617

The World Psychiatric Association (WPA)

The WPA is an association of national psychiatric societies aimed to increase knowledge and skills necessary for work in the field of mental health and the care for the mentally ill. Its member societies are presently 147, spanning 123 different countries and representing more than 250,000 psychiatrists.

The WPA organizes the World Congress of Psychiatry every year. It also organizes international and regional congresses and meetings, and thematic conferences. It has 66 scientific sections, aimed to disseminate information and promote collaborative work in specific domains of psychiatry. It has produced several educational programmes and series of books. It has developed ethical guidelines for psychiatric practice, including the Madrid Declaration (1996).

Further information on the WPA can be found on the website www.wpanet.org.

WPA Secretariat
Geneva University Psychiatric Hospital, 2 Chemin du Petit Bel-Air, 1226 Thônex, Geneva, Switzerland. Phone: +41223055737; Fax: +41223055735; E-mail: wpasecretariat@wpanet.org.

World Psychiatry

World Psychiatry is the official journal of the World Psychiatric Association. It is published in three issues per year and is sent free of charge to psychiatrists whose names and addresses are provided by WPA member societies and sections.

Research Reports containing unpublished data are welcome for submission to the journal. They should be subdivided into four sections (Introduction, Methods, Results, Discussion). References should be numbered consecutively in the text and listed at the end according to the following style:

1. Cuijpers P, Sijbrandij M, Koole SL et al. Adding psychotherapy to antidepressant medication in depression and anxiety disorders: a meta-analysis. World Psychiatry 2014;13: 56-67.
2. McRae TW. The impact of computers on accounting. London: Wiley, 1964.
3. Fraeijs de Veubeke B. Displacement and equilibrium models in the finite element method. In: Zienkiewicz OC, Hollister GS (eds). Stress analysis. London: Wiley, 1965:145-97.

All submissions should be sent to the office of the Editor.

Office of the Editor – Department of Psychiatry, University of Campania "L. Vanvitelli", Largo Madonna delle Grazie, 80138 Naples, Italy. Phone: +390815666502; Fax: +390815666523; E-mail: mario.maj@unicampania.it.

World Psychiatry **is indexed in PubMed, Current Contents/Clinical Medicine, Current Contents/Social and Behavioral Sciences, Science Citation Index, and EMBASE.**
All back issues of *World Psychiatry* can be downloaded free of charge from the WPA website (https://www.wpanet.org/world-psychiatry-volumes/?language=english).

The cross-national mental disorder vulnerability paradox

The World Mental Health (WMH) surveys suggest that common mental disorders are more prevalent in high-income than lower-income countries[1]. This is striking, because individual-level socioeconomic status (SES) within countries is inversely associated with mental disorder prevalence. Based on this within-country association, one might expect that higher national economic development, all else equal, would lead to lower national prevalence of mental disorders. That this opposite-sign pattern exists has been called the cross-national mental disorder vulnerability paradox[2].

The term *paradox* has long been used to describe cases where opposite-sign patterns exist across different levels of aggregation. Such cases are often referred to as examples of Simpson's paradox, named after the statistician who formalized its requirements in 1951. However, calling such patterns paradoxes can be misleading. As Simpson noted, individual-level and aggregate-level associations align only if *all else is equal* across levels. If there are unmeasured third variables that are unevenly distributed across units and associated with the outcome at the individual level, the aggregate association can reverse.

Well-known examples abound. One involves university admissions, where women sometimes have higher acceptance rates within individual departments of a university, but a lower acceptance rate overall. How is this possible? Because women apply disproportionately to more competitive departments. In this case, department choice is the unmeasured third variable. Another well-known example involves fertility: while higher SES is associated with fewer children overall, income is often associated positively with the number of children within SES groups, because higher income allows more children to be afforded.

Both methodological and substantive explanations have been proposed to account for the opposite-sign associations in the WMH surveys[3]. Methodological explanations suggest that there might be greater reluctance to report psychological problems in low-income countries or there are problems in WMH survey question wording that obscure a true negative association between economic development and mental disorder prevalence. Substantive explanations suggest that features of modern life, such as social disconnection and competitive pressures, increase mental disorder prevalence and outweigh the benefits of having more objective economic resources.

Support for the methodological explanation comes from large cross-national public opinion surveys, such as the annual Gallup World Poll of over 100,000 respondents per year conducted in more than 150 countries, which find that country-level economic development is associated with higher levels of psychological well-being, as measured by brief survey questions about happiness, sadness, worry, life satisfaction, and hope[4]. This has led some commentators to argue that the true association between SES and mental disorders is negative at both the individual and country levels, and that the WMH finding of an opposite-sign pattern is an artifact of measurement error[5].

However, a similar opposite-sign pattern was found in a recent 16-country study of ICD-11 prolonged grief disorder (PGD), which did not include WMH data[6]. Reluctance to admit distress seems an unlikely explanation in this case, as nearly all bereaved respondents reported some distress. Diagnoses of PGD instead depended on the persistence and severity of symptoms, which were more prevalent in economically developed countries.

A comparable pattern appears in WMH data: the endorsement of diagnostic stem questions for common mental disorders (e.g., two weeks of persistent dysphoria or anhedonia for major depressive disorder) does not differ across country income levels. What differs is the conditional probability of meeting full diagnostic criteria given endorsement of stem questions for major depressive disorder, generalized anxiety disorder, and post-traumatic stress disorder (PTSD). In the case of PTSD, for example, some potentially traumatic life experiences (e.g., wars, sectarian violence, natural disasters, premature deaths of loved ones) are less common in high-income countries, but conditional risk of PTSD given trauma exposure is significantly higher in high-income countries. This higher conditional risk overwhelms the lower prevalence of trauma exposure to result in a higher PTSD prevalence in high-income countries. Such an opposite-sign pattern could account for the simultaneous existence of a negative country-level association between economic development and unhappiness, as found in the Gallup surveys, and a positive country-level association between economic development and clinically significant mental disorder, as found in the WMH surveys.

It is interesting to note that a similar opposite-sign cross-level pattern exists for suicide. Within countries, lower SES is associated with higher suicide risk. Yet the suicide rate is higher in high-income countries compared to upper-middle-income countries (with estimates unavailable for most lower-income countries). The positive country-level association likely reflects increased stress and decreased coping resources in high-income societies, that offset the objective benefits of higher economic resources. An example of such a deficit is that perceived low social position is more strongly associated with mental disorders in high-income countries than lower-income countries[7]. Consistent with this observation, there is an opposite-sign association between SES and suicide in the US, with the suicide rate negatively associated with personal income but positively associated with mean average area income after controlling for personal income[8].

A growing literature supports the notion that economic development can have negative effects that offset, or even outweigh, the protective effects of higher economic resources on mental health[9]. These same negative effects contribute to lifestyle-related physical disorders such as obesity and its associated sequelae (e.g., cardiometabolic and musculoskeletal disorders). These disorders are more prevalent in high-income than lower-income countries, despite being inversely associated with SES within countries. Why does this opposite-sign pattern exist? Because not all else is equal across levels. Although higher SES supports healthier lifestyles at the individual level, high-income countries create environments

that promote obesity at a scale that overwhelms the benefits of higher income and better health care access.

The same logic applies to mental disorders. Features of modern high-income societies – social disconnection, loneliness, competition, reduced meaning – are known to be associated with increased vulnerability to some mental disorders even though higher SES has a protective effect at the individual level within countries. But much of the research in this area is ambiguous, because these presumed risk factors are reciprocally related to mental disorders at the individual level within countries. Studying their effects systematically at a time-space cross-national level would require data at a scale comparable to that of the cross-national public opinion surveys used to study well-being (i.e., many countries assessed annually) rather than at the scale of psychiatric epidemiological surveys (i.e., a small number of countries assessed once or, in a few cases, twice a decade apart) along with the country-level and regional data on economic and policy variables.

A call has been made for creating parallel national mental wealth observatories, that would implement ongoing surveys to study societal influences through the investigation of time-space variation in mental disorders and make policy recommendations aimed at influencing mental well-being[9]. Any such effort would need to go beyond assessing well-being to include information about clinically significant mental disorders, as time-space variation in prevalence of mental disorders almost certainly differs from variation in well-being. There would be challenges in doing this with the long diagnostic assessments used in epidemiological surveys, but new hybrid methods exist that allow integrating epidemiological assessments of clinically significant mental disorders into the kinds of annual cross-national public opinion surveys that are used to monitor well-being. This hybrid approach needs to be the way of the future, if we are to expand our understanding of societal-level influences on mental disorders.

Ronald C. Kessler[1], Shelby Borowski[1], Nilson N. Mendes Neto[2]
[1]Department of Health Care Policy, Harvard Medical School, Boston, MA, USA; [2]Clinical Investigation, Harvard Medical School, Boston, MA, USA

1. Scott KM, de Jonge P, Stein DJ et al. Mental disorders around the world: facts and figures from the WHO World Mental Health surveys. Cambridge: Cambridge University Press, 2018.
2. Dückers ML, Alisic E, Brewin CR. Br J Psychiatry 2016;209:300-5.
3. Mulder RT, Scott KM, Jorm AF. Aust N Z J Psychiatry 2020;54:664-6.
4. Oparina E, Clark AE, Layard R. The Easterlin paradox at 50. Washington: Brookings Center on Social Dynamics and Policy, 2025.
5. Jorm AF, Mulder RT. Glob Ment Health 2022;9:65-71.
6. Comtesse H, Smid GE, Rummel AM et al. J Affect Disord 2024;350:359-65.
7. Scott KM, Al-Hamzawi AO, Andrade LH et al. JAMA Psychiatry 2014;71:1400-8.
8. Daly MC, Wilson DJ, Johnson NJ. Rev Econ Stat 2013;95:1480-500.
9. Occhipinti JA, Skinner A, Doraiswamy PM et al. Bull World Health Organ 2024;102:323-9.

DOI:10.1002/wps.21375

What can practitioners do about social determinants of mental health?

A vast body of epidemiological and anthropological research shows that a variety of social determinants, from maltreatment in early childhood to an array of adversities in adulthood, are the strongest risk factors for the emergence and maintenance of mental health problems[1,2]. But no mental health practitioner needs to read the epidemiological literature to appreciate this profound relationship. After all, we experience it first-hand through the life histories of virtually every person who seeks our care. One may diagnose a person with depression, but being aware of the intimate partner violence that is a key driver. One may document that a person's psychosis has relapsed because of failure to take prescribed medication, but being aware that the individual's social isolation is a critical factor. One may offer treatment for alcohol use disorder, but being aware that this habit is fuelled by the person's enduring unemployment.

And yet, despite this awareness that simply treating the symptoms is an incomplete and inadequate solution to the mental health condition, most practitioners do just that, by writing out a prescription or offering a psychotherapy. They may feel constrained to directly address social determinants for a variety of reasons, most commonly because these determinants appear to be outside their locus of control; or because they have no idea on how these can be addressed within brief consultation times, and they fear that doing so will unlock a Pandora's box which could overwhelm the clinical process; or, more fundamentally, because they feel that this is not the business of a clinician. The consequences are all too visible: our patients often do not engage in clinical treatments because they perceive them to be superfluous to the drivers that are fuelling their symptoms, which contributes to drop-outs, relapses and chronicity.

The challenge, then, is identifying exactly what a busy practitioner can do when the social determinants seem so complex and remote from the clinic. This is a question which has not been adequately addressed in our research, not least because of the complexity of applying study designs aimed to evaluate clinical treatments to interventions targeting social determinants[3,4]. This is why most treatment guidelines say so little in this respect. While this is a monumental gap in the landscape of evidence-based interventions, which must be prioritized in mental health research, we do not need to wait for research evidence to guide the way. We can use our common sense, in particular when the interventions we might offer do little or no harm.

At the level of individual patients, simply eliciting their illness narrative, in particular their view of what is causing symptoms, is an essential first step, if only because we cannot target an issue we are unaware of. Encouraging patients to engage with reward-

ing and pleasurable activities that they may have withdrawn from, which is the essence of behavioral activation (one of the most cost-effective of all mental health interventions)[5], in particular activities which expose the person to the social world, could be considered[6]. Think of activities which involve engaging with one's community (for example, through volunteering), being in natural surroundings (for example, through walks in parks), or seeking social support (for example, arranging to meet with people who matter to them).

At the level of people in the patient's close social network, we may offer to engage with them to discuss how they might support the recovery process. We may offer to refer the patient to a peer support organization to promote connection with others who had a similar experience. We may be pro-active in leveraging social welfare schemes for persons with disabilities. We may keep updated lists of agencies in our geographical area which address social determinants, such as job centres, domestic violence shelters, and social welfare schemes; establish personal relationships with key contacts in these agencies, and be pro-active in referring patients to these services.

Naturally, we should consider factors such as cultural background, co-occurring impairments, and personal interests when recommending activities or services. One important strategy is to build clinical skills in providers who are already engaged in addressing social determinants (such as community health workers and peer support workers)[7]. This means that practitioners should collaborate with or, if there are none existing, support the creation of such front-line worker programs.

Of course, there are challenges for practitioners to seamlessly collaborate with social sectors. Communication gaps between mental health practitioners, community health workers and social welfare organizations can hinder effective implementation. Some patients may prefer to be involved exclusively in traditional medical treatments. Or, the problems may be so overwhelming that there is a risk of practitioner and patient defeat. Despite these limitations, the risks are small and the benefits often visible through improved patient engagement and better health outcomes.

Addressing social determinants during the clinical encounter can be an empowering experience for the patients, who will feel that care is sensitive to their real-world concerns and makes them active partners in their own recovery journeys. Any additional investment of time that a practitioner makes, or financial resources that a health system provides, would pay for themselves handsomely (which is why a growing number of health systems are facilitating such opportunities for health care practitioners[8]), and would make the practice of mental health care more holistic and interesting than the reductive process that it has become for many. It would help connect the clinic with the community in what could be a mutually reinforcing dynamic which destigmatizes mental health care and makes it more meaningful to its beneficiaries.

Would these practices dilute the mission for psychiatry to be more grounded in targeting the core mechanisms which propel mental health conditions? I think not, since our emphasis on psychopharmacology seems to almost entirely focus on dampening down symptoms. On the other hand, psychotherapies often target an individual's inner world of cognitions and behaviors on the basis of theories, not observable facts, suggesting that those are the mediators of the mental health condition in that particular person. In my mind, not dealing with social determinants in mental health care is equivalent to treating tuberculosis with only antitussive medication.

Indeed, addressing social determinants, far from unmooring psychiatry from its siblings in other medical specialties, would do the exact opposite, by offering a mechanistically informed approach to clinical practice. By fully embracing the need to attend to social determinants, clinical mental health disciplines would finally realize the true potential of the biopsychosocial approach to mental health care, which, while widely touted, remains largely unfulfilled. Moreover, doing so would also align clinical practice with what matters most to patients themselves, moving away from a one-size-fits-all approach predicated on a narrowly defined, diagnosis-driven, clinical treatment model to person-centred care.

This, in turn, would offer an opportunity to bridge the gulf between the clinical outcomes which dominate practice and the recovery outcomes which are championed by persons with the lived experience. Most of all, it would make the experience of mental health care fulfilling for both the patient and the practitioner.

Vikram Patel
Department of Global Health and Social Medicine, Harvard Medical School, Boston, MA, USA

1. Kirkbride JB, Anglin DM, Colman I et al. World Psychiatry 2024;23:58-90.
2. Lund C, Brooke-Sumner C, Baingana F et al. Lancet Psychiatry 2018;5:357-69.
3. Bickerdike L, Booth A, Wilson PM et al. BMJ Open 2017;7:e013384.
4. Husk K, Elston J, Gradinger F et al. Br J Gen Pract 2019;69:6-7.
5. Cuijpers P, van Straten A, Warmerdam L. Clin Psychol Rev 2007;27:318-26.
6. Holt-Lunstad J. World Psychiatry 2024;23:312-32.
7. Patel V, Saxena S, Lund C et al. Lancet 2023;402:656-66.
8. UK National Health System. Social prescribing. www.england.nhs.uk/personalisedcare/social-prescribing.

DOI:10.1002/wps.21337

Beyond symptom improvement: transdiagnostic and disorder-specific ways to assess functional and quality of life outcomes across mental disorders in adults

Christoph U. Correll[1-5], Samuele Cortese[6-10], Marco Solmi[1,6,11-14], Tommaso Boldrini[11,15], Koen Demyttenaere[16], Katharina Domschke[5,17], Paolo Fusar-Poli[18-21], Philip Gorwood[22], Philip D. Harvey[23,24], Richard S.E. Keefe[25], Christine Knaevelsrud[5,26], Roman Kotov[27], Laura Nohr[5,26], Taeho Greg Rhee[28], David Roe[29], Matthias Rose[30,31], Lon S. Schneider[32,33], Mike Slade[34,35], Dan J. Stein[36], Charlene Sunkel[37], Roger S. McIntyre[38,39]

[1]Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin Berlin, Berlin, Germany; [2]Department of Psychiatry, Northwell Health, Zucker Hillside Hospital, Glen Oaks, NY, USA; [3]Department of Psychiatry and Molecular Medicine, Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, USA; [4]Center for Psychiatric Neuroscience, Feinstein Institutes for Medical Research, Manhasset, NY, USA; [5]DZPG, German Center for Mental Health, partner site Berlin, Berlin, Germany; [6]Centre for Innovation in Mental Health, School of Psychology, Faculty of Environmental and Life Sciences, University of Southampton, Southampton, UK; [7]Clinical and Experimental Sciences (CNS and Psychiatry), Faculty of Medicine, University of Southampton, Southampton, UK; [8]Solent NHS Trust, Southampton, UK; [9]Hassenfeld Children's Hospital at NYU Langone, New York University Child Study Center, New York, NY, USA; [10]Department of Precision and Regenerative Medicine and Ionian Area, University of Bari "Aldo Moro", Bari, Italy; [11]Department of Psychiatry, University of Ottawa, Ottawa, ON, Canada; [12]Regional Centre for Treatment of Eating Disorders, and On Track: Champlain First Episode Psychosis Program, Department of Mental Health, Ottawa Hospital, Ottawa, ON, Canada; [13]Clinical Epidemiology Program, Ottawa Hospital Research Institute, University of Ottawa, Ottawa, ON, Canada; [14]School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada; [15]Department of Psychology and Educational Science, Pegaso Telematic University, Naples, Italy; [16]University Psychiatric Center KU Leuven, Department of Neurosciences, Faculty of Medicine, Leuven, Belgium; [17]Department of Psychiatry and Psychotherapy, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany; [18]Early Psychosis: Interventions and Clinical-detection (EPIC) Lab, Department of Psychosis Studies, King's College London, London, UK; [19]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy; [20]Outreach and Support in South-London (OASIS) service, South London and Maudsley NHS Foundation Trust, London, UK; [21]Department of Psychiatry and Psychotherapy, Ludwig-Maximilian-University, Munich, Germany; [22]Université Paris Cité, INSERM U1266, Paris, France; [23]Department of Psychiatry and Behavioral Sciences, Miller School of Medicine, University of Miami, Miami, FL, USA; [24]Research Service, Miami VA Healthcare System, Miami, FL, USA; [25]Department of Psychiatry, Duke University Medical Center, Durham, NC, USA; [26]Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany; [27]Department of Psychiatry, Stony Brook University, Stony Brook, NY, USA; [28]Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA; [29]Department of Community Mental Health, Faculty of Social Welfare and Health Sciences, University of Haifa, Haifa, Israel; [30]Department of Psychosomatic Medicine, Charité Universitätsmedizin Berlin, Berlin, Germany; [31]Center for Patient-Centred Outcome Research, Charité Universitätsmedizin Berlin, Berlin, Germany; [32]Department of Psychiatry and Behavioral Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; [33]Department of Neurology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; [34]School of Health Sciences, University of Nottingham, Nottingham, UK; [35]Faculty of Nursing and Health Sciences, Health and Community Participation Division, Nord University, Namsos, Norway; [36]SA Medical Research Council Unit on Risk and Resilience in Mental Disorders, Department of Psychiatry and Neuroscience Institute, University of Cape Town, Cape Town, South Africa; [37]Global Mental Health Peer Network, Paarl, Western Cape, South Africa; [38]Department of Psychiatry, University of Toronto, Toronto, ON, Canada; [39]Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, Canada

Improving meaningful outcomes is the main goal of clinical care for mental disorders. Traditionally, the focus in clinical research and practice has been on outcome domains that refer to symptom severity or service use (e.g., hospitalization), relate to categorical diagnoses, and favour clinician-rated measures. More recently, self-rated and dimensional as well as transdiagnostic outcome domains have gained traction, and functioning, quality of life and well-being/life satisfaction, along with the construct of personal recovery, have become a stronger focus. These key multidimensional outcome domains need to be properly defined and assessed. Further, the concepts of "functional" and "personal" recovery need to be differentiated. "Functional recovery" is defined by observed functioning across the domains of self-care, social interactions, leisure time activities, and educational or vocational activities. "Personal recovery" involves the subjective sense of living a personally meaningful life, irrespective of whether symptoms continue, or ongoing/intermittent support is needed. Despite the multi-stakeholder relevance of these outcome domains, no comprehensive account of how to measure them is available. To fill this gap, we provide here an overview of the main tools to assess functioning, quality of life/well-being/life satisfaction, and personal recovery outcomes across mental disorders in adults, aiming to also identify additional needs that should be addressed. We identified tools that can be used in clinical and research practice to assess people with the following mental health conditions: anxiety disorders, bipolar disorder, dementias, eating disorders, major depressive disorder, obsessive-compulsive and related disorders, personality disorders, post-traumatic stress disorder, schizophrenia, and substance use disorders. Both transdiagnostic and disorder-specific measures are described. Suggested tools were selected keeping feasibility and scalability needs in mind. The incorporation of these measures in both research and clinical care will enrich patient assessment as well as treatment planning and evaluation, increasing the likelihood of enhanced outcomes in people living with mental disorders.

Key words: Outcome domains, functioning, quality of life, well-being, life satisfaction, personal recovery, transdiagnostic measures, measurement-based care

(World Psychiatry 2025;24:296–318)

Improving meaningful outcomes represents the main goal of clinical care and mental health service provision[1]. The choice of the outcome domains to be assessed often reflects assumptions about what is deemed to be relevant, which raises questions as to who determines this relevance, and how. Related and equally important questions are why, in what time frame and setting, and for whom is the outcome relevant. For example, hierarchies of outcome domains differ in an acute illness exacerbation from times of symptomatic stabilization or from living with a long-term condition. Also, the importance attached to particular outcome domains may differ between the person with a mental health condition, his/her family members, care partners and peers, clinical and professional groups, cultural subgroups, and society at large. Moreover, measurement of outcome domains can be made by clinicians, caregivers who know the person with a mental health condition, and/or the person him/herself.

Categorical diagnoses based on symptoms, as used in current ICD-11 and DSM-5-TR manuals, have been instrumental in mov-

ing the field from a non-standardized art-form of assessment linked to nuanced interpretations of symptom meaning, to a more technical and structured process based on data gathered from systematic observations. However, the relevance given to symptoms grouped in categorical diagnoses has been a matter of intense debate and criticism by researchers and clinicians in the field, as well as by persons living with a mental health condition[2].

The focus in clinical research has traditionally been on the severity of clinical symptomatology of categorical disorders and on clinician-rated outcomes, as well as on service use data, such as (re)-admission. Indeed, in people with mental disorders, symptom severity is the most commonly assessed outcome in randomized controlled trials of pharmacological as well as non-pharmacological interventions. However, concerns have been expressed that symptom-based diagnoses and treatment goals may shift the focus away from the outcome domains that are prioritized by people receiving mental health interventions[3].

Critics of the symptom-based categorical diagnoses have also expressed concerns that the mere enumeration of symptoms "produces negative value judgment, promotes conformity and has no meaning for treatment"[4] and, ultimately, may risk dehumanizing the patient[5]. It has also been pointed out that the strict application of a categorical diagnostic approach may result in individuals with significant symptoms and/or impairments, but who fall short of the diagnostic criteria, being denied support and treatment[6].

These concerns and a strengthening of mental health service user involvement have contributed to shifting the interest towards outcome domains beyond symptom severity, better fitting people's reality[7]. The focus on outcome assessment beyond symptoms has moved in particular to functioning as an observable phenomenon and to the personal evaluation of functional status, which can be captured by the constructs of quality of life, well-being and life satisfaction[8]. Furthermore, personal recovery, as a subjectively experienced and evaluated state that incorporates a person's individual values and preferences, has become an important additional domain[9].

The term "functioning", as used starting from the DSM-III, refers to the observable ability of an individual to carry out self-care and daily life activities, and to be involved in interpersonal relationships and educational/vocational activities[10]. The concept of functional impairment indicates the limitations in the personal, interpersonal and societal functioning due to the illness, the environment and/or treatment adverse effects. This concept is related to the term "disability" in the World Health Organization (WHO)'s International Classification of Functioning, Disability and Health (ICF), where decrements in the individual's functioning are termed "impairments" at the body level, "activity limitations" at the person level, and "participation restrictions" at the societal level[11].

The related concept of functional recovery refers to a continuum where some symptoms may persist but do not severely impact daily functioning, similar to what may happen in chronic physical conditions such as asthma[12]. Harvey and Bellack[13] suggested to evaluate functional recovery by looking at improvements in daily functioning across domains such as independent living, work, and social relationships, maintained for at least six months. This broader approach emphasizes practical aspects of life and aligns mental health interventions with the true needs and aspirations of individuals.

The concept of "quality of life" originated in oncology more than 30 years ago, in relation to the impact of symptoms of illness as well as treatment side effects[14]. Using this concept as a foundation for outcome measurement presents a challenge, due to its potential for numerous definitions and measurements[15,16]. Various models of quality of life exist, encompassing observable and subjective indicators, needs satisfaction, psychological and subjective well-being models, as well as health, functioning, and social models[17]. A persistent debate revolves around the question of whether a measure should lean towards an observable or subjective orientation. An approach focused on subjectivity may be centred on the experience of immediate happiness or pleasure, or broader aspects such as self-fulfilment, realization or actualization[18]. Along those lines, the so-called subjective or, better, evaluation-based quality of life has been referred to as the ability to satisfy one's needs – physical, emotional and social – which is a personal cognitive-emotional construct mediating between observable indices (e.g., living conditions) and personal expectations and aspirations. Accordingly, the WHO defines quality of life as an individual's perception of his/her position in life in the context of the culture and value systems in which he/she lives and in relation to his/her goals, expectations, standards and concerns[15].

Overall, while some aspects of the above conceptualizations have blurred borders and do overlap in some instances, the shift from symptoms to observable functioning and self-rated quality of life is consistent with the well-known definition of mental health from the WHO as "a state of well-being in which the individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to his or her community"[19].

The concept of well-being is very close to (and basically is synonymous with) that of "flourishing", which is a combination of feeling good (the *hedonic* component) and functioning meaningfully (the *eudaimonic* component)[20]. High levels of well-being have been shown to be associated with a range of positive outcomes, including effective learning, productivity and creativity, good relationships, pro-social behaviors, good health, and greater life expectancy[21]. Well-being is usually regarded as a multi-dimensional construct, in contrast with the long-standing assumption that positive human experience can be adequately assessed using a single item about life satisfaction or happiness[20]. Indeed, there is evidence that multi-dimensional measures of well-being correlate only moderately with standard life satisfaction questions[22].

More recently, the concept of "personal recovery" has gained central importance. This goes beyond not only the focus on symptom improvement, but also that on performance-based or self-rated functioning. Personal recovery has been defined as "a deeply personal, unique process of changing one's attitudes, values, feelings, goals, skills, and/or roles" and "a way of living a satisfying, hopeful, and contributing life even within the limitations caused by illness"[23,24]. A systematic review identified five personal recovery processes: connectedness, hope and optimism about the future,

identity, meaning in life, and empowerment (CHIME), adding that cultural sensitivity (stigma, religiosity factors) are also relevant to consider[25].

The transition towards supporting personal recovery involves an increased emphasis on the promotion of autonomy, patient choice and, importantly, personally identified rather than "system-imposed" goals, including meaning and purpose and a sense of agency, also recently termed "life engagement"[26-28], irrespective of whether symptom persist and people continue to need help and support[23]. Specific measures related to subdomains of personal recovery have been developed for empowerment[29] and hope[30].

In addition to the crucial choice regarding the type of outcome to be assessed – symptom-based, performance-based (functional) or appraisal-based (quality of life/well-being, personal recovery) – there is an ongoing debate about whose perspective is used in assessing the outcome. It has been argued that the patient's perspective should be central, rather than the perspective of the clinician or informal carer[31-33]. To stress this position, the term patient-reported outcome or patient-rated outcome (PRO) was introduced more than a decade ago. The most recognized definition comes from the US Food and Drug Administration (FDA), which defines PRO as an umbrella term for all patient self-reported health information[34]. Thus, the term also includes the use of self-reports for screening purposes, and for the monitoring of symptoms, well-being or quality of life during the treatment process or follow-up.

Patient-rated measures encompass assessment of the experience of using mental health services and systems (patient-rated experience measures, PREMs) and assessment of health gain (patient-rated outcome measures, PROMs), including the important patient-generated PROMs (PGPROMs)[35-37]. Currently available PROMs include both clinical and recovery outcome domains [35], thus being more likely to capture also the content most relevant to people with mental health conditions.

Another relevant issue is the use of digital technologies[38,39]. For example, passive sensing of performance measures related to functioning (e.g., activity, geolocation, social media and e-mail contacts)[40] can provide a more accurate and continuous evaluation that takes place in the person's real-world environment, increasing the external validity of the assessment. Moreover, the experience sampling method (ESM)[41,42] can further be used to gather information on subjective experiences of people with mental health conditions via a collection of self-reports on activities, emotions or other data related to daily life at various points throughout the day, preferably randomly timed to avoid behavioral adaptation to fixed intervals of questioning. Ecological momentary assessment[43] is a subtype of ESM based on high-frequency, real-time collection of data. While the use of digital technologies provides new and exciting opportunities for mental health assessment, this area has been reviewed elsewhere[38,44]. Thus, we will not consider this aspect in more detail in this paper.

Overall, the shift towards outcome domains beyond symptom severity in mental health is promising on several fronts. First, it has been recently highlighted that this expansion of assessment domains may actually allow us to better understand the symptoms themselves. Indeed, it has been pointed out that "the way these core phenomena are perceived, elaborated and verbalized by the affected person likely depends upon how that person generally functions and appraises her functioning (e.g., how rich and articulated her cognitive life is, or how much she is focused on her body and its functioning), upon the influence of the cultural environment in which she is immersed, and upon the pattern of predisposing and precipitating factors at work in that individual case"[45].

Second, this broadening shift in outcome domains and their assessments could ultimately contribute to the success of precision psychiatry, i.e. the ability to stratify or individualize treatment according to specific characteristics of subgroups or individual people with mental conditions[46-48]. Indeed, studies using machine learning have shown that non-symptom variables may contribute to the identification of people who are likely to respond to a given antidepressant drug[49,50]. In this regard, variables related to functioning and quality of life/well-being have been proposed as relevant in the fields of depression[51], bipolar disorder[52] and anxiety disorders[53].

Third, the move beyond symptoms resonates with efforts to strengthen the assessment of mental health beyond specialist care, in particular in primary care, which should not limit its assessment paradigm to symptomatic pathology or disordered behavioral responses, but also help each individual cope, and if possible thrive, within his/her context, managing personal limitations while building on strengths[54].

To the best of our knowledge, a comprehensive account of available measures of observable functioning, quality of life, well-being and personal recovery outcome domains is not available. To fill this gap, we aimed to provide an overview of the main tools to assess these outcome domains across mental disorders in adults.

We identified both transdiagnostic and disorder-specific scales, questionnaires and interviews targeting these domains by searching in PubMed for the name of the disorder(s) together with "psychosocial" or "function*", or "recovery" or "quality of life" or "well-being", and "assessment" or "tool" or "instrument" or "interview". This approach was supplemented by a hand search of measures known to the authors from their own work and from the literature. Ultimately, we included in this report those measures that were deemed most relevant and practical for research and clinical purposes by expert consensus.

We excluded from this review direct performance-based assessments of functioning, as these are lengthy or laboratory-based research tools and would not be applicable in clinical care.

TRANSDIAGNOSTIC FUNCTIONAL ASSESSMENT TOOLS

Transdiagnostic measures capturing real-world functional outcome domains can be used in routine clinical practice without posing a significant burden on the system, providers or people with mental health conditions. These measures can be completed by the patient, the clinician and/or a formal or informal carer.

There is a long tradition for the assessment of self-reported health perceptions independent of the underlying condition. The

first studies were published half a century ago. The transdiagnostic tools that we deemed most relevant are presented below in order of historical development.

The *Global Assessment Scale (GAS)*, developed by Endicott et al in 1976[55], is a clinician-administered tool that assesses global functioning through two domains: symptom severity and psychosocial functioning. It is a very simple 1-100-point scale that generates one number that can be scored quickly. Since the final score is determined by the more severely impacted domain, one does not know which of the two domains is responsible for the score, and what the score of the less severe domain would have been. However, in the case of the categorical symptom-function recovery dyad, this drawback is less relevant, as a score of ≥71 would reflect borderline severity of both illness symptoms ("if symptoms are present, they are transient and expectable reactions to psychosocial stressors – e.g., difficulty concentrating after family argument") and functioning ("no more than slight impairment in social, work, or school functioning – e.g. temporarily falling behind in school or work")[55].

The GAS was then modified into the *Global Assessment of Functioning scale (GAF)*, which was first included in the DSM-III-R[56] and then in the DSM-IV[57]. The GAF includes two scores: the GAF-Symptoms (GAF-S) and the GAF-Functioning (GAS-F). Nevertheless, one will still not know which subdomains of functioning (self-care, social interactions, leisure time, school/work) are responsible for the ultimately lowest score and, most importantly, whether all domains were really asked about or considered (equally) in the scoring. Therefore, the GAF has been removed from the DSM-5, after having been a core component of the Axis 5 assessment in the DSM-IV.

The GAF has been followed by the *Social and Occupational Functioning Assessment scale (SOFAS)*[58,59], which is more functionally oriented, having two subdomains (social and occupational). This scale has good concurrent and predictive validity[59]. It is simple and quick to administer (1 min), but, as the GAF, has been criticized for being contaminated by clinical symptom severity[60].

The *Sheehan Disability Scale (SDS)*[61], developed in 1983, is a patient-reported measure. It includes three items regarding work/school, social life, and family life/home responsibilities, which are used the most, along with two additional items on "days lost" and "days unproductive". It uses an 11-point analogue scale and takes less than 5 min to complete. It is licensed and has demonstrated good reliability and validity[62].

The *Specific Level of Functioning (SLOF)*[63] is a 43-item clinician-, patient- or informant-rated scale exploring the following subdomains: physical functioning, personal care skills, interpersonal relationships, social acceptability, activities, and work skills. It demonstrates strong reliability and validity[64]. Additionally, it is sensitive to changes over time, making it a valuable tool for tracking progress and treatment outcomes in clinical settings[64].

The Medical Outcomes Study (MOS)[65] from 1989 can be considered the landmark investigation moving the field forward. One main contribution was the identification of eight health domains: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health. The instrument resulting from this study was the *MOS 36-item Short-Form Health Survey (SF-36)*[66], which is now the most used PROM worldwide, with more than 25,000 papers listed in PubMed. This tool includes 36 items scored on a mixed Likert and dichotomous scale, with physical and mental health component scores (PCS and MCS). It takes 10-15 min to complete and is available for free with a user agreement. It has demonstrated high reliability, internal consistency and convergent validity[67].

The *MOS 12-item Short-Form Health Survey (SF-12)*[68], published in 1996, is a shorter version of the SF-36, measuring the same domains of physical and mental health. It includes 12 items scored similarly to the SF-36, taking about 3 min to complete. It is licensed and has high reliability[69]. As the SF-36, it is accepted by the FDA and the European Medicines Agency (EMA).

In 1999, M. Weissman developed the *Social Adjustment Scale - Self-Report (SAS-SR)*[70], a patient-reported tool measuring psychosocial functioning across six domains: work, social and leisure activities, relationships with extended family, role as a marital partner, parental role, and role within the family unit. The scale is available in three versions: a 54-item full-length, a 24-item short, and a 14-item screening version. It is scored on a Likert scale from 1 to 5, and takes between 5 and 20 min to complete, depending on the version. It has demonstrated good reliability[71]. Validity studies have shown significant correlations of the short and screening versions with various measures of psychosocial functioning[72].

The *Global Functioning (GF) scale*[73,74], published in the early 2000s, is a clinician-rated instrument measuring two areas: "role" (performance and amount of support needed in a specific role) and "social" (focusing on quantity and quality of peer relationships, level of peer conflict, age-appropriate intimate relationships, and involvement with family members). Its administration takes about 5 min, and it has shown good psychometric properties[75,76].

The *Personal and Social Performance scale (PSP)*[77], also developed in the early 2000s, is a 4-item clinician-rated tool (administration time: 5 min) exploring four areas: socially useful activities, personal and social relationships, self-care, and disturbing and aggressive behaviors. It has shown good reliability and validity[77].

The *Work and Social Adjustment Scale (WSAS)*[78], published in 2002, is a patient-rated scale with five items assessing function in terms of work, home management, social leisure activities, private leisure activities, and interpersonal relationships. Its administration time is about 5 min. It has been shown to have good reliability[79].

The *Functional Assessment Short Test (FAST)*[80], developed in 2007, is a clinician-administered tool designed to assess psychosocial functioning across six domains: autonomy, occupation, cognition, financial issues, interpersonal relationships, and leisure time. It comprises 24 items scored on a 4-point Likert scale, taking about 5-10 min to complete. It is freely available and demonstrates strong psychometric properties, with excellent test-retest reliability and good sensitivity and specificity[80].

In 2004, the US National Institutes of Health jointly funded an initiative aimed to provide a measurement framework for key health domains relevant to both mental and physical disorders. The result was a highly comprehensive *Patient-Reported Outcome Measurement Information System (PROMIS)*[81], supported by the

Patient-Centered Outcomes Research Institute (PCORI)[82] since 2009. With more than 3,600 studies published to date, PROMIS scales are increasingly used, and start to be recommended by national health authorities, e.g. in Germany[83] and the Netherlands[84].

The PROMIS utilizes methods derived from item-response theory to provide item banks, which allow estimating a scale score based on various combinations of items[85]. In clinical practice, instruments with fewer items are preferred, whereas study settings favor more precise, mostly longer tools. PROMIS item banks enable the use of computer adaptive tests[86], which tailor the assessment to the individual patient responses. By administering only informative items, measurement precision can be increased, and response time and burden decreased.

Currently, the PROMIS provides item banks for >90 health domains, and for most of them investigators have suggested some item combinations, i.e., short forms similar to conventional assessment measures. The most commonly used short form is the *PROMIS-29 Profile*[87], which combines the rating of seven health domains (physical function, fatigue, sleep disturbance, pain interference, depression, anxiety, social participation), normed to a representative general population for intuitive score interpretation on a 50 (mean)/10 (standard deviation) metric. The PROMIS measures have shown good psychometric properties in patients with various health conditions[88].

In 2003, the WHO initiated the development of tools for the assessment of key health domains. The *World Health Organization Disability Assessment Schedule II (WHODAS 2.0)*[89], developed in 2010 and translated into many languages, focuses on the assessment of physical and social components, and thus can be combined with some of the instruments commonly recommended for the assessment of specific mental health symptoms with minimal overlap. The tool can be used by both clinicians and patients to assess six domains: cognition, mobility, self-care, relational, life activities, and participation. It includes 36 items (with a 12-item short form) scored on a 5-point Likert scale. The 36-item version takes about 20 min to complete, while the 12-item version takes about 5 min. The WHODAS 2.0 is available for free with a user agreement and has shown high test-retest reliability and internal consistency[90-92].

FUNCTIONAL ASSESSMENT IN SPECIFIC MENTAL DISORDERS

We now highlight selected tools to assess functioning in individuals with specific mental disorders or groups of related disorders. We first discuss the use of transdiagnostic tools, followed, when available, by specific tools designed to evaluate functioning in these conditions. Specific disorders are listed in alphabetical order.

Anxiety disorders

Anxiety disorders affect the individual not just through the presence of distressing thoughts, emotions and behaviors, but also by their effects on the ability to perform daily tasks, maintain relationships, and engage in work or social activities[93]. These impacts vary widely among individuals, and may not be fully reflected in standard assessment tools measuring core symptoms of anxiety[93]. Therefore, a comprehensive evaluation – with respect to diagnostic assessment and treatment response monitoring – should include both measures of core symptoms and functional assessments, recognizing the intricate ways in which anxiety can disrupt daily life[94].

The WHODAS 2.0 has been explicitly recommended by the International Consortium for Health Outcomes Measurement (ICHOM) Depression and Anxiety Working Group as an outcome domain in anxiety disorders, especially to monitor treatment progress[95]. The Canadian guidelines for the management of anxiety disorders[96] suggest integrating the evaluation of functional impairment measured by the SDS or the SF-36 into the definition of treatment response. In the EMA guideline on the clinical investigation of medicinal products indicated for generalized anxiety disorder, panic disorder and social anxiety disorder, the use of the SDS and quality of life measures is recommended as secondary outcome parameters[97,98]. The UK National Institute for Health and Care Excellence (NICE) guidelines for the management of generalized anxiety disorder/panic disorder and social anxiety disorder also point to the necessity of including measures of functioning and quality of life as outcome parameters[99,100].

We highlight here two disorder-specific tools for the assessment of functioning in anxiety disorders. The *Disability Profile (DP)*[101], developed in 1994, is a clinician-administered tool designed to evaluate psychosocial functioning in individuals with anxiety disorders. It covers eight domains (school, work, family, marriage/dating, friendships, other interests, activities of daily living, and suicidal behavior) to be scored on a 5-point Likert scale, taking approximately 5 min to complete. It is available for free and demonstrates strong psychometric properties[101].

The *Liebowitz Self-Rated Disability Scale (LSRDS)*[101], also published in 1994, is a patient-reported instrument that assesses psychosocial functioning and dysfunctional behaviors across eleven domains (alcohol abuse, drug abuse, mood dysregulation, education, career, family relationships, romantic relationships, friendships, hobbies, activities of daily living, and suicidality), to be scored on a 4-point Likert scale, taking less than 5 min to complete. It is freely available and has excellent psychometric properties, for both current and lifetime assessments, being validated especially for social anxiety disorder[101].

Functional recovery in anxiety disorders could be defined as a total score ≤6 on the first three items of the SDS plus a score ≤2 on each of those items, persisting for at least one year[102,103].

Bipolar disorder

Bipolar disorder is associated with high rates of functional impairment. The burden of illness at the individual and population level is largely mediated by impairment in functional outcome domains[104]. For example, impairment in workplace attendance and performance differentially contributes to the overall cost of

illness[52]. It is also well established that functional recovery in bipolar disorder lags behind symptomatic recovery, and that persons who are in symptomatic recovery and still evidencing functional impairment are at higher risk of symptomatic relapse and recurrence[52]. Moreover, people with lived experience of the disorder assign priority to improvement in psychosocial functioning as a critical therapeutic objective[52].

Thus, measurement-based care in bipolar disorder should include the assessment of functional outcome domains in addition to symptomatic ones. Indeed, assessing functioning is a guiding principle commonly emphasized across clinical practice guidelines on this disorder[105,106].

Interestingly, the degree of mood instability (including subclinical symptoms) has been shown to be at least as predictive of functional recovery as the number of full-blown episodes[107]. Furthermore, the various domains of functional recovery seem to be affected by different variables, with social recovery being better predicted by clinical variables, and occupational recovery by sociodemographic ones[108].

Two transdiagnostic functional measures, the FAST and the SDS, can be considered in patients with bipolar disorder. Both of them have shown good psychometric properties in these patients[80,109], as well as ability to detect change over time[110]. No disorder-specific functional measure can currently be proposed for use in this condition.

Functional recovery in bipolar disorder could be defined as a total score ≤11 on the FAST, or a score ≤2 on each of the first three items of the SDS, yielding a total score ≤6. These criteria should be met for at least one year. Identifying and prioritizing therapeutic objectives with respect to functional outcomes should be collaborative between provider and patient, and adapted to the individual person and illness characteristics. For example, for individuals with chronic and treatment-resistant bipolar disorder, recalibrating functional outcome objectives will be required.

Dementias

The assessment of functioning in people with mild cognitive impairment and different types of dementia has recently gained traction. Since dementia involves impaired insight or cognitive ability to recall and adequately describe functional and psychosocial performance, informal caregiver and clinician ratings are generally given preference over patient self-report.

Two frequently used functional assessment measures are the *Functional Activities Questionnaire (FAQ)*[111] and the *Disability Assessment for Dementia (DAD)*[112].

The FAQ, developed in 1982, evaluates the current level of performance in daily tasks, such as handling personal finance, shopping alone, remembering arrangements, using transport, and preparing a meal. It consists of 10 items scored from 0 to 3, taking about 15 min to complete, and is freely available. It shows high inter-rater reliability and validity, with good sensitivity and specificity in distinguishing levels of functional impairment[111].

The DAD, published in 1999, measures cognitive processes related to activities of daily living (ADLs) in patients with Alzheimer's disease or other dementias. It consists of 40 items covering both basic and instrumental ADLs, scored dichotomously and normalized to a 0-100 scale. Basic ADLs include hygiene, dressing, continence and eating. Instrumental ADLs include meal preparation, telephoning, going on an outing, finance and correspondence, medications, and leisure and housework. The scale takes about 15-20 min to complete and is freely available. It shows excellent test-retest reliability and validity[112].

Three additional rating scales of ADLs in patients with dementias are the *Alzheimer's Disease Cooperative Study - Activities of Daily Living Inventory (ADCS-ADL)*[113], the *Amsterdam Instrumental Activities of Daily Living Questionnaire (A-IADL-Q)*[114,115], and the *Amsterdam Instrumental Activities of Daily Living Questionnaire - Short Version (A-IADL-Q-SV)*[116].

The ADCS-ADL, developed in 1997, evaluates competence in both basic and instrumental ADLs in patients with Alzheimer's disease or mild cognitive impairment. It includes 45 items scored on a mixed scale, taking 30-45 min to complete, and is freely available. It demonstrates good test-retest reliability and significant correlations with cognitive measures[113].

The A-IADL-Q, published in 2013, assesses instrumental ADLs in the early stages of dementia. It features 47-70 items scored on a 5-point scale based on item response theory, taking about 23 min to complete. It is freely available and shows high test-retest reliability and significant correlations with other dementia measures[114,115].

The A-IADL-Q-SV, developed in 2017, is a shorter version of the A-IADL-Q, containing 30 items scored on a 5-point scale. It takes about 10 min to complete and is freely available. It has strong reliability and validity, with good construct validation against other cognitive and functional measures[116].

In regulatory trials for the approval of pharmacological treatments of cognitive symptoms or associated clinical features (such as behavioral dysregulation), functional assessments are generally co-primary or main secondary outcome measures. The most frequently used rating scale for functional performance in this context is the ADCS-ADL[117].

Eating disorders

In eating disorders, functioning has been defined in various ways, and its assessment poses specific challenges[118]. Notably, different from other mental disorders, the DSM-5 definition of these disorders does not include a criterion for functional impairment. In fact, although people diagnosed with eating disorders may revolve a lot of time around eating and body shape, and may have serious physical problems, their symptoms are mostly ego-syntonic, and functioning in several academic or professional activities is often preserved, at least compared to general expectations[118]. Nevertheless, in more severe clinical presentations, the impairment in functioning can be pervasive, extending to cognition and personal ADLs, due to low body weight or physical complications[119].

Among transdiagnostic tools, the WHODAS 2.0 has been recommended to assess functional impairment in persons with an

eating disorder by international expert consensus[120]. However, the SF-36 and SF-12 have been used more frequently in these disorders, and have been often adopted as an external validator for other tools[121].

In terms of disorder-specific tools, the *Clinical Impairment Assessment 3.0 (CIA)*[122] is the gold standard to measure functioning in eating disorders, as recommended by various policy-making stakeholders across continents, including the American Psychiatric Association (APA)[123]. The CIA measures impairment in functioning broadly, capturing social, cognitive and personal impairment due to eating disorder symptoms, and is designed to be administered after a questionnaire measuring symptoms[122]. Admission CIA scores have been found to be correlated with those on the Eating Disorder Examination Questionnaire (EDE-Q)[124], a self-report questionnaire used to assess the range and severity of eating disorder symptoms. Changes in CIA scores from admission to discharge were also positively correlated with changes in EDE-Q. Moreover, patients with lower admission CIA scores were more likely to be classified as "recovered" at discharge[122].

The CIA probably best captures functional impairment due to eating disorders. On the other hand, the WHODAS 2.0 and the SF-36/12 are able to also account for comorbid conditions, such as depression or anxiety. The use of one of these instruments should thus be tailored to each case, accounting in particular for comorbidities.

Recently, an eating disorder-specific questionnaire with a focus on recovery has been introduced, the *Eating Disorders Recovery Questionnaire (EDRQ)*[125]. This is a 28-item tool with four subscales: lack of symptomatic behavior, acceptance of self and body, social and emotional connection, and physical health. It has excellent psychometric properties and goes beyond merely body weight and symptoms[125]. However, its implementation has been limited so far, and more research is needed.

Functional recovery in eating disorders could be defined as a total score ≤16 on the CIA, persisting for at least one year.

Major depressive disorder

In addition to emotional, cognitive and behavioral symptoms, functional impairment is widely associated with major depressive disorder (MDD), representing an important target for interventions[51]. Out of 80 outcome domains for depression identified by a large group of patients, carers and health professionals, 16 were related to functioning[126]. Moreover, symptomatically remitted patients often still have significant functional limitations[127,128]. However, a systematic review has found that functional domains are considered only in a very small minority of trials of pharmacotherapies and psychotherapies[129].

The DSM-IV proposed the use of the GAF for the assessment of functioning in MDD, but this tool has shown inadequate interrater reliability (r=0.26) and poor discriminant validity in outpatients with the disorder[130].

Among other transdiagnostic measures of functioning, the SF-36 has been used in over a thousand depression studies, while the SDS is cited in over 600 PubMed publications. The WHODAS 2.0 has also been successfully used in patients with MDD[131], although too little focus on employment has been criticized[132]. The WSAS has been validated in depression, with a Cronbach's alpha ranging from 0.807 (screening) to 0.942 (week 10)[78].

If functioning in people with MDD is to be assessed more systematically, it is important that the results obtained are comparable, hence the need to use similar assessment measures. Consensus may be easier to reach for measures that are already widely used, assess different and relevant domains, have high reliability and validity, and can also be used in other disorders to facilitate comparisons. No functional measure is at present universally agreed upon as the preferred one in MDD.

Functional recovery in MDD could be defined as a score ≤2 on each of the first three items of the SDS, yielding a total score ≤6, persisting for at least one year. A difference of 2.8 on the SDS between active treatment and placebo has been suggested to be clinically relevant in relation to functional improvement[133]. In difficult-to-treat depression, more modest outcomes with respect to functioning may be satisfactory[134,135].

Obsessive-compulsive and related disorders

Assessing functional aspects beyond core symptom severity in people with obsessive-compulsive disorder (OCD) is crucial for a comprehensive understanding of the disorder's impact on daily life. Although functional impairments – such as difficulties in social interactions, work, and academic performance – may improve when core symptoms are managed[136], a broader assessment is useful in tailoring treatment plans and supporting the individual's overall well-being. However, challenges in this approach include the subjective nature of functional assessments, the variability in daily functioning, and the influence of co-occurring conditions. Additionally, standardized tools may not capture the nuanced experiences of all individuals, necessitating a more personalized evaluation strategy[136].

Some clinical guidelines for the treatment of OCD and related disorders have emphasized that it is useful to assess both symptomatic response and functional recovery. Thus, for example, the Canadian guidelines for the management of anxiety disorders also addressed OCD and noted that recovery from the illness should be defined as loss of diagnostic status, a low score on a disorder-specific measure, and no functional impairment[96]. Similarly, the APA Practice Guideline for OCD and the EMA guideline on the clinical investigation of medicinal products for OCD note the value of employing measures of observable functioning and subjective quality of life to assess treatment outcome[137,138].

Functional impairment in OCD and related disorders can be assessed by the SDS, which has shown improvements on most dimensions in trials of selective serotonin reuptake inhibitors (SSRIs)[139], or the WSAS, which has been validated for use in OCD, with a test-retest correlation of 0.73 and a Cronbach's alpha ranging from 0.789 (screening) to 0.882 (week 10)[78]. Further, the ICHOM Depression and Anxiety Working Group included OCD in its ex-

pert consensus on measures, and recommended the use of the WHODAS 2.0[95]. The development of specific functional impairment tools for OCD and related disorders could be useful.

Functional recovery in OCD could be defined as a total score ≤6 on the first three items of the SDS plus a score ≤2 on each of these items, persisting for at least one year[102,103].

Personality disorders

Although the focus on functioning in people with personality disorders has increased only recently, functional disability and, especially, interpersonal dysfunction are at the core of each of these disorders[140]. Clinical guidelines, mainly concerning borderline and antisocial personality disorders, indicate the need to assess outcome domains beyond symptom improvement, but do not provide guidance on specific tools to measure them.

The NICE guideline for borderline personality disorder[141] recommends utilizing outcome measures of functioning that are relevant to users of services and families/caregivers. Improving role functioning by reaching long-term educational and employment goals is further advised for both borderline and antisocial personality disorders, with the proportion of individuals in contact with secondary mental health services who are in paid employment as an outcome parameter[142]. Similarly, the Global Alliance for Prevention and Early Intervention for Borderline Personality Disorder recommends to fully quantify the educational, vocational and social outcome domains for young people with this disorder[143].

The transdiagnostic tool mostly used in personality disorders is the GAF, with validation for internal consistency and sensitivity to change in clinical trials and cohort studies of borderline personality disorder[144-146]. The WSAS has also shown internal reliability, with a Cronbach's alpha of 0.79-0.90 and an inter-item correlation of 0.43-0.65, in personality disorders[147].

In terms of disorder-specific tools, the *Shedler-Westen Assessment Procedure (SWAP-200, SWAP-II, SWAP-200-A; SWAP-II-A)*[148,149] is a 200-item tool including an overall measure of personality functioning (the *High Functioning Scale*) which takes up to 45 min for administration. It has been validated in studies including personality disorders, with good test-retest and inter-rater reliability[150-152], and is used in both clinical and research settings.

The Criterion A of the *Level of Personality Functioning Scale (LPFS)* from the DSM-5 Alternative Model for Personality Disorders[153] requires the clinician to select the level that most closely captures the individual's current overall level of impairment in personality functioning (i.e., self and interpersonal core functional impairments). The rating is aimed for diagnosing a personality disorder; however, it can also be utilized as a general indicator of personality functioning without specifying a personality disorder diagnosis, or when the level of personality impairment falls below the threshold for such a diagnosis[153].

Clinical interviews have been explicitly developed to assess the DSM-5 LPFS, including the *Clinical Interview for the DSM-5 AMPD Module*[154], the *Semi-Structured Interview for Personality Functioning DSM-5*[155], and the *Clinical Assessment of the Level of Personality Functioning Scale*[156]. A range of self-report measures of the LPFS have also been developed[157-159], even though none is currently considered as the gold standard[160].

Post-traumatic stress disorder

Assessing functioning beyond core symptoms in individuals with post-traumatic stress disorder (PTSD) is crucial, as it provides a holistic view of their well-being, encompassing social, occupational and daily living aspects. The disorder often impacts these areas, hindering recovery and quality of life despite symptom management. Understanding functional impairments aids in developing comprehensive treatment plans that address the full spectrum of a patient's needs[161].

However, the evaluation of functioning in this population poses challenges. Assessment tools may not capture the complexity and variability of functional impairments, and individuals with PTSD might underreport or misinterpret their functional limitations due to stigma, avoidance or cognitive biases[162]. Additionally, cultural differences and personal contexts may influence responses, making it difficult to standardize assessments across diverse populations[163]. Thus, careful consideration and possibly multi-method approaches are necessary for accurate evaluation.

Transdiagnostic measures validated in individuals with PTSD include the WHODAS 2.0, which showed good psychometric properties in samples of veterans[164] and scores sensitive to changes in a pharmacological trial of PTSD[165], and the GAF, which demonstrated high inter-rater reliability between clinicians in veterans with PTSD[166]. The SF-36 and SF-12 are likewise used in the context of PTSD, but studies validating them in this condition are not available.

In terms of disorder-specific tools, three measures should be highlighted. The first is the *Posttraumatic Stress Related Functioning Inventory (PRFI)*[167], developed in 2016, a patient-rated tool covering three domains: work and school, relationships, and lifestyle, each with subscales assessing the impact of specific PTSD symptom clusters. The tool contains 26 items plus an additional one for detailed information, scored on a Likert scale from 0 to 4, taking 10-15 min to complete. It is freely available and demonstrates high reliability and validity, with strong correlations to total PTSD symptom scores and moderate correlations to measures of depression and substance use[167].

The *Inventory of Psychosocial Functioning (IPF)*[168], developed in 2018, is a patient-rated tool that evaluates psychosocial functioning across seven areas: romantic relationships, family relationships, work, friendships and socializing, parenting, education, and self-care. It includes 80 items scored on a Likert scale from 0 to 6, with both total and domain-specific scores. It takes about 25-30 min to complete and is freely available. It has excellent psychometric properties and construct validity[168].

The *Brief Inventory of Psychosocial Functioning (B-IPF)*[169], published in 2020, is a shortened version of the IPF, evaluating the same functional domains. Participants skip any non-relevant items. It consists of seven items scored on a 0-6 Likert scale, tak-

ing less than 5 min to complete. It is freely available and has good reliability and validity, with strong correlations to the IPF and other measures of mental health impairment and quality of life[169].

Since the WHODAS 2.0 is validated in people with PTSD, it represents a reasonable transdiagnostic choice. Depending on the desire to include more disorder-specific aspects and time requirements for the completion of the scale, the PRFI can be employed instead. These instruments have been mainly tested in populations of veterans, and further studies on more diverse PTSD samples could provide further insight into their proposed use.

Schizophrenia

Given the pervasiveness of disability in schizophrenia-spectrum disorders, there have been many attempts to develop tools to assess functioning in these conditions. Within a systematic review and worldwide expert consensus effort, 59 different functional status assessment tools were nominated by experts[170]. Ultimately, the review process examining available data identifed ten assessment measures with evidence of adequate reliability and validity, as well as convergence with other measures of functioning (other scales, real-world milestones) and important predictors of functioning, such as cognition and performance-based measures of functional capacity. A more recent review[171] identified many of the same measures as relevant. Notably, these instruments are designed to capture either performance-based or multi-perspective ratings (informal caregivers, clinicians), in order to avoid challenges associated with self-reports in people with schizophrenia.

Although there have been systematic reviews of newer PROMs [172], their overall usefulness is challenged by consistent findings that self-reports have only modest convergence with the reports of other informants[172,173], and with observable data such as functional milestones[174] or performance-based measures of critical everyday skills[175]. High-contact formal or informal caregivers of people with schizophrenia have been found to generate ratings that are more convergent with observable data from other sources[172]. Perspectives on experience and personal goals constitute critical information and are not necessarily required to be linked to real-world observable behavior, but such perspectives should be complemented by measures of clinical recovery to provide a comprehensive view.

During the above systematic review process[170], it became clear that the available tools differed in their functional targets, with some focusing on everyday activities or on social functioning, and others capturing a range of elements. Further, many tools included subscales that measured symptoms and disruptive behavior, which are outside the domain of functioning *per se*. Some instruments generated composite indices rather than individual domain scores for ADLs, productive activities, and social functioning. Finally, some tools did not provide definitions or guidance for ratings, meaning that different studies could employ different definitions of domains of functioning, and different weightings of outcome domains to lead to a single global tool.

One factor that varied considerably across the tools was the extent to which prior knowledge of the person with schizophrenia was required for ratings. Some instruments have detailed interview strategies and others are summary scales. Moreover, some of the summary scale measures do not specify the origin of information required to generate the ratings. Without this specification, it is possible that not only the basis of ratings is different across raters, but that information sources in longitudinal assessments of the same participant vary as well.

Evidence that different domains defined by achievement of functional milestones (marriage or equivalent, competitive employment, and living independently) are only weakly correlated with each other[176] suggests that global ratings across domains provide inadequate resolution for functional assessments. Thus, even if composite scores are available, they may not reflect all elements of functioning, and averaging across domains may not provide an accurate full-range functional assessment.

A transdiagnostic tool that has been specifically used and validated in clinical populations of individuals with schizophrenia is the SLOF, which has been found to be a reliable and valid instrument in an Italian[177,178] and a Japanese[179] sample. The PSP showed an intraclass correlation of 0.79 for subjects whose underlying condition did not change, and correlated (r=.61) with alternative global measures[77]. The WHODAS 2.0 could also be used, but data in schizophrenia are limited, despite the fact that it has been selected by an ICHOM consensus[37].

A specific tool that has been used to measure social functioning in family intervention programmes for patients with schizophrenia is the *Social Functioning Scale (SFS)*[180], which has seven subdomains: withdrawal/social engagement, interpersonal communication, independence-performance, independence-competence, recreation, prosocial, and employment/occupation. However, its 79 items makes the tool unpractical for administration, which takes about 45 min.

Substance use disorders

Assessing functioning beyond core symptoms is essential for individuals with substance use disorders, as it provides a holistic view of their impairments and strengths, including daily activities, relationships, and quality of life[181,182]. This comprehensive approach helps tailor interventions for sustainable recovery. However, challenges include the subjective nature of self-reports, cognitive impairments from substance use, and potential stigma or denial affecting accurate reporting[181,182]. Clinicians must use a combination of self-assessments, clinical interviews, and objective measures to overcome these challenges.

Among transdiagnostic tools, the SF-36 and SF-12 have been validated in clinical populations of individuals with a substance use disorder (alcohol dependence[183], and alcohol, cannabis and cocaine use disorders[69], respectively). Further studies suggest that these tools are useful in evaluating the effects of treatment across different settings (e.g., detoxification programs[184] and outpatient interventions[185]).

Regarding disorder-specific tools, the *Addiction Severity Index (ASI)*[186,187], developed in 1980, assesses psychosocial functioning

related to substance use, covering seven domains (medical status, employment/support status, alcohol or drug status, legal problems, family and social relationships, and psychiatric problems), through 163 items. The ASI generates a composite score ranging from 0 to 9, and takes about 60 min to complete. It has excellent psychometric properties[186] and is freely available. It has been validated in studies with multiple substance use disorders (e.g., alcohol and opioid use disorder[188-191]).

TRANSDIAGNOSTIC QUALITY OF LIFE, WELL-BEING AND LIFE SATISFACTION ASSESSMENT TOOLS

Quality of life

Despite the multitude of available quality of life assessment tools, concerns have been raised regarding their content validity and suitability for use in the mental health field. A systematic review[192] examined the available instruments using the evaluation framework by Connell et al[193], to assess whether they cover the dimensions highly valued by people with mental health problems: well-being and ill-being; relationships and belonging; activity; self-perception; autonomy; hope and hopelessness; and physical health. A total of 44 quality of life instruments were identified, of which 12 were adapted versions of original instruments. None of the identified instruments fully covered all dimensions of the evaluation framework. The review highlighted the challenge of adequately measuring the current quality of life status, particularly due to the insufficient coverage of key dimensions such as "hope and hopelessness" and "self-perception".

We describe here eight key instruments for the assessment of quality of life, in historical order of development. The first is the *European Quality of Life 5-Dimension 5-Level questionnaire (EQ-5D-5L)*[194,195], developed by the EuroQuality of Life Group in 1990 and updated by the EuroQuality of Life Research Foundation in 2009. This is a patient-reported tool designed to assess quality of life across five domains: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Responses in these five dimensions can be converted into a single summary index by weighting each level in each category. The tool includes five items and one visual analogue scale, with item scoring from 1 (no problem) to 5 (unable to/extreme problems). The higher the score, the lower the quality of life. It takes about 5 min to complete and is available for free. It has demonstrated very high test-retest validity and good internal consistency[196]. Validity studies have shown a reasonable degree of content and construct validity[196].

The *EuroQuality of Life 5-Dimension - Proxy Version (EQ-5D Proxy)*[197] is an informant-reported version of the EQ-5D-5L designed for use when patients cannot self-report. It assesses the same five domains as the EQ-5D-5L and is also free. The *EuroQuality of Life 5-Dimension 3-Level (EQ-5D-3L)*[198] is another version of the EQ-5D, assessing the same five domains but with simplified three levels for each dimension: no problems, some problems, and extreme problems. It includes an additional analogue scale ranging from 0 (worst imaginable health state) to 100 (best imaginable health state). This tool also takes about 5 min to complete and is free. It has demonstrated high reliability and good validity, with moderate to strong correlations between most dimensions and physical measures[199].

The *Quality of Life Inventory (QOLI)*[200], developed in 1992, is a patient-reported measure assessing 17 areas: health, self-regard, philosophy of life, standard of living, work, recreation, learning, creativity, social service, civic action, love relationships, friendships, relationships with children, relationships with relatives, home, neighbourhood, and community. It includes importance ratings (0 to 2) and satisfaction ratings (–3 to 3), which are multiplied to form weighted total ratings. It takes about 15 min to complete and is licensed. It has shown good reliability and validity, with internal consistency coefficients ranging from 0.77 to 0.89[200].

The *World Health Organization Quality of Life Assessment (WHOQOL-100)*[201], developed by the WHO in the mid-1990s, is designed to be a cross-culturally applicable instrument covering six broad domains: physical health, psychological health, level of independence, social relationships, environment, and spirituality/religion/personal beliefs. It consists of 100 questions and takes about 20-30 min to complete. It is available for free. It has shown good test-retest reliability and fair relationship with the SF-36[202,203].

The *World Health Organization Quality of Life Brief Version (WHOQOL-BREF)*[204], developed by the WHO in 1998, is an abbreviated version of the WHOQOL-100, assessing four dimensions: physical health, psychological health, social relationships, and environment. It includes 26 items (24 specific and 2 general), scored from 1 (worst) to 5 (best), with scores transformed into a 0-100 range. It takes about 5-10 min to complete and is free with a user agreement. It has demonstrated good reliability and validity across various populations[e.g.,205].

The *Manchester Short Assessment of Quality of Life (MANSA)*[206], developed in the late 1990s, is a 12-item patient-rated instrument (no subscales) that takes <5 min to complete, designed to be a concise tool for evaluating quality of life in clinical settings, with good test-retest reliability and validity. It is well known for its brevity and ease of use, while providing reliable and valid measures of quality of life, particularly in clinical populations[207].

Well-being

Despite the widespread interest and interdisciplinary research, there is no consensus on the best methods for assessing well-being. The challenge lies in the variety of definitions and theories, which range from basic human needs to capabilities, leading to a proliferation of instruments without a universally accepted standard, although it should be noted that different individuals may actually have different views of what is well-being for them.

A systematic review[208] identified 99 generic measures of well-being used for adults, spanning from the 1960s to recent years. It highlighted the lack of agreement on what constitutes well-being and how it should be assessed. Instruments vary significantly in content, dimensions, and theoretical foundations, reflecting a

broad range of perspectives and disciplines. While new tools continue to emerge, many older measures remain in use, reflecting a growing but fragmented field. The review emphasized the need for further research to address the ambiguity in well-being measurement and to refine the conceptual and psychometric aspects of existing tools. Understanding the diverse dimensions and their overlap can help researchers choose the most suitable instruments and improve the consistency and relevance of well-being assessments.

We have selected three key assessment tools. The first is the *Personal Wellbeing Index (PWI)*[209], developed by the International Wellbeing Group in 2013. This is a patient-reported measure of well-being across seven domains: standard of living, health, achieving in life, relationships, safety, community connectedness, and future security. It includes seven items scored from 0 (completely dissatisfied) to 10 (completely satisfied), with higher scores representing a stronger sense of well-being. It takes about 2 min to complete and is free. It has shown good reliability and validity[210].

The *Quality of Well-being Scale - Self-Administered (QWB-SA)*[211], developed in 1997, is a patient-reported measure of well-being across five domains: chronic illness, self-care, mobility, physical activity, and usual activities. It includes yes/no questions for illness and a 1 to 5 scale for other domains. It takes about 5 min to complete and is free. It has demonstrated good reliability and validity[211].

The *World Health Organization Well-Being Index (WHO-5)*[212], developed by the WHO Regional Office for Europe in 1998, is a patient-reported measure of well-being across five domains: cheerful/good spirits, calm/relaxed, active/vigorous, fresh/rested, and filled with things of interest. It includes five items scored from 0 (never) to 5 (all the time), with total scores ranging from 0 to 25 and multiplied by 4 to obtain a percentage scale value. It takes about 2-3 min to complete and is free. It has demonstrated good reliability (Cronbach's alpha of 0.923) and high clinical validity[213,214].

Life satisfaction

Developing scales to measure life satisfaction in mental health contexts faces several challenges[215,216]. Conceptually, defining life satisfaction and differentiating it from related constructs such as happiness and well-being is complex. Methodologically, selecting appropriate items, avoiding response biases, and ensuring reliability and validity are significant hurdles. Culturally, developing a scale that is sensitive to different cultural norms and translating it accurately poses difficulties. Practically, choosing the right administration method, and balancing the scale's comprehensiveness with its length are critical considerations. We have selected here four transdiagnostic instruments measuring life satisfaction.

The *Satisfaction With Life Scale (SWLS)*[217], developed in 1985, is a patient-reported measure of global life satisfaction. It includes five items scored from 1 (strongly disagree) to 7 (strongly agree), with higher scores indicating higher global life satisfaction. It takes 2-3 min to complete and is free. It has demonstrated good reliability and validity, and significant correlations with social support. It is also sensitive to changes in PTSD symptoms and alcohol use during treatment[218-220].

The *Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q)*[221], developed in the early 1990s, is a patient-reported measure of life satisfaction across 13 domains (physical health, mood, work, household duties, school/course work, leisure activities, social relationships, family relationships, general activities, economic status, living situation, sexual drive and interest, and self-care). It consists of 93 items, with an administration time of about 10-15 min. It has shown good test-retest reliability[222].

The *Quality of Life Enjoyment and Satisfaction Questionnaire - Short Form (Q-LES-Q-SF)*[223], published in 1998, includes 14 domains: physical health, mood, work, household activities, social relationships, family relationships, leisure time activities, ability to function, sexual interest and performance, economic status, living/housing situation, mobility, vision, and overall sense of well-being. It includes 16 items, scored from 1 (very poor) to 5 (very good), and takes less than 5 min to complete. It is free and has demonstrated high reliability and validity[224].

The *Satisfaction Profile (SAT-P)*[225], published in 2000, is a patient-reported measure of life satisfaction across five domains: psychological functioning, physical functioning, work, sleep/eating/free time, and social functioning. It includes 32 items scored on a visual analogue scale ranging from "extremely dissatisfied" to "extremely satisfied". It takes about 10 min to complete and is free. It has shown high reliability and good test-retest reliability[225].

QUALITY OF LIFE, WELL-BEING AND LIFE SATISFACTION ASSESSMENT IN SPECIFIC MENTAL DISORDERS

We present here a selection of instruments for measuring quality of life, well-being and life satisfaction in specific mental disorders, listed in alphabetical order.

Anxiety disorders

Developing measures of quality of life, well-being and life satisfaction in people with anxiety disorders is challenging, due to the subjective nature of these experiences and the variability in individual symptoms. A key difficulty is capturing the nuanced impact of anxiety on daily functioning, distinguishing between anxiety-related impairments and other co-occurring issues[226].

In the EMA guideline on the clinical investigation of medicinal products indicated for generalized anxiety disorder, panic disorder and social anxiety disorder[97,98], the use of quality of life measures is explicitly recommended as secondary outcome parameters. Similarly, the NICE guidelines for the management of generalized anxiety disorder/panic disorder and social anxiety disorder[99,100] point to the necessity of including measures of quality of life when assessing recovery.

Some transdiagnostic quality of life tools have been used in clinical populations of individuals with anxiety disorders. For example, the QOLI has shown a Cronbach's alpha of 0.85 for the over-

all score, 0.81 for the first factor (self-oriented quality of life), and 0.75 for the second factor (externally oriented quality of life)[227]. The QOLI has been validated in generalized anxiety disorder, panic disorder and social anxiety disorder, and has been shown to be sensitive to change after Internet-administered psychological interventions in social anxiety disorder[228].

Another useful tool in people with anxiety disorders is the WHOQOL-BREF, whose scores correlated negatively with anxiety scores after cognitive behavior therapy[229]. Furthermore, in generalized anxiety disorder and panic disorder/agoraphobia, the Q-LES-Q-SF was shown to be sensitive to change after pharmacological treatment[230]. No anxiety-specific tool was deemed to be suitable.

Bipolar disorder

Improvements in quality of life, well-being and life satisfaction are prioritized as therapeutic objectives in persons living with bipolar disorder[231]. However, measuring these aspects is complex. Challenges include accounting for the impact of mood swings on daily functioning, distinguishing between the effects of the disorder and treatment side effects, capturing the subjective experience during different mood states, and considering the influence of concurrent physical diseases. Additionally, measures need to address the impact of relationships and work on quality of life, and interactions with self-esteem, and be sensitive to individual differences in how bipolar disorder affects life satisfaction[232].

In terms of transdiagnostic instruments measuring life satisfaction in people with bipolar disorder, the Q-LES-Q has shown Cronbach's alpha values ranging between 0.88 to 0.96, and intra-class correlations between 0.80 and 0.97[222]. Among tools assessing well-being, the WHO-5 showed high test-retest reliability (r=0.83) in a sample of individuals with bipolar disorder[233].

The only instrument measuring quality of life specifically in bipolar disorder is the *Quality of Life in Bipolar Disorder (QoL.BD)*, published in 2010[234]. This is a 56-item questionnaire developed in consultation with individuals with lived experience, assessing the impact of symptoms (mood, sleep, physical health, cognition), role functioning (household management, finances, and the optional areas of work and education), and the domains of leisure, relationships, self-esteem, spirituality, identity and independence. A systematic review[235] has confirmed its good psychometric properties. A brief version of this tool, the *Brief Quality of Life in Bipolar Disorder (Brief QoL.BD)*, including 12 items, takes only 4 min to administer and is suited for outcome assessment in clinical practice.

Dementias

Concerns have been expressed about the ability of several persons with dementias to evaluate their own quality of life or well-being. Self-rated quality of life measures may not correlate well with observable life circumstances or behaviors, and need to be interpreted with this potential disconnect in mind. Informant reports that cover longer periods of time than the situational rating of the individual may need to be taken into account. Nevertheless, some people with dementia may have a wide range of preserved cognitive functions, including the ability to remember situations and judge their own affective, cognitive and physical well-being[236].

Among the transdiagnostic tools, the EQ-5D Proxy and the EQ-5D-5L have shown convergent validity with some of the disorder-specific instruments described below[237,238].

We selected three disorder-specific tools. The first is the *Alzheimer's Disease Related Quality of Life scale (ADRQL)*[239], developed in 1991, a clinician- and informant-administered tool designed to assess health-related quality of life in people with Alzheimer's disease. This scale covers five domains: social interaction, awareness of self, feelings and mood, enjoyment of activities, and response to surroundings. It includes 47 dichotomous items, with higher scores indicating higher quality of life. It takes about 30-45 min to complete and is free. It has shown good reliability and validity, with significant correlations with other instruments, and the ability to discriminate between groups with different levels of cognitive and physical functioning[240].

The *Dementia Quality of Life - Proxy scale (DEMQOL-Proxy)*[241], developed in 2007, is a clinician- and informant-administered measure assessing quality of life across five domains: health and well-being, cognitive functioning, daily activities, social relationships, and self-concept. It includes 29 items, with higher scores indicating higher quality of life. It takes about 10 min to complete and is free. It has demonstrated good reliability and validity, with significant correlations with other similar instruments[240].

The *Quality of Life in Alzheimer's Disease (QoL-AD)*[242], developed in 1999, is a clinician- and informant-rated tool assessing quality of life in multiple domains: physical health, energy, mood, living situation, memory, family, marriage, friends, self as a whole, ability to do chores, ability to do things for fun, money, and life as a whole. It includes 13 items scored from 1 (poor) to 4 (excellent), with higher scores indicating better quality of life. It takes about 10 min to complete and is free. It has shown good reliability and validity, and significant correlations with external measures of cognitive and functional status, depression, and pleasant events[242].

Eating disorders

Developing tools to measure quality of life, well-being and life satisfaction in people with eating disorders presents several challenges. These include addressing the complex and often fluctuating nature of eating disorder symptoms, ensuring sensitivity to the emotional and psychological state of respondents, and capturing the multifaceted impact of the disorder on various life domains[243].

We selected four disorder-specific instruments. The first is the *Eating Disorders Quality of Life Scale (EDQLS)*[244], developed in 2007, a patient-reported tool designed to assess quality of life across 12 domains: cognitive, education/vocation, family and close relationships, relationships with others, future outlook, appearance, leisure, psychological, emotional, values and beliefs, physical, and

eating. It includes 40 items, scored on a scale from 1 to 5 (39 items) or 1 to 10 (one global quality of life item), taking about 5-10 min to complete. It is free to use and has shown high validity, and significant associations with psychiatric comorbidity, psychiatric symptoms, eating disorder symptoms, time in treatment, and stage of change[244].

The *Eating Disorders Quality of Life (EDQoL)*[245], developed in 2006, is a patient-rated measure including four subscales: physical/cognitive, psychological, work/school, and financial. It consists of 25 items scored from 0 to 4, with an average score derived from these items. It takes about 5-10 min to complete and is available for free. It has demonstrated high reliability and validity, with confirmatory factor analysis showing strong fit indices[245].

The *Health-Related Quality of Life in Eating Disorders (HeRQoLED)*[246], version 2, developed in 2006, is a patient-reported measure of quality of life consisting of nine subscales: symptoms, restrictive behavior, body image, mental health and functionality, emotional role, physical role, personality traits, social relations, and binge. It includes 55 items scored from 0 to 4/5, with scores standardized to 100. It takes about 15-20 min to complete and is free. It has shown high reliability and validity, with good fit indices from confirmatory factor analysis[246].

The *Health-Related Quality of Life in Eating Disorders - short form (HeRQoLED-s)*[247], developed in 2007, is a shorter version of the HeRQoLED. It assesses quality of life across two subscales: social maladjustment, and mental health and functionality. It includes 20 items scored from 0 to 4/5, with scores standardized to 100. It takes about 5 min to complete and is free. It has demonstrated high validity and good fit indices from confirmatory factor analysis[247].

The choice among these instruments should be based on clinician experience and the availability of a validated version in the native language of the patient.

Major depressive disorder

Quality of life is highly relevant to MDD, but is often overlooked, despite being largely independent of symptoms and being regarded as a priority by people with a lived experience of this condition[3,248]. Indeed, clinical remission in MDD does not necessarily imply an improvement in quality of life, and it has been argued that the construct of "remission" should include quality of life and well-being in addition to symptom level[249-252].

Among transdiagnostic tools, the QOLI has shown a Cronbach's alpha of 0.85 in patients with MDD[228], and the WHO-5 has demonstrated a significant association with self- and observer-rated measures of depressive symptoms[253]. Other widely used tools to assess quality of life and life satisfaction in patients with depression are the Q-LES-Q and the EQ-5D-3L[199].

No disorder-specific tool for the assessment of quality of life, well-being and life satisfaction in patients with MDD can be recommended. Tools specifically aimed to assess depressive symptoms such as interest or pleasure in activities are not regarded as relevant here.

Obsessive-compulsive and related disorders

Developing tools to assess quality of life, well-being and life satisfaction in individuals with OCD and related disorders is a complex endeavour. The subjective nature of these constructs, compounded by the variability of obsessive-compulsive and related symptoms and their impact on daily functioning, makes it difficult to create universally applicable and sensitive instruments. Additionally, the presence of comorbid conditions, such as depression and anxiety, can confound assessments, requiring tools that distinguish between the effects of OCD and those of other disorders[136].

Among the transdiagnostic instruments, the Q-LES-Q-SF, the WHOQOL-BREF and the QOLI have shown some evidence of sensitivity to change in patients with OCD[136].

Personality disorders

Few studies have delved into quality of life, well-being and life satisfaction in personality disorders[254]. However, cross-sectional community-based investigations have identified these disorders as significant negative predictors of quality of life, surpassing the predictive value of sociodemographic variables, physical health, and other comorbid mental disorders[255,256]. This predictive role is also supported by evidence over an 17-year period showing that any personality disorder in young adulthood was independently associated with significant impairment in overall quality of life, with antisocial, borderline and schizotypal personality disorder symptoms being linked to higher quality of life impairment[257]. Moreover, evidence from a multicenter study on clinical samples indicates that quality of life in adults with personality disorders (measured by the EQ-5D-3L) is similar to that of people suffering from severe physical diseases, including rheumatic disease, lung cancer and Parkinson's disease[258].

The social domain of quality of life seems to be most affected in people with personality disorders. A meta-synthesis of qualitative studies on loneliness among these people described an intense sense of disconnect and struggle with unmet social needs[259]. This finding is also corroborated by quantitative studies reporting that (except for narcissistic personality disorder) people with full-blown or subthreshold personality disorders have higher levels of perceived loneliness, lower relationship satisfaction, and poorer social support than the general population or other clinical samples[260].

The physical domain of quality of life has also been shown to be specifically impaired in people with borderline personality disorder, even in relatively young samples (range 21-54 years)[261]. This finding can be due to lifestyle and risk-taking behaviors, such as self-harm and heavy alcohol or other substance use[261]. However, the role of behavioral symptoms in affecting physical quality of life is controversial, as cohort studies document that borderline personality disorder is highly predictive of physical health diseases even when controlling for unhealthy behavior[262].

Evidence regarding quality of life as a modifiable outcome in trials of pharmacological and psychosocial treatments for personal-

ity disorders is limited to people with borderline personality disorder. A meta-analysis found that specialized psychotherapies for this disorder outperformed control conditions in improving quality of life (Cohen's d = 0.32)[263]. Also, meta-regression analysis showed a lack of association between changes in borderline personality disorder symptom severity and changes in quality of life, supporting treatment approaches addressing quality of life beyond symptom reduction[263].

Among transdiagnostic tools, the EQ-5D-5L is by far the most adopted as an outcome parameter in borderline personality disorder trials[261]. The WHOQOL-BREF has also been used as an outcome measure for quality of life in clinical trials testing specialized psychotherapies for borderline personality disorder[263,264]. Regarding life satisfaction, the SWLS is widely adopted in clinical populations, albeit no validation studies in samples with personality disorders have been conducted so far.

No instruments assessing quality of life, well-being or life satisfaction have been specifically developed for personality disorders.

Post-traumatic stress disorder

Subjective measures of quality of life, well-being and life satisfaction add significantly to the full picture of mental health in the context of PTSD. However, PTSD symptoms – such as hyperarousal, avoidance, and intrusive thoughts – can vary widely among individuals, affecting their perception of life quality in different ways[162].

Several transdiagnostic tools have shown to be reliable and valid measures in the context of PTSD. The PWI had good reliability in studies on Israeli civilians living within the range of fire from the Gaza Strip[265], in inpatients from a residential PTSD treatment program in Australia[266], and in refugee populations with PTSD in Australia and New Zealand[210], with Cronbach's alpha values, respectively, of 0.94, 0.84 and 0.85.

The Q-LES-Q-SF had good reliability in veterans with PTSD (alpha = 0.93)[267,268]. The SWLS showed good reliability (alpha = 0.88) in Croatian war veterans[269]. The WHOQOL-BREF had moderate reliability in Colombian ex-combatants from illegal armed groups, with alpha values ranging from 0.60 (physical health, social relationships) to 0.80 (environment)[270], and in male tortured refugees, with alpha values ranging from 0.76 to 0.85[271]. The WHO-5 also showed good reliability (alpha = 0.93) in PTSD[272].

So far, no PTSD-specific instrument has been developed to measure quality of life, well-being or life satisfaction. The choice of the instrument should be based on the specific construct to be addressed and the psychometric properties of each tool (for a given population).

Schizophrenia

People with schizophrenia have reports of quality of life that correlate only very minimally with observable functional information, and most studies find that the predominant response bias is to underestimate illness burden. One of the negative symptoms of schizophrenia, lack of normal distress, is critically interwoven with subjective quality of life[273]. This symptom is rated when individuals report minimal concern with major life challenges, such as lack of friends, minimal productive activities, and financial challenges. If these problems, as well as living in substandard conditions, do not elicit any subjective concern, quality of life indices must be interpreted accordingly.

Study results[274] have been quite consistent, with self-reports on quality of life measures in participants with schizophrenia generally being much more strongly correlated with current mood states than with other elements of everyday functioning. These data suggest that quality of life self-reports are not invalid, but rather that they provide a window into experiences of mood states rather than disability. Interestingly, mood symptoms are also consistently correlated with self-reports of everyday disability. Thus, both quality of life and disability measures may provide an index of subjective distress and impairment associated with mood states, but not with other elements of the illness[275,276]. As a result, commonplace quality of life measures can be completed by participants with schizophrenia, but the results need careful interpretation.

In terms of disorder-specific instruments, we have selected four tools. The first is the *Heinrichs-Carpenter Quality of Life Scale (QLS)*[277] a 21-item tool (45 min to complete) exploring sense of purpose, motivation, curiosity, interpersonal relations, extent of functioning, accomplishment, employment, and common objects and activities, with excellent psychometric properties when there is sufficient time to conduct a clinician-led interview[278].

The *Schizophrenia Quality of Life Questionnaire 41 (S-QoL 41)*[279] is a self-reported instrument sensitive to change that takes 15-20 min to complete, with 41 items composing eight subscales (psychological well-being, self-esteem, family relationships, relationships with friends, resilience, physical well-being, autonomy, and sentimental life) and a total score.

The *Schizophrenia Quality of Life Scale (SQLS)*[280] is a 30-item self-reported measure with three subscales (psychosocial, motivation and energy, symptoms and side effects), each with a range from 0 (best possible health state) to 100 (worst possible health state). It takes 10-15 min to complete. This scale has undergone further development resulting in the *Schizophrenia Quality of Life Scale Revision 4 (SQLS-R4)*[281], comprising 33 items with two subscales (psychosocial feelings, cognition and vitality), that also takes 10-15 min to complete. While the SQLS exhibited favourable convergent validity, it had limitations regarding structural validity, internal consistency, reliability, and known-groups validity. The SQLS-R4 demonstrated promising reliability and convergent validity, yet it faced challenges in structural validity, internal consistency, cross-cultural validity, known-groups validity, and responsiveness[282].

Substance use disorders

Individuals affected by substance use disorders experience a substantial decline in their quality of life, including – but not limited to – negative impacts on physical, social, emotional and eco-

nomic domains[283].

While no substance use disorder-specific tools for the assessment of quality of life, well-being or life satisfaction exist, some transdiagnostic measures have been explored. Among them, the EQ-5D-5L has been validated across different disorders, but a systematic review concluded that high-quality evidence is further needed[284], with a study reporting small to medium responsiveness over time[285]. In the same review[284], the WHOQOL-BREF, the most widely used tool for quality of life assessment in this field, showed replicated good internal consistency, with all values for the separation indices above 2, and evidence of structural validity, with the original factor structure being supported.

PERSONAL RECOVERY MEASURES IN MENTAL HEALTH

It is increasingly recognized that observable or self-reported illness symptoms and functional outcome domains that can be used to define clinical recovery are on a different level from self-perceived personal recovery[286]. Empirical evidence confirms that personal recovery and clinical recovery are distinct constructs[287]. Personal recovery is not unrelated to clinical recovery, but incorporates personal values and appraisals, and is not solely informed by illness symptoms and functional behavior. This distinction has two implications. First, the type of recovery (clinical versus personal) should be clarified when reporting individual or aggregated data on recovery. Second, the centrality of diagnosis within traditional psychiatric research is not mirrored in the development of recovery measures, which are almost always transdiagnostic. Rather, individual and often dynamic attitudes, values, goals and preferences define the concept of personal recovery. Considering personal recovery in the assessment, alliance formation, diagnostic formulation, as well as treatment planning and implementation is key in order to enhance treatment personalization, satisfaction, adherence and disease outcomes[288].

Here we provide a brief description of selected tools to measure personal recovery, followed by some remarks on tools that are currently not recommended for use.

The *Brief INSPIRE-O*[289], based on the INSPIRE tool developed in 2015[290], is a patient-reported measure of personal recovery that has been translated into many languages. The tool comprises five items ("I feel supported by other people", "I have hopes and dreams for the future", "I feel good about myself", "I do things that mean something to me", and "I feel in control of my life"), each scored on a 5-point Likert scale ranging from 0 to 4. The total score is calculated by summing the item scores and multiplying by 20, resulting in a range from 0 (low recovery) to 100 (full recovery). Completion takes approximately 2 min, and the tool is available for free. The Brief INSPIRE-O has demonstrated good internal consistency, with a Cronbach's alpha of 0.83, and moderate correlation with social contacts and self-reported general health[289].

The *Questionnaire about the Process of Recovery (QPR)*[291], developed in 2009, is the most widely used patient-reported measure in English-speaking countries for research purposes. It is designed to assess personal recovery across intrapersonal and interpersonal domains. It is available in two versions: a 22-item and a 15-item one. Each item is rated on a 5-point scale from 0 to 4. For the 22-item version, scores are summed to provide intrapersonal (0 to 68) and interpersonal (0 to 20) recovery scores. The total score for the 15-item version ranges from 0 (low recovery) to 60 (full recovery). The tool takes about 10 min to complete for the 22-item version and 5 min for the 15-item version, both available for free. It has shown high internal consistency and good test-retest reliability. Construct validity is evidenced by positive correlations with empowerment and quality of life, and negative correlations with symptoms. Increased recovery scores were associated with increased self-esteem and functioning, and decreased psychopathology, hopelessness, and depression[292-294].

The *Recovery Assessment Scale (RAS)*, initially developed in 2004 [295], is the most widely used patient-reported measure of recovery in English-speaking countries for clinical purposes. It includes 24 items rated on a 5-point Likert scale from 1 to 5, with total scores ranging from 24 (low recovery) to 120 (full recovery). The *Recovery Assessment Scale - Domains and Stages (RAS-DS)*[296] consists of 38 items scored on a 4-point scale from 1 to 4, with total scores ranging from 38 (low recovery) to 152 (full recovery). The RAS takes about 10-15 min to complete and is free. It has demonstrated a consistent factor structure, positive associations with related constructs, and sensitivity to change over time. The RAS-DS has shown good internal consistency and ability to detect changes over time, with significant increases in scores observed between assessments[297-300]. The RAS-DS is available for free in 18 languages.

The *Camberwell Assessment of Need (CAN)*[301,302] assesses health and social needs, from the perspective of staff, service user and informal carer. While this family of tools does not directly measure recovery, it has been argued that there is an overlap between the concepts of recovery and unmet needs (when an individual has many unmet needs, this is likely to hamper his/her recovery; when these needs are met, this is likely to support recovery[302]). It has also been suggested that having few unmet needs can be regarded as an "objective measure of recovery"[303].

The CAN family of tools, translated into many languages, includes: the standard CAN for adults with mental health problems, CAN for the elderly (CANE)[304], CAN for developmental and intellectual disabilities (CANDID)[305], CAN-forensic (CANFOR)[306] for forensic patients, CAN-M for mothers and pregnant women with mental health issues[307], and Humanitarian Emergency Settings Perceived Needs Scale (HESPER) for individuals in disaster situations[308].

The "patient-rated unmet need" total in the CAN has emerged from both a recovery values and an empirical perspective as the most valuable and informative score, as it both centres attention on the service user's perceptions, and identifies life domains beyond the traditional focus of psychiatry to target clinical support towards.

Targeting unmet needs is important, as reducing patient-rated unmet needs causally impacts on quality of life[309], therapeutic alliance[310], and satisfaction with care[311]. There is also evidence of an association between unmet needs and symptomatology (e.g., anhedonia)[312], worse social outcomes (e.g., lower employment)[313],

worse service experience (e.g., more compulsory admissions)[314], more use of emergency department services[315], and more rehospitalizations[316].

Each version of the CAN assesses health and social needs across multiple domains, whose number ranges from 22 to 26. Items are scored on a 5-point scale from 0 (not at all) to 4, and the total score ranges from 0 to 100. Completion times vary from 5 to 30 min, and all versions are available for free. The standard CAN demonstrated excellent inter-rater reliability, and good and excellent test-retest reliability for staff-rated total needs and for patient-rated total needs, respectively. Criterion validity was supported by significant correlations with the GAF. The CAN tools also showed good predictive validity, with reduced patient-rated unmet needs predicting improved subjective quality of life over time[317,318].

Apart from these recommended tools, there are numerous other recovery-related measures which may be suitable for use in specific settings. They are not recommended here due to their psychometric properties not being fully established; their length, complexity and burden of administration; or their narrow focus. Examples include the *Recovery Self-Assessment*[319], the *Mental Health Recovery Measure*[320], the *Recovering Quality of Life (ReQoL)*[321], and the *Individual Recovery Outcomes Counter (I.ROC)*[322].

DISCUSSION

Functioning, quality of life, well-being and life satisfaction are highly clinically and individually relevant outcome domains in mental disorders. We defined and reviewed each of these domains, presenting a selection of transdiagnostic assessment tools and, where specifically helpful, disorder-specific measures. We then conceptualized personal recovery with its main measurement tools.

While many measures exist, we selected the included ones on the basis of their frequency of use, psychometric properties, and clinical utility, including consideration of ease of use, time required for administration, usability for measurement-based care, and trans-setting (inpatient, outpatient, specialty and primary care) as well as trans-purpose (clinical care, research) characteristics. We excluded performance-based evaluations of functionality, as they do not qualify as measures that can be used in ordinary clinical care. We only focused on adults, as different considerations would be required concerning mental disorders in children and adolescents.

Based on our review, functional and quality of life/well-being measures are in most mental disorders restricted to non-disorder-specific tools. Notable exceptions include some functional assessment tools for anxiety disorders, PTSD, substance use disorders and dementias, as well as quality of life and well-being measures for eating disorders and dementias. While predominantly transdiagnostic assessments of functioning and quality of life/well-being may make sense (since these domains are to some degree universal, and capturing them transdiagnostically allows for comparison across disorders), individual tools may not have been validated in specific mental health conditions or capture all relevant aspects of a specific disorder.

In disorders where insight or cognition may be impaired, self-assessments and even interviews may need to be adapted to yield reliable and valid information. This situation may apply to people with active psychosis, eating disorders, and dementia. Moreover, although the broad dimensions of functioning (such as self-care, social interactions, leisure time activities, and educational/vocational activities) are not necessarily disorder-specific, baseline levels and functional expectations may vary across disorders and subgroups or illness stages. For example, in people with dementia, lack of worsening on an instrumental ADL scale may be an appropriate goal. Similarly, in people with schizophrenia and early illness onset, certain functional milestones may have never been achieved, so that functional disability may be assessed, but improvement may not be as easy and achievable as for people with other disorders. Such potential ceiling effects of a functional measure may also need to be considered when choosing or developing a functional assessment tool across people at different stages of their illness.

In assessing functional outcome domains, informal carer perspectives are largely missing, although arguably being even more important than for illness symptoms, as observation of behaviors that occur intermittently and outside of a clinical encounter would otherwise solely be based on patient report or recall. Hence, specific carer-rated instruments are needed.

To capture the needs of people with mental health conditions, more functional and quality of life/well-being measures need to be co-developed with these persons and their caregivers. To help implementation of such tools, measures need to be identified that can be completed with minimal effort or in a waiting room setting by people who may have cognitive difficulties or who are so burdened by their illness that lengthy self-assessments are unlikely finished. In this context, in order to address personal or rater bias, non-persistence or non-adherence with ratings, and lack of ecologic validity of the obtained information, both the research and clinical field should take advantage of new digital technologies, if feasible and adequately validated, safe and ethically acceptable[38,323].

While personal recovery is an increasingly important concept, and measures have been developed that can be used transdiagnostically, the assessment of this domain in research and, especially, in clinical care remains very limited. This situation is unfortunate, as taking into consideration the individual recovery concept of a person with a mental health condition can help orient the treatment plan in ways that are more likely to keep that person engaged in care.

The functional, quality of life/well-being and personal recovery measures summarized in this paper provide a starting point for research and clinical care, especially inasmuch as the tools are scalable. Additionally, the identified gaps highlight areas for active investigation and creativity to develop, extend, validate and, if necessary, shorten or refine the assessment tools, and to define the various domains in more meaningful and measurable ways in people living with mental health conditions.

Future research should integrate functional and quality of life assessments as much as possible in randomized as well as obser-

vational studies. Adopting brief and clinically informative measures in clinical practice would enable both measurement-based care and more granular database studies.

Across all the reviewed mental disorders, a combination of transdiagnostic and disorder-specific measures to assess psychosocial functioning might be advisable. However, future research is needed to identify the mental disorders that can benefit from disorder-specific assessments, and validate the relevant measures where this has not taken place yet. Moreover, in addition to instruments for the assessment of functioning that can be useful for research, tools that are brief and scalable and can be proposed for clinical care should become a greater focus of attention. These latter tools, which would ideally be co-developed with multiple stakeholders, could facilitate measurement-based care and bridge the gap between research knowledge and applicability to ordinary care, as well as facilitate implementation of research findings.

The next stage of personal recovery research involves longitudinal investigations to provide evidence-based prognostic advice to individuals experiencing mental health difficulties, as well as allowing the development of staged models of personal recovery support. The evidence has been clear for some time that disease outcomes in relation to personal recovery are better than clinical training and clinician attitudes would assume[324], and that there has been too much emphasis on pessimistic diagnostically-driven conceptualizations of prognosis. At least some of this favourable outcome potential may be related to the re-evaluation of personal needs and priorities when faced with a mental disorder, which is an adaptive task. An understanding of the factors involved in the dynamic process of personal recovery is increasingly becoming a focus of research efforts[325].

Since recovery includes both the clinical dimension (which can be assessed based on behaviors that are observable by the person with a mental health condition, informal caregivers or clinicians) and the personal recovery dimension (which is based on the subjective standpoint of the person), research should assess the relationship between these perspectives using multivariate repeated-measure designs to investigate long-term illness trajectories and disease outcomes.

In conclusion, this report summarizes what we consider at the current time the most relevant as well as valid and reliable measures to assess functioning, quality of life/well-being and personal recovery domains in people with mental disorders. Further research is needed to co-develop more meaningful measures with people who have lived experience of mental disorders, identify and implement assessment measures that can be used in research as well clinical practice by virtue of being sufficiently brief and broad, and leverage digital health tools. Cultural sensitivity[326] and cost-effectiveness should be additional concerns.

It is hoped that the use of a set of measures for the assessment of functioning, quality of life/well-being and personal recovery will become standard in research and clinical care settings, including trials of new therapeutics, ultimately aiming to improve care, care experiences and care outcomes of people living with mental disorders.

ACKNOWLEDGEMENTS

C.U. Correll and S. Cortese contributed equally to this work. S. Cortese is supported by the UK National Institute for Health and Care Research (NIHR) (grants nos. NIHR203684, NIHR203035, NIHR130077, NIHR128472, and RP-PG-0618-20003) and by the European Research Executive Agency (grant no. 101095568-HORIZON HLTH-2022-DISEASE-07-03). M. Slade is supported by the NIHR Nottingham Biomedical Research Centre; T.G. Rhee by the US National Institute on Aging (grants nos. R21AG070666, R21AG078972 and R01AG088647), the US National Institute of Mental Health (grant n. R01MH131528), the US National Institute on Drug Abuse (grant no. R21DA057540) and the US Health Resources and Services Administration (grant no. R42MC53154); R.S. McIntyre by the Canadian Institutes for Health Research and the Milken Institute; P. Fusar-Poli by the Italian Ministry of University and Research (MNESYS project). The views expressed in this paper are those of the authors and not necessarily those of the funding bodies.

REFERENCES

1. Thornicroft G, Slade M. New trends in assessing the outcomes of mental health interventions. World Psychiatry 2014;13:118-24.
2. Stein DJ, Shoptaw SJ, Vigo DV et al. Psychiatric diagnosis and treatment in the 21st century: paradigm shifts versus incremental integration. World Psychiatry 2022;21:393-414.
3. Demyttenaere K, Donneau AF, Albert A et al. What is important in being cured from depression? Discordance between physicians and patients (1). J Affect Disord 2015;174:390-6.
4. Verhaeghe P. On being normal and other disorders: a manual for clinical psychodiagnostics. Milton Park: Routledge, Taylor & Francis, 2004.
5. Sadler JZ, Hulgus YF. Clinical problem solving and the biopsychosocial model. Am J Psychiatry 1992;149:1315-23.
6. Coghill D, Sonuga-Barke EJ. Annual research review: Categories versus dimensions in the classification and conceptualisation of child and adolescent mental disorders – implications of recent empirical study. J Child Psychol Psychiatry 2012;53:469-89.
7. Slade M. What outcomes to measure in routine mental health services, and how to assess them: a systematic review. Aust N Z J Psychiatry 2002;36:743-53.
8. Lousada J. Mental health: moving beyond remission of symptoms. Lancet Psychiatry 2015;2:e6-7.
9. Bejerholm U, Roe D. Personal recovery within positive psychiatry. Nord J Psychiatry 2018;72:420-30.
10. Ustün B, Kennedy C. What is "functional impairment"? Disentangling disability from clinical significance. World Psychiatry 2009;8:82-5.
11. World Health Organization. International Classification of Functioning, Disability and Health. www.who.int.
12. Swildens WE, Visser E, Bähler M et al. Functional recovery of individuals with serious mental illnesses: development and testing of a new short instrument for routine outcome monitoring. Psychiatr Rehabil J 2018;41:341-50.
13. Harvey PD, Bellack AS. Toward a terminology for functional recovery in schizophrenia: is functional remission a viable concept? Schizophr Bull 2009;35:300-6.
14. Schipper H. Quality of life: principles of the clinical paradigm. J Psychosoc Oncol 1990;8:171-85.
15. World Health Organization. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med 1995;41:1403-9.
16. Pequeno NPF, Cabral NLA, Marchioni DM et al. Quality of life assessment instruments for adults: a systematic review of population-based studies. Health Qual Life Outcomes 2020;18:208.
17. Brown J, Brown A, Flynn T. Models of quality of life. a taxonomy, overview and systematic review of the literature. London: European Forum on Population Ageing Research, 2004.
18. Philips D. Quality of life: concept, policy and practice. New York: Routledge, 2006.
19. World Health Organization. Promoting mental health: concepts, emerging evidence, practice. A report. Geneva: World Health Organization, 2005.
20. Huppert FA, So TT. Flourishing across Europe: application of a new conceptual framework for defining well-being. Soc Indic Res 2013;110:837-61.
21. Chida Y, Steptoe A. Positive psychological well-being and mortality: a quantitative review of prospective observational studies. Psychosom Med 2008;70:741-56.
22. Diener E, Seligman ME. Beyond money: toward an economy of well-being.

Psychol Sci Public Interest 2004;5:1-31.
23. Boodoo R, Saunders EFH, Thompson KS et al. Recovery at 30: perspectives from psychiatry clinicians and senior faculty. Community Ment Health J 2025; 61:113241.
24. World Health Organization. Guidance on community mental health services: promoting person-centred and rights-based approaches. Geneva: World Health Organization, 2021.
25. Leamy M, Bird V, Le Boutillier C et al. Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. Br J Psychiatry 2011;199:445-52.
26. Correll CU, Ismail Z, McIntyre RS et al. Patient functioning and life engagement: unmet needs in major depressive disorder and schizophrenia. J Clin Psychiatry 2022;83:LU21112AH1.
27. Correll CU, Ismail Z, McIntyre RS et al. Patient functioning, life engagement, and treatment goals in schizophrenia. J Clin Psychiatry 2022;83:LU21112AH2.
28. Correll CU, Ismail Z, McIntyre RS et al. Patient functioning, life engagement, and treatment goals in major depressive disorder. J Clin Psychiatry 2022; 83:LU21112AH3.
29. Castelein S, van der Gaag M, Bruggeman R et al. Measuring empowerment among people with psychotic disorders: a comparison of three instruments. Psychiatr Serv 2008;59:1338-42.
30. Herth K. Abbreviated instrument to measure hope: development and psychometric evaluation. J Adv Nurs 1992;17:1251-9.
31. Mainz J, Kristensen S, Roe D. The power of the patient's voice in the modern health care system. Int J Qual Health Care 2022;34:ii1-2.
32. Fusar-Poli P, Estradé A, Stanghellini G et al. The lived experience of psychosis: a bottom-up review co-written by experts by experience and academics. World Psychiatry 2022;21:168-88.
33. Fusar-Poli P, Estradé A, Esposito CM et al. The lived experience of mental disorders in adolescents: a bottom-up review co-designed, co-conducted and co-written by experts by experience and academics. World Psychiatry 2024;23:191-208.
34. US Food and Drug Administration. Patient-reported outcome measures: use in medical product development to support labeling claims. www.fda.gov.
35. McIntyre RS, Ismail Z, Watling CP et al. Patient-reported outcome measures for life engagement in mental health: a systematic review. J Patient Rep Outcomes 2022;6:62.
36. Roe D, Mazor Y, Gelkopf M. Patient-reported outcome measurements (PROMs) and provider assessment in mental health: a systematic review of the context of implementation. Int J Qual Health Care 2021;34:ii28-39.
37. McKenzie E, Matkin L, Sousa Fialho L et al. Developing an international standard set of patient-reported outcome measures for psychotic disorders. Psychiatr Serv 2022;73:249-58.
38. Torous J, Bucci S, Bell IH et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021;20:318-35.
39. Torous J, Myrick K, Aguilera A. The need for a new generation of digital mental health tools to support more accessible, effective and equitable care. World Psychiatry 2023;22:1-2.
40. Sükei E, Norbury A, Perez-Rodriguez MM et al. Predicting emotional states using behavioral markers derived from passively sensed data: data-driven machine learning approach. JMIR Mhealth Uhealth 2021;9:e24465.
41. van Os J, Verhagen S, Marsman A et al. The experience sampling method as an mHealth tool to support self-monitoring, self-insight, and personalized health care in clinical practice. Depress Anxiety 2017;34:481-93.
42. Deakin E, Ng F, Young E et al. Design decisions and data completeness for experience sampling methods used in psychosis: systematic review. BMC Psychiatry 2022;22:669.
43. Stone AA, Schneider S, Smyth JM. Evaluation of pressing issues in ecological momentary assessment. Annu Rev Clin Psychol 2023;19:107-31.
44. Martin-Key NA, Spadaro B, Funnell E et al. The current state and validity of digital assessment tools for psychiatry: systematic review. JMIR Ment Health 2022;9:e32824.
45. Maj M. Understanding depression beyond the "mind-body" dichotomy. World Psychiatry 2023;22:349-50.
46. Baldwin H, Loebel-Davidsohn L, Oliver D et al. Real-world implementation of precision psychiatry: a systematic review of barriers and facilitators. Brain Sci 2022;12:934.
47. Kessler RC, Luedtke A. Pragmatic precision psychiatry – a new direction for optimizing treatment selection. JAMA Psychiatry 2021;78:1384-90.
48. Salazar de Pablo G, Studerus E, Vaquerizo-Serrano J et al. Implementing precision psychiatry: a systematic review of individualized prediction models for clinical practice. Schizophr Bull 2021;47:284-97.
49. Chekroud AM, Zotti RJ, Shehzad Z et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. Lancet Psychiatry 2016; 3:243-50.
50. Iniesta R, Malki K, Maier W et al. Combining clinical variables to optimize prediction of antidepressant treatment outcomes. J Psychiatr Res 2016;78:94-102.
51. Maj M, Stein DJ, Parker G et al. The clinical characterization of the adult patient with depression aimed at personalization of management. World Psychiatry 2020;19:269-93.
52. McIntyre RS, Alda M, Baldessarini RJ et al. The clinical characterization of the adult patient with bipolar disorder aimed at personalization of management. World Psychiatry 2022;21:364-87.
53. Stein DJ, Craske MG, Rothbaum BO et al. The clinical characterization of the adult patient with an anxiety or related disorder aimed at personalization of management. World Psychiatry 2021;20:336-56.
54. Lynch JM, Askew DA, Mitchell GK et al. Beyond symptoms: defining primary care mental health clinical assessment priorities, content and process. Soc Sci Med 2012;74:143-9.
55. Endicott J, Spitzer RL, Fleiss JL et al. The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. Arch Gen Psychiatry 1976;33:766-71.
56. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 3rd ed., revised. Washington: American Psychiatric Association, 1987.
57. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed., Washington: American Psychiatric Publishing, 1994.
58. Goldman HH, Skodol AE, Lave TR. Revising axis V for DSM-IV: a review of measures of social functioning. Am J Psychiatry 1992;149:1148-56.
59. Morosini PL, Magliano L, Brambilla L et al. Development, reliability and acceptability of a new version of the DSM-IV Social and Occupational Functioning Assessment Scale (SOFAS) to assess routine social functioning. Acta Psychiatr Scand 2000;101:323-9.
60. Lam RW, Filteau MJ, Milev R. Clinical effectiveness: the importance of psychosocial functioning outcomes. J Affect Disord 2011;132(Suppl. 1):S9-13.
61. Sheehan D. The anxiety disease. New York: Charles Scribner's Sons, 1983.
62. Leon AC, Olfson M, Portera L et al. Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. Int J Psychiatry Med 1997;27:93-105.
63. Schneider LC, Struening EL. SLOF: a behavioral rating scale for assessing the mentally ill. Soc Work Res Abstr 1983;19:9-21.
64. Bowie CR, Reichenberg A, Patterson TL et al. Determinants of real-world functional performance in schizophrenia subjects: correlations with cognition, functional capacity, and symptoms. Am J Psychiatry 2006;163:418-25.
65. Tarlov AR, Ware JE Jr, Greenfield S et al. The Medical Outcomes Study. An application of methods for monitoring the results of medical care. JAMA 1989;262:925-30.
66. Ware JE Jr, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-83.
67. McHorney CA, Ware JE Jr, Lu JF et al. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994;32:40-66.
68. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220-33.
69. Salyers MP, Bosworth HB, Swanson JW et al. Reliability and validity of the SF-12 health survey among people with severe mental illness. Med Care 2000;38: 1141-50.
70. Weissman M. Social Adjustment Scale - Self Report technical manual: Toronto: Multi-Health Systems, 1999.
71. Weissman MM. Social Adjustment Scale - Self-Report: Short & Social Adjustment Scale - Self-Report: screener technical manual. Toronto: Multi-health Systems, 2007.
72. Gameroff MJ, Wickramaratne P, Weissman MM. Testing the Short and Screener versions of the Social Adjustment Scale - Self-Report (SAS-SR). Int J Methods Psychiatr Res 2012;21:52-65.
73. Auther A, Smith C, Cornblatt BA. Global Functioning: Social Scale (GF: social). Glen Oaks: Zucker Hillside Hospital, 2006.
74. Piskulic D, Addington J, Auther A et al. Using the Global Functioning Social and Role scales in a first-episode sample. Early Interv Psychiatry 2011;5:219-23.
75. Carrión RE, Auther AM, McLaughlin D et al. The Global Functioning: Social and Role Scales – further validation in a large sample of adolescents and young adults at clinical high risk for psychosis. Schizophr Bull 2019;45:763-72.
76. Cornblatt BA, Auther AM, Niendam T et al. Preliminary findings for two new measures of social and role functioning in the prodromal phase of schizophre-

nia. Schizophr Bull 2007;33:688-702.
77. Nasrallah H, Morosini P, Gagnon DD. Reliability, validity and ability to detect change of the Personal and Social Performance Scale in patients with stable schizophrenia. Psychiatry Res 2008;161:213-24.
78. Mundt JC, Marks IM, Shear MK et al. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. Br J Psychiatry 2002;180:461-4.
79. Zahra D, Qureshi A, Henley W et al. The Work and Social Adjustment Scale: reliability, sensitivity and value. Int J Psychiatry Clin Pract 2014;18:131-8.
80. Rosa AR, Sánchez-Moreno J, Martínez-Aran A et al. Validity and reliability of the Functioning Assessment Short Test (FAST) in bipolar disorder. Clin Pract Epidemiol Ment Health 2007;3:5.
81. US National Institutes of Health. Patient-Reported Outcomes Measurement Information System (PROMIS). https://commonfund.nih.gov/promis/index.
82. Bingham CO 3rd, Bartlett SJ, Merkel PA et al. Using patient-reported outcomes and PROMIS in research and clinical applications: experiences from the PCORI pilot projects. Qual Life Res 2016;25:2109-16.
83. Obbarius A, Klapproth CP, Liegl G et al. Measuring PROMIS pain interference in German patients with chronic conditions: calibration, validation, and cross-cultural use of item parameters. Qual Life Res 2023;32:2839-52.
84. Terwee CB, Roorda LD, de Vet HC et al. Dutch-Flemish translation of 17 item banks from the Patient-Reported Outcomes Measurement Information System (PROMIS). Qual Life Res 2014;23:1733-41.
85. Rose M, Bjorner JB, Becker J et al. Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). J Clin Epidemiol 2008;61:17-33.
86. van Muilekom MM, Luijten MAJ, van Oers HA et al. From statistics to clinics: the visual feedback of PROMIS® CATs. J Patient Rep Outcomes 2021;5:55.
87. Hays RD, Spritzer KL, Schalet BD et al. PROMIS®-29 v2.0 profile physical and mental health summary scores. Qual Life Res 2018;27:1885-91.
88. HealthMeasures. PROMIS®. www.healthmeasures.net.
89. World Health Organization. Measuring health and disability: manual for WHO disability assessment schedule WHODAS 2.0. Geneva: World Health Organization, 2010.
90. Axelsson E, Lindsäter E, Ljótsson B et al. The 12-item Self-Report World Health Organization Disability Assessment Schedule (WHODAS) 2.0 administered via the Internet to individuals with anxiety and stress disorders: a psychometric investigation based on data from two clinical trials. JMIR Ment Health 2017;4:e58.
91. Bovin MJ, Meyer EC, Kimbrel NA et al. Using the World Health Organization Disability Assessment Schedule 2.0 to assess disability in veterans with posttraumatic stress disorder. PLoS One 2019;14:e0220806.
92. MacLeod MA, Tremblay PF, Graham K et al. Psychometric properties and a latent class analysis of the 12-item World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) in a pooled dataset of community samples. Int J Methods Psychiatr Res 2016;25:243-54.
93. Penninx BW, Pine DS, Holmes EA et al. Anxiety disorders. Lancet 2021;397:914-27.
94. Domschke K, Seuling PD, Schiele MA et al. The definition of treatment resistance in anxiety disorders: a Delphi method-based consensus guideline. World Psychiatry 2024;23:113-23.
95. Obbarius A, van Maasakkers L, Baer L et al. Standardization of health outcomes assessment for depression and anxiety: recommendations from the ICHOM Depression and Anxiety Working Group. Qual Life Res 2017;26:3211-25.
96. Katzman MA, Bleau P, Blier P et al. Canadian clinical practice guidelines for the management of anxiety, posttraumatic stress and obsessive-compulsive disorders. BMC Psychiatry 2014;14(Suppl. 1):S1.
97. European Medicines Agency. Clinical investigation of medicinal products indicated for panic disorder – Scientific guideline. https://www.ema.europa.eu.
98. European Medicines Agency. Clinical investigation of medicinal products indicated for generalised anxiety disorder – Scientific guideline. https://www.ema.europa.eu.
99. National Institute for Health and Care Excellence. Generalised anxiety disorder and panic disorder in adults: management. Clinical guideline. www.nice.org.uk.
100. National Institute for Health and Care Excellence. Social anxiety disorder: recognition, assessment and treatment. Clinical guideline. https://www.ema.europa.eu.
101. Schneier FR, Heckelman LR, Garfinkel R et al. Functional impairment in social phobia. J Clin Psychiatry 1994;55:322-31.
102. Sheehan DV, Harnett-Sheehan K, Spann ME et al. Assessing remission in major depressive disorder and generalized anxiety disorder clinical trials with the discan metric of the Sheehan Disability Scale. Int Clin Psychopharmacol 2011;26:75-83.
103. Gérard A, Liard F, Crochard A et al. Disability in patients consulting for anxiety or mood disorders in primary care: response to antidepressant treatment. Neuropsychiatr Dis Treat 2012;8:605-14.
104. McIntyre RS, Berk M, Brietzke E et al. Bipolar disorders. Lancet 2020;396:1841-56.
105. Yatham LN, Kennedy SH, Parikh SV et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) 2018 guidelines for the management of patients with bipolar disorder. Bipolar Disord 2018;20:97-170.
106. Malhi GS, Bell E, Boyce P et al. The 2020 Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders: bipolar disorder summary. Bipolar Disord 2020;22:805-21.
107. Strejilevich SA, Martino DJ, Murru A et al. Mood instability and functional recovery in bipolar disorders. Acta Psychiatr Scand 2013;128:194-202.
108. Mezes B, Lobban F, Costain D et al. Recovery beyond clinical improvement – Recovery outcomes measured for people with bipolar disorder between 1980 and 2020. J Affect Disord 2022;309:375-92.
109. Montejo L, Orhan M, Chen P et al. Functioning in older adults with bipolar disorder: a report on recommendations by the International Society of Bipolar Disorder (ISBD) older adults with bipolar disorder (OABD) task force. Bipolar Disord 2023;25:457-68.
110. Sheehan DV, Harnett-Sheehan K, Raj BA. The measurement of disability. Int Clin Psychopharmacol 1996;11(Suppl. 3):89-95.
111. Pfeffer RI, Kurosaki TT, Harrah CH Jr et al. Measurement of functional activities in older adults in the community. J Gerontol 1982;37:323-9.
112. Gélinas I, Gauthier L, McIntyre M et al. Development of a functional measure for persons with Alzheimer's disease: the Disability Assessment for Dementia. Am J Occup Ther 1999;53:471-81.
113. Galasko D, Bennett D, Sano M et al. An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. The Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord 1997;11(Suppl. 2):S33-9.
114. Sikkes SA, Knol DL, Pijnenburg YA et al. Validation of the Amsterdam IADL Questionnaire©, a new tool to measure instrumental activities of daily living in dementia. Neuroepidemiology 2013;41:35-41.
115. Sikkes SA, Pijnenburg YA, Knol DL et al. Assessment of instrumental activities of daily living in dementia: diagnostic value of the Amsterdam Instrumental Activities of Daily Living Questionnaire. J Geriatr Psychiatry Neurol 2013;26:244-50.
116. Jutten RJ, Peeters CFW, Leijdesdorff SMJ et al. Detecting functional decline from normal aging to dementia: development and validation of a short version of the Amsterdam IADL Questionnaire. Alzheimers Dement 2017;8:26-35.
117. Rodriguez-Porcel F, Wyman-Chick KA, Abdelnour Ruiz C et al. Clinical outcome measures in dementia with Lewy bodies trials: critique and recommendations. Transl Neurodegener 2022;11:24.
118. Solmi M, Monaco F, Højlund M et al. Outcomes in people with eating disorders: a transdiagnostic and disorder-specific systematic review, meta-analysis and multivariable meta-regression analysis. World Psychiatry 2024;23:124-38.
119. Rome ES, Ammerman S. Medical complications of eating disorders: an update. J Adolesc Health 2003;33:418-26.
120. Ackard DM, Richter S, Egan A et al. The meaning of (quality of) life in patients with eating disorders: a comparison of generic and disease-specific measures across diagnosis and outcome. Int J Eat Disord 2014;47:259-67.
121. Quiles Marcos Y, Ruiz Maciá Á, Manchón López J et al. Validation of the Spanish version of the Eating Disorders Quality of Life instrument (EDQOL). J Eat Disord 2023;11:103.
122. Maraldo TM, Fewell L, Vander Wal JS. Factor structure and psychometric properties of the Clinical Impairment Assessment 3.0 (CIA) in a clinical eating disorder sample. Eat Behav 2021;40:101469.
123. Crone C, Fochtmann LJ, Attia E et al. The American Psychiatric Association practice guideline for the treatment of patients with eating disorders. Am J Psychiatry 2023;180:167-71.
124. Fairburn CG, Cooper Z. The eating disorder examination (12th ed.). In: Fairburn CG, Wilson GT (eds). Binge eating: nature, assessment and treatment. New York: Guilford, 1993:317-56.
125. Bachner-Melman R, Lev-Ari L, Zohar AH et al. The Eating Disorders Recovery Questionnaire: psychometric properties and validity. Eat Weight Disord 2021;26:2633-43.
126. Chevance A, Ravaud P, Tomlinson A et al. Identifying outcomes for depression that matter to patients, informal caregivers, and health-care professionals: qualitative content analysis of a large international online survey. Lancet Psychiatry 2020;7:692-702.

127. McKnight PE, Kashdan TB. The importance of functional impairment to mental health outcomes: a case for reassessing our goals in depression treatment research. Clin Psychol Rev 2009;29:243-59.
128. Collard RM, Wassink-Vossen S, Schene AH et al. Symptomatic and functional recovery in depression in later life. Soc Psychiatry Psychiatr Epidemiol 2018;53:1071-9.
129. Kamenov K, Cabello M, Coenen M et al. How much do we know about the functional effectiveness of interventions for depression? A systematic review. J Affect Disord 2015;188:89-96.
130. Grootenboer EM, Giltay EJ, van der Lem R et al. Reliability and validity of the Global Assessment of Functioning Scale in clinical outpatients with depressive disorders. J Eval Clin Pract 2012;18:502-7.
131. Abdin E, Seet V, Jeyagurunathan A et al. Validation of the 12-item World Health Organization Disability Assessment Schedule 2.0 in individuals with schizophrenia, depression, anxiety, and diabetes in Singapore. PLoS One 2023;18: e0294908.
132. Chiang YC, Liou TH, Lee HC et al. Using WHODAS 2.0 to assess functional impairment in people with depression: should employment receive more attention? Int J Environ Res Public Health 2021;18:4552.
133. Guico-Pabia CJ, Fayyad RS, Soares CN. Assessing the relationship between functional impairment/recovery and depression severity: a pooled analysis. Int Clin Psychopharmacol 2012;27:1-7.
134. McIntyre RS, Alsuwaidan M, Baune BT et al. Treatment-resistant depression: definition, prevalence, detection, management, and investigational interventions. World Psychiatry 2023;22:394-412.
135. Rush AJ, Sackeim HA, Conway CR et al. Clinical research challenges posed by difficult-to-treat depression. Psychol Med 2022;52:419-32.
136. Subramaniam M, Soh P, Vaingankar JA et al. Quality of life in obsessive-compulsive disorder: impact of the disorder and of treatment. CNS Drugs 2013;27:367-83.
137. Koran LM, Hanna GL, Hollander E et al. Practice guideline for the treatment of patients with obsessive-compulsive disorder. Am J Psychiatry 2007;164:5-53.
138. European Medicines Agency. Clinical investigation of medicinal products for the treatment of obsessive compulsive disorder – Scientific guideline. https://www.ema.europa.eu.
139. Hollander E, Stein DJ, Fineberg NA et al. Quality of life outcomes in patients with obsessive-compulsive disorder: relationship to treatment response and symptom relapse. J Clin Psychiatry 2010;71:784-92.
140. Tyrer P, Reed GM, Crawford MJ. Classification, assessment, prevalence, and effect of personality disorder. Lancet 2015;385:717-26.
141. National Institute for Health and Care Excellence. Borderline personality disorder: recognition and management. Clinical guideline. www.nice.org.uk.
142. National Institute for Health and Care Excellence. Personality disorders: borderline and antisocial – Quality standard. www.nice.org.uk.
143. Chanen A, Sharp C, Hoffman P. Prevention and early intervention for borderline personality disorder: a novel public health priority. World Psychiatry 2017; 16:215-6.
144. Jørgensen CR, Freund C, Bøye R et al. Outcome of mentalization-based and supportive psychotherapy in patients with borderline personality disorder: a randomized trial. Acta Psychiatr Scand 2013;127:305-17.
145. Paris J, Zweig-Frank H. A 27-year follow-up of patients with borderline personality disorder. Compr Psychiatry 2001;42:482-7.
146. Gunderson JG, Stout RL, McGlashan TH et al. Ten-year course of borderline personality disorder: psychopathology and function from the Collaborative Longitudinal Personality Disorders study. Arch Gen Psychiatry 2011;68:827-37.
147. Pedersen G, Kvarstein EH, Wilberg T. The Work and Social Adjustment Scale: psychometric properties and validity among males and females, and outpatients with and without personality disorders. Personal Ment Health 2017; 11:215-28.
148. Blagov PS, Bi W, Shedler J et al. The Shedler-Westen Assessment Procedure (SWAP): evaluating psychometric questions about its reliability, validity, and impact of its fixed score distribution. Assessment 2012;19:370-82.
149. De Fife JA, Malone JC, DiLallo J et al. Assessing adolescent personality disorders with the Shedler-Westen Assessment Procedure for Adolescents. Clin Psychol Sci Pract 2013;20:4:393-407.
150. Cogan R, Porcerelli JH. Test-retest reliability and discriminant validity of the SWAP-200 in a psychoanalytic treatment sample. Psychol Psychother 2012; 85:36-47.
151. Shedler J. Integrating clinical and empirical perspectives on personality: the Shedler-Westen Assessment Procedure (SWAP). In: Huprich SK (ed). Personality disorders: toward theoretical and empirical integration in diagnosis and assessment. Washington: American Psychological Association, 2015:225-52.
152. Westen D, Waller NG, Shedler J et al. Dimensions of personality and personality pathology: factor structure of the Shedler-Westen Assessment Procedure-II (SWAP-II). J Pers Disord 2014;28:281-318.
153. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 5th ed., text revision. Washington: American Psychiatric Association, 2022.
154. Bender DS, Skodol AE, First MB. Structured Clinical Interview for the DSM-5® Alternative Model for Personality Disorders (SCID-5-AMPD), Module I: Level of Personality Functioning Scale. Washington: American Psychiatric Publishing, 2018.
155. Hutsebaut J, Kamphuis JH, Feenstra DJ et al. Assessing DSM-5-oriented level of personality functioning: development and psychometric evaluation of the Semi-Structured Interview for Personality Functioning DSM-5 (STiP-5.1). Personal Disord 2017;8:94-101.
156. Thylstrup B, Simonsen S, Nemery C et al. Assessment of personality-related levels of functioning: a pilot study of clinical assessment of the DSM-5 level of personality functioning based on a semi-structured interview. BMC Psychiatry 2016;16:298.
157. Morey LC, McCredie MN, Bender DS et al. Criterion A: level of personality functioning in the alternative DSM-5 model for personality disorders. Personal Disord 2022;13:305-15.
158. Birkhölzer M, Schmeck K, Goth K. Assessment of Criterion A. Curr Opin Psychol 2021;37:98-103.
159. Sharp C, Wall K. DSM-5 level of personality functioning: refocusing personality disorder on what it means to be human. Annu Rev Clin Psychol 2021;17:313-37.
160. Waugh MH, McClain CM, Mariotti EC et al. Comparative content analysis of self-report scales for level of personality functioning. J Pers Assess 2021;103:161-73.
161. Benfer N, Litz BT. Assessing and addressing functioning and quality of life in PTSD. Curr Treat Options Psych 2023;10:1-20.
162. Rodriguez P, Holowka DW, Marx BP. Assessment of posttraumatic stress disorder-related functional impairment: a review. J Rehabil Res Dev 2012;49: 649-65.
163. Patel AR, Hall BJ. Beyond the DSM-5 diagnoses: a cross-cultural approach to assessing trauma reactions. Focus 2021;19:197-203.
164. Marx BP, Wolf EJ, Cornette MM et al. Using the WHODAS 2.0 to assess functioning among veterans seeking compensation for posttraumatic stress disorder. Psychiatr Serv 2015;66:1312-7.
165. Mitchell JM, Ot'alora GM, van der Kolk B et al. MDMA-assisted therapy for moderate to severe PTSD: a randomized, placebo-controlled phase 3 trial. Nat Med 2023;29:2473-80.
166. Jovanović AA, Gasić MJ, Ivković M et al. Reliability and validity of DSM-IV Axis V scales in a clinical sample of veterans with posttraumatic stress disorder. Psychiatr Danub 2008;20:286-300.
167. McCaslin SE, Maguen S, Metzler T et al. Assessing posttraumatic stress related impairment and well-being: the Posttraumatic Stress Related Functioning Inventory (PRFI). J Psychiatr Res 2016;72:104-11.
168. Bovin MJ, Black SK, Rodriguez P et al. Development and validation of a measure of PTSD-related psychosocial functional impairment: the Inventory of Psychosocial Functioning. Psychol Serv 2018;15:216-29.
169. Kleiman SE, Bovin MJ, Black SK et al. Psychometric properties of a brief measure of posttraumatic stress disorder-related impairment: the Brief Inventory of Psychosocial Functioning. Psychol Serv 2020;17:187-94.
170. Leifker FR, Patterson TL, Heaton RK et al. Validating measures of real-world outcome: the results of the VALERO expert survey and RAND panel. Schizophr Bull 2011;37:334-43.
171. Long M, Stansfeld JL, Davies N et al. A systematic review of social functioning outcome measures in schizophrenia with a focus on suitability for intervention research. Schizophr Res 2022;241:275-91.
172. Sabbag S, Twamley EM, Vella L et al. Assessing everyday functioning in schizophrenia: not all informants seem equally informative. Schizophr Res 2011;131:250-5.
173. Sabbag S, Twamley EW, Vella L et al. Predictors of the accuracy of self assessment of everyday functioning in people with schizophrenia. Schizophr Res 2012;137:190-5.
174. Gould F, Sabbag S, Durand D et al. Self-assessment of functional ability in schizophrenia: milestone achievement and its relationship to accuracy of self-evaluation. Psychiatry Res 2013;207:19-24.
175. Durand D, Strassnig M, Sabbag S et al. Factors influencing self-assessment of cognition and functioning in schizophrenia: implications for treatment studies. Eur Neuropsychopharmacol 2015;25:185-91.
176. Harvey PD, Sabbag S, Prestia D et al. Functional milestones and clinician rat-

ings of everyday functioning in people with schizophrenia: overlap between milestones and specificity of ratings. J Psychiatr Res 2012;46:1546-52.
177. Mucci A, Galderisi S, Gibertoni D et al. Factors associated with real-life functioning in persons with schizophrenia in a 4-year follow-up study of the Italian Network for Research on Psychoses. JAMA Psychiatry 2021;78:550-9.
178. Mucci A, Rucci P, Rocca P et al. The Specific Level of Functioning Scale: construct validity, internal consistency and factor structure in a large Italian sample of people with schizophrenia living in the community. Schizophr Res 2014; 159:144-50.
179. Sumiyoshi T, Nishida K, Niimura H et al. Cognitive insight and functional outcome in schizophrenia; a multi-center collaborative study with the Specific Level of Functioning Scale - Japanese version. Schizophr Res Cogn 2016;6:9-14.
180. Birchwood M, Smith J, Cochrane R et al. The Social Functioning Scale. The development and validation of a new scale of social adjustment for use in family intervention programmes with schizophrenic patients. Br J Psychiatry 1990;157:853-9.
181. Volkow ND, Jones EB, Einstein EB et al. Prevention and treatment of opioid misuse and addiction: a review. JAMA Psychiatry 2019;76:208-16.
182. Volkow ND, Blanco C. Substance use disorders: a comprehensive update of classification, epidemiology, neurobiology, clinical aspects, treatment and prevention. World Psychiatry 2023;22:203-29.
183. Daeppen JB, Krieg MA, Burnand B et al. MOS-SF-36 in evaluating health-related quality of life in alcohol-dependent patients. Am J Drug Alcohol Abuse 1998;24:685-94.
184. Richter D, Eikelmann B, Berger K. Use of the SF-36 in the evaluation of a drug detoxification program. Qual Life Res 2004;13:907-14.
185. Buchholz A, Krol A, Rist F et al. An assessment of factorial structure and health-related quality of life in problem drug users using the Short Form 36 Health Survey. Qual Life Res 2008;17:1021-9.
186. McLellan AT, Cacciola JC, Alterman AI et al. The Addiction Severity Index at 25: origins, contributions and transitions. Am J Addict 2006;15:113-24.
187. McLellan AT, Luborsky L, Woody GE et al. An improved diagnostic evaluation instrument for substance abuse patients. The Addiction Severity Index. J Nerv Ment Dis 1980;168:26-33.
188. Daeppen JB, Burnand B, Schnyder C et al. Validation of the Addiction Severity Index in French-speaking alcoholic patients. J Stud Alcohol 1996;57:585-90.
189. Zapata JP, García J, Arroyave CA et al. Validation of the sixth version of the Addiction Severity Index (ASI-6) for patients in a clinical population of Colombia. Biomedica 2019;39:385-404.
190. Mäkelä K. Studies of the reliability and validity of the Addiction Severity Index. Addiction 2004;99:398-410.
191. Krenz S, Dieckmann S, Favrat B et al. French version of the Addiction Severity Index (5th edition): validity and reliability among Swiss opiate-dependent patients. Eur Addict Res 2004;10:173-9.
192. van Krugten FCW, Feskens K, Busschbach JJV et al. Instruments to assess quality of life in people with mental health problems: a systematic review and dimension analysis of generic, domain- and disease-specific instruments. Health Qual Life Outcomes 2021;19:249.
193. Connell J, O'Cathain A, Brazier J. Measuring quality of life in mental health: are we asking the right questions? Soc Sci Med 2014;120:12-20.
194. EuroQol Group. EuroQol – a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199-208.
195. EuroQol Research Foundation. EQ-5D-5L user guide, 2019. https://euroqol.org/publications/user-guides.
196. Feng YS, Kohlmann T, Janssen MF et al. Psychometric properties of the EQ-5D-5L: a systematic review of the literature. Qual Life Res 2021;30:647-73.
197. EuroQoL. Proxy version 1. https://euroqol.org/faq/proxy-version-1.
198. EuroQoL. EQ-5D-3L. https://euroqol.org/information-and-support/euroqol-instruments/eq-5d-3l.
199. Zare F, Ameri H, Madadizadeh F et al. Validity and reliability of the EQ-5D-3L (a generic preference-based instrument used for calculating quality-adjusted life-years) for patients with type 2 diabetes in Iran. Diabetes Metab Syndr 2021;15:319-24.
200. Frisch MC, Cornell J, Villanueva M et al. Clinical validation of the Quality of Life Inventory. A measure of life satisfaction for use in treatment planning and outcome assessment. Psychol Assess 1992;4:1:92-101.
201. World Health Organization. WHOQOL-100. www.who.int.
202. Carpiniello B, Lai GL, Pariante CM et al. Symptoms, standards of living and subjective quality of life: a comparative study of schizophrenic and depressed out-patients. Acta Psychiatr Scand 1997;96:235-41.
203. Orsel S, Akdemir A, Dağ I. The sensitivity of quality-of-life scale WHOQOL-100 to psychopathological measures in schizophrenia. Compr Psychiatry 2004;45:57-61.
204. The WHOQOL Group. Development of the World Health Organization WHOQOL-BREF quality of life assessment. Psychol Med 1998;28:551-8.
205. Akanni OO, Edeh AN, Agbir MT et al. The quality of life and its inter-relationship with posttraumatic stress disorder and social support in two post-conflict communities in Nigeria. J Health Psychol 2024; 29:1519-32.
206. Priebe S, Huxley P, Knight S et al. Application and results of the Manchester Short Assessment of Quality of Life (MANSA). Int J Soc Psychiatry 1999;45:7-12.
207. Björkman T, Svensson B. Quality of life in people with severe mental illness. Reliability and validity of the Manchester Short Assessment of Quality of Life (MANSA). Nord J Psychiatry 2005;59:302-6.
208. Linton MJ, Dieppe P, Medina-Lara A. Review of 99 self-report measures for assessing well-being in adults: exploring dimensions of well-being and developments over time. BMJ Open 2016;6:e010641.
209. International Wellbeing Group. Personal Wellbeing Index - Adult, 6th ed: Version 2. Geelong: Australian Centre on Quality of Life, School of Psychology, Deakin University, 2024.
210. Sulaiman-Hill CM, Thompson SC. Selecting instruments for assessing psychological wellbeing in Afghan and Kurdish refugee groups. BMC Res Notes 2010;3:237.
211. Kaplan RM, Sieber WJ, Ganiats TG. The Quality of Well-Being Scale: comparison of the interviewer-administered version with a self-administered questionnaire. Psychol Health 1997;12:783-91.
212. Child Outcomes Research Consortium. The World Health Organisation - Five Well-Being Index (WHO-5). www.corc.uk.net.
213. Lara-Cabrera ML, Mundal IP, De Las Cuevas C. Patient-reported well-being: psychometric properties of the World Health Organization Well-Being Index in specialised community mental health settings. Psychiatry Res 2020;291: 113268.
214. Topp CW, Østergaard SD, Søndergaard S et al. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 2015;84:167-76.
215. Diener E, Lucas RE, Oishi S. Subjective well-being: the science of happiness and life satisfaction. In: Lopez SSC (ed). Oxford handbook of positive psychology. New York: Oxford University Press, 2009:69-73.
216. Meule A, Voderholzer U. Life satisfaction in persons with mental disorders. Qual Life Res 2020;29:3043-52.
217. Diener E, Emmons RA, Larsen RJ et al. The Satisfaction With Life Scale. J Pers Assess 1985;49:71-5.
218. Arrindell WA, Meeuwesen L, Huyse FJ. The Satisfaction With Life Scale (SWLS): psychometric properties in a non-psychiatric medical outpatients sample. Personal Individ Diff 1991;12:117-23.
219. Besser A, Neria Y. PTSD symptoms, satisfaction with life, and prejudicial attitudes toward the adversary among Israeli civilians exposed to ongoing missile attacks. J Trauma Stress 2009;22:268-75.
220. Pavot W, Diener E, Colvin CR et al. Further validation of the Satisfaction With Life Scale: evidence for the cross-method convergence of well-being measures. J Pers Assess 1991;57:149-61.
221. Endicott J, Nee J, Harrison W et al. Quality of Life Enjoyment and Satisfaction Questionnaire: a new measure. Psychopharmacol Bull 1993;29:321-6.
222. Rossi A, Rucci P, Mauri M et al. Validity and reliability of the Italian version of the Quality of Life, Enjoyment and Satisfaction Questionnaire. Qual Life Res 2005;14:2323-8.
223. Medworks Media. Quality of Life Enjoyment and Satisfaction Questionnaire - Short-Form (Q-LES-Q-SF). https://medworksmedia.com.
224. Stevanovic D. Quality of Life Enjoyment and Satisfaction Questionnaire - short form for quality of life assessments in clinical practice: a psychometric study. J Psychiatr Ment Health Nurs 2011;18:744-50.
225. Majani G, Pierobon A, Giardni A et al. Satisfaction Profile (SAT-P) in 732 patients: focus on subjectivity in HRQoL assessment. Psychol Health 2000;15: 409-22.
226. Wilmer MT, Anderson K, Reynolds M. Correlates of quality of life in anxiety disorders: review of recent research. Curr Psychiatry Rep 2021;23:77.
227. McAlinden NM, Oei TP. Validation of the Quality of Life Inventory for patients with anxiety and depression. Compr Psychiatry 2006;47:307-14.
228. Lindner P, Andersson G, Ost LG et al. Validation of the internet-administered Quality of Life Inventory (QOLI) in different psychiatric conditions. Cogn Behav Ther 2013;42:315-27.
229. Freidl M, Wegerer M, Litvan Z et al. Determinants of quality of life improvements in anxiety and depressive disorders – A longitudinal study of inpatient psychotherapy. Front Psychiatry 2022;13:937194.
230. Pohl RB, Wolkow RM, Clary CM. Sertraline in the treatment of panic disorder: a double-blind multicenter trial. Am J Psychiatry 1998;155:1189-95.

231. Bonnín CDM, Reinares M, Martínez-Arán A et al. Improving functioning, quality of life, and well-being in patients with bipolar disorder. Int J Neuropsychopharmacol 2019;22:467-77.
232. Rosenblat JD, Simon GE, Sachs GS et al. Treatment effectiveness and tolerability outcomes that are most important to individuals with bipolar and unipolar depression. J Affect Disord 2019;243:116-20.
233. Bonnín CM, Yatham LN, Michalak EE et al. Psychometric properties of the Well-Being Index (WHO-5) Spanish version in a sample of euthymic patients with bipolar disorder. J Affect Disord 2018;228:153-9.
234. Michalak EE, Murray G. Development of the QoL.BD: a disorder-specific scale to assess quality of life in bipolar disorder. Bipolar Disord 2010;12:727-40.
235. Morton E, Murray G, Yatham LN et al. The Quality of Life in Bipolar Disorder (QoL.BD) questionnaire a decade on – a systematic review of the measurement of condition-specific aspects of quality of life in bipolar disorder. J Affect Disord 2021;278:33-45.
236. Smith S. Measuring health-related quality of life in dementia. In: Fisher WP Jr, Cano SJ (eds). Person-centered outcome metrology. Cham: Springer, 2022:35-52.
237. Cheung PWH, Wong CKH, Samartzis D et al. Psychometric validation of the EuroQoL 5-Dimension 5-Level (EQ-5D-5L) in Chinese patients with adolescent idiopathic scoliosis. Scoliosis Spinal Disord 2016;11:19.
238. Michalowsky B, Xie F, Kohlmann T et al. Acceptability and validity of the EQ-5D in patients living with dementia. Value Health 2020;23:760-7.
239. Rabins PV, Kasper JD, Kleinman L et al. Concepts and methods in the development of the ADRQL: an instrument for assessing health-related quality of life in persons with Alzheimer disease. J Mental Health Aging 1991;5:33-48.
240. Kasper JD, Black BS, Shore AD et al. Evaluation of the validity and reliability of the Alzheimer Disease-related Quality of Life Assessment Instrument. Alzheimer Dis Assoc Disord 2009;23:275-84.
241. Smith SC, Lamping DL, Banerjee S et al. Development of a new measure of health-related quality of life for people with dementia: DEMQOL. Psychol Med 2007;37:737-46.
242. Logsdon RG, Gibson LE, McCurry SM et al. Quality of life in Alzheimer's disease: patient and caregiver reports. J Mental Health Aging 1999;5:1:21-32.
243. Milic J, Stankic D, Stefanovic D. Eating disorder and quality of life. In: Patel VB, Preedy VR (eds). Eating disorders. Cham: Springer, 2023:353-65.
244. Adair CE, Marcoux GC, Cram BS et al. Development and multi-site validation of a new condition-specific quality of life measure for eating disorders. Health Qual Life Outcomes 2007;5:23.
245. Engel SG, Wittrock DA, Crosby RD et al. Development and psychometric validation of an eating disorder-specific health-related quality of life instrument. Int J Eat Disord 2006;39:62-71.
246. Las Hayas C, Quintana JM, Padierna A et al. The new questionnaire Health-Related Quality of Life for Eating Disorders showed good validity and reliability. J Clin Epidemiol 2006;59:192-200.
247. Las Hayas C, Quintana JM, Padierna JA et al. Health-Related Quality of Life for Eating Disorders questionnaire version-2 was responsive 1-year after initial assessment. J Clin Epidemiol 2007;60:825-33.
248. Zimmerman M, McGlinchey JB, Posternak MA et al. How should remission from depression be defined? The depressed patient's perspective. Am J Psychiatry 2006;163:148-50.
249. Cohen RM, Greenberg JM, IsHak WW. Incorporating multidimensional patient-reported outcomes of symptom severity, functioning, and quality of life in the Individual Burden of Illness Index for Depression to measure treatment impact and recovery in MDD. JAMA Psychiatry 2013;70:343-50.
250. Zimmerman M, Martinez J, Attiullah N et al. Why do some depressed outpatients who are not in remission according to the Hamilton Depression Rating Scale nonetheless consider themselves to be in remission? Depress Anxiety 2012;29:891-5.
251. Demyttenaere K, Kiekens G, Bruffaerts R et al. Outcome in depression (I): why symptomatic remission is not good enough. CNS Spectr 2021;26:393-9.
252. Sackeim HA, Rush AJ, Greco T et al. Alternative metrics for characterizing longer-term clinical outcomes in difficult-to-treat depression: I. Association with change in quality of life. Psychol Med 2023;53:1-13.
253. Krieger T, Zimmermann J, Huffziger S et al. Measuring depression with a well-being index: further evidence for the validity of the WHO Well-Being Index (WHO-5) as a measure of the severity of depression. J Affect Disord 2014;156:240-4.
254. Papaioannou D, Brazier J, Parry G. How to measure quality of life for cost-effectiveness analyses of personality disorders: a systematic review. J Pers Disord 2013;27:383-401.
255. Cramer V, Torgersen S, Kringlen E. Personality disorders and quality of life. A population study. Compr Psychiatry 2006;47:178-84.
256. Cramer V, Torgersen S, Kringlen E. Socio-demographic conditions, subjective somatic health, Axis I disorders and personality disorders in the common population: the relationship to quality of life. J Pers Disord 2007;21:552-67.
257. Chen H, Cohen P, Kasen S et al. Impact of adolescent mental disorders and physical illnesses on quality of life 17 years later. Arch Pediatr Adolesc Med 2006;160:93-9.
258. Soeteman DI, Verheul R, Busschbach JJ. The burden of disease in personality disorders: diagnosis-specific quality of life. J Pers Disord 2008;22:259-68.
259. Ikhtabi S, Pitman A, Toh G et al. The experience of loneliness among people with a "personality disorder" diagnosis or traits: a qualitative meta-synthesis. BMC Psychiatry 2022;22:130.
260. Ikhtabi S, Pitman A, Maconick L et al. The prevalence and severity of loneliness and deficits in perceived social support among who have received a 'personality disorder' diagnosis or have relevant traits: a systematic review. BMC Psychiatry 2024;24:21.
261. IsHak WW, Elbau I, Ismail A et al. Quality of life in borderline personality disorder. Harv Rev Psychiatry 2013;21:138-50.
262. Dokucu ME, Cloninger CR. Personality disorders and physical comorbidities: a complex relationship. Curr Opin Psychiatry 2019;32:435-41.
263. Chakhssi F, Zoet JM, Oostendorp JM et al. Effect of psychotherapy for borderline personality disorder on quality of life: a systematic review and meta-analysis. J Pers Disord 2021;35:255-69.
264. Oud M, Arntz A, Hermens ML et al. Specialized psychotherapies for adults with borderline personality disorder: a systematic review and meta-analysis. Aust N Z J Psychiatry 2018;52:949-61.
265. Weinberg M, Harel H, Shamani M et al. War and well-being: the association between forgiveness, social support, posttraumatic stress disorder, and well-being during and after war. Soc Work 2017;62:341-8.
266. Berle D, Hilbrink D, Russell-Williams C et al. Personal wellbeing in posttraumatic stress disorder (PTSD): association with PTSD symptoms during and following treatment. BMC Psychol 2018;6:7.
267. Pietrzak RH, Tsai J, Armour C et al. Functional significance of a novel 7-factor model of DSM-5 PTSD symptoms: results from the National Health and Resilience in Veterans study. J Affect Disord 2015;174:522-6.
268. Silverberg ND, Wojtowicz M, Bui E et al. Contribution of perceived cognitive functioning to quality of life in service members and veterans with posttraumatic stress disorder. J Trauma Stress 2017;30:318-22.
269. Jakšić N, Aukst Margetić B, Šimunović Filipčić I et al. Temperament, character, and subjective well-being in Croatian war veterans suffering from posttraumatic stress disorder. J Nerv Ment Dis 2020;208:340-3.
270. Avendaño Vásquez CJ, Aguilar Ticona JP, Lins-Kusterer L et al. Posttraumatic stress disorder in demobilized ex-combatants from illegal armed groups in Colombia: cross-sectional study. Psychol Trauma 2023;15:265-70.
271. Carlsson JM, Olsen DR, Mortensen EL et al. Mental health and health-related quality of life: a 10-year follow-up of tortured refugees. J Nerv Ment Dis 2006; 194:725-31.
272. Ben-Ezra M, Karatzias T, Hyland P et al. Posttraumatic stress disorder (PTSD) and complex PTSD (CPTSD) as per ICD-11 proposals: a population study in Israel. Depress Anxiety 2018;35:264-74.
273. Eack SM, Newhill CE. Psychiatric symptoms and quality of life in schizophrenia: a meta-analysis. Schizophr Bull 2007;33:1225-37.
274. Siu CO, Harvey PD, Agid O et al. Insight and subjective measures of quality of life in chronic schizophrenia. Schizophr Res Cogn 2015;2:127-32.
275. Strassnig M, Kotov R, Fochtmann L et al. Associations of independent living and labor force participation with impairment indicators in schizophrenia and bipolar disorder at 20-year follow-up. Schizophr Res 2018;197:150-5.
276. Oliveri LN, Awerbuch AW, Jarskog LF et al. Depression predicts self assessment of social function in both patients with schizophrenia and healthy people. Psychiatry Res 2020;284:112681.
277. Heinrichs DW, Hanlon TE, Carpenter WT Jr. The Quality of Life Scale: an instrument for rating the schizophrenic deficit syndrome. Schizophr Bull 1984; 10:388-98.
278. Cramer JA, Rosenheck R, Xu W et al. Quality of life in schizophrenia: a comparison of instruments. Department of Veterans Affairs Cooperative Study Group on Clozapine in Refractory Schizophrenia. Schizophr Bull 2000;26:659-66.
279. Auquier P, Simeoni MC, Sapin C et al. Development and validation of a patient-based health-related quality of life questionnaire in schizophrenia: the S-QoL. Schizophr Res 2003;63:137-49.
280. Wilkinson G, Hesdon B, Wild D et al. Self-report quality of life measure for people with schizophrenia: the SQLS. Br J Psychiatry 2000;177:42-6.
281. Martin CR, Allan R. Factor structure of the Schizophrenia Quality of Life Scale Revision 4 (SQLS-R4). Psychol Health Med 2007;12:126-34.
282. Zúñiga Le-Bert M, Wiessner M, Wehr S et al. Schizophrenia Quality of Life

Scale and Schizophrenia Quality of Life Scale Revision 4: a systematic review of measurement properties. Schizophr Bull 2025;51:997-1008.
283. Armoon B, Fleury MJ, Bayat AH et al. Quality of life and its correlated factors among patients with substance use disorders: a systematic review and meta-analysis. Arch Public Health 2022;80:179.
284. Dyer A, Böhnke JR, Curran D et al. A systematic review of quality of life and health-related quality of life as outcomes in substance and behavioural addictions. Drug Alcohol Rev 2023;42:1680-700.
285. Günther O, Roick C, Angermeyer MC et al. The EQ-5D in alcohol dependent patients: relationships among health-related quality of life, psychopathology and social functioning. Drug Alcohol Depend 2007;86:253-64.
286. Castelein S, Timmerman ME, van der Gaag M et al. Clinical, societal and personal recovery in schizophrenia spectrum disorders across time: states and annual transitions. Br J Psychiatry 2021;219:401-8.
287. Skar-Fröding R, Clausen H, Šaltytė Benth J et al. Associations between personal recovery and service user-rated versus clinician-rated clinical recovery, a cross-sectional study. BMC Psychiatry 2022;22:42.
288. Slade M, Amering M, Farkas M et al. Uses and abuses of recovery: implementing recovery-oriented practices in mental health systems. World Psychiatry 2014;13:12-20.
289. Moeller SB, Larsen PV, Austin S et al. Scalability, test-retest reliability and validity of the Brief INSPIRE-O measure of personal recovery in psychiatric services. Front Psychiatry 2024;15:1327020.
290. Williams J, Leamy M, Bird V et al. Development and evaluation of the INSPIRE measure of staff support for personal recovery. Soc Psychiatry Psychiatr Epidemiol 2015;50:777-86.
291. Neil ST, Kilbride M, Pitt L et al. The Questionnaire about the Process of Recovery (QPR): a measurement tool developed in collaboration with service users. Psychosis 2009;1:145-55.
292. Dehmahdi N, Law H, Pyle M et al. Estimating the minimum important difference for the Questionnaire about the Process of Recovery (QPR): an anchor-based approach. Psychosis 2021;13:3:220-30.
293. Law H, Neil ST, Dunn G et al. Psychometric properties of the Questionnaire about the Process of Recovery (QPR). Schizophr Res 2014;156:184-9.
294. Williams J, Leamy M, Pesola F et al. Psychometric evaluation of the Questionnaire about the Process of Recovery (QPR). Br J Psychiatry 2015;207:551-5.
295. Corrigan PW, Salzer M, Ralph RO et al. Examining the factor structure of the Recovery Assessment Scale. Schizophr Bull 2004;30:1035-41.
296. Hancock N, Scanlan JN, Honey A et al. Recovery Assessment Scale - Domains and Stages (RAS-DS): its feasibility and outcome measurement capacity. Aust N Z J Psychiatry 2015;49:624-33.
297. Fukui S, Salyers MP. Assessing the Recovery Assessment Scale across time. Psychiatr Serv 2021;72:716-9.
298. Hancock N, Scanlan JN, Kightley M et al. Recovery Assessment Scale - Domains and Stages: measurement capacity, relevance, acceptability and feasibility of use with young people. Early Interv Psychiatry 2020;14:179-87.
299. Salzer MS, Brusilovskiy E. Advancing recovery science: reliability and validity properties of the Recovery Assessment Scale. Psychiatr Serv 2014;65:442-53.
300. Scanlan JN, Hancock N, Honey A. The Recovery Assessment Scale - Domains and Stages (RAS-DS): sensitivity to change over time and convergent validity with level of unmet need. Psychiatry Res 2018;261:560-4.
301. Phelan M, Slade M, Thornicroft G et al. The Camberwell Assessment of Need: the validity and reliability of an instrument to assess the needs of people with severe mental illness. Br J Psychiatry 1995;167:589-95.
302. Slade M, Thornicroft G. Camberwell Assessment of Need, 2nd ed. Cambridge: Cambridge University Press, 2020.
303. Lloyd C, King R, Moore L. Subjective and objective indicators of recovery in severe mental illness: a cross-sectional study. Int J Soc Psychiatry 2010;56:220-9.
304. Hoe J, Orell M. Camberwell Assessment of Need for the Elderly (CANE), 2nd ed. Cambridge: Cambridge University Press, 2021.
305. Xenitidis K, Thornicroft G, Bouras N. Camberwell Assessment of Need for adults with Developmental and Intellectual Disabilities (CANDID), 2nd ed. Cambridge: Cambridge University Press, 2021.
306. Thomas S. Camberwell Assessment of Need Forensic Version (CANFOR), 2nd ed. Cambridge: Cambridge University Press, 2021.
307. Howard L, Slade M, O'Keane V et al. The Camberwell Assessment of Need for Pregnant Women and Mothers with Severe Mental Illness. London: Gaskell, 2008.
308. Semrau M, van Ommeren M, Blagescu M et al. The development and psychometric properties of the Humanitarian Emergency Settings Perceived Needs (HESPER) Scale. Am J Public Health 2012;102:e55-63.
309. Slade M, Leese M, Taylor R et al. The association between needs and quality of life in an epidemiologically representative sample of people with psychosis. Acta Psychiatr Scand 1999;100:149-57.
310. Junghan UM, Leese M, Priebe S et al. Staff and patient perspectives on unmet need and therapeutic alliance in community mental health services. Br J Psychiatry 2007;191:543-7.
311. Wiersma D, van Busschbach J. Are needs and satisfaction of care associated with quality of life? An epidemiological survey among the severely mentally ill in the Netherlands. Eur Arch Psychiatry Clin Neurosci 2001;251:239-46.
312. Ritsner MS, Ratner Y. Predicting predischarge anhedonia among inpatients with schizophrenia and schizoaffective disorders: a large-scale analysis. J Nerv Ment Dis 2019;207:12-21.
313. Kortrijk HE, Mulder NL, Kamperman AM et al. Employment rates in flexible assertive community treatment teams in The Netherlands: an observational study. Community Ment Health J 2019;55:350-9.
314. Ritsner MS, Farkash H, Rauchberger B et al. Assessment of health needs, satisfaction with care, and quality of life in compulsorily admitted patients with severe mental disorders. Psychiatry Res 2018;267:541-50.
315. Schmidt M, Ekstrand J, Tops AB. Self-reported needs for care, support and treatment of persons who frequently visit psychiatric emergency rooms in Sweden. Issues Ment Health Nurs 2018;39:738-45.
316. Guzman-Parra J, Moreno-Küstner B, Rivas F et al. Needs, perceived support, and hospital readmissions in patients with severe mental illness. Community Ment Health J 2018;54:189-96.
317. McCrone P, Leese M, Thornicroft G et al. Reliability of the Camberwell Assessment of Need -European Version. EPSILON Study 6. European Psychiatric Services: Inputs Linked to Outcome Domains and Needs. Br J Psychiatry 2000;177(Suppl. 39):s34-40.
318. Slade M, Leese M, Cahill S et al. Patient-rated mental health needs and quality of life improvement. Br J Psychiatry 2005;187:256-61.
319. Kidd SA, George L, O'Connell M et al. Recovery-oriented service provision and clinical outcomes in assertive community treatment. Psychiatr Rehabil J 2011;34:194-201.
320. Weber RL, Abraham KM. An assessment of the 10-item Mental Health Recovery Measure in a predominantly African American sample of adults with serious mental illness. Psychiatr Q 2022;93:849-60.
321. Keetharuth AD, Brazier J, Connell J et al. Recovering Quality of Life (ReQoL): a new generic self-reported outcome measure for use with people experiencing mental health difficulties. Br J Psychiatry 2018;212:42-9.
322. Monger B, Hardie SM, Ion R et al. The Individual Recovery Outcomes Counter: preliminary validation of a personal recovery measure. Psychiatrist 2013;37:221-7.
323. Galderisi S, Appelbaum PS, Gill N et al. Ethical challenges in contemporary psychiatry: an overview and an appraisal of possible strategies and research needs. World Psychiatry 2024;23:364-86.
324. Van Eck RM, van Velden J, Vellinga A et al. Personal recovery suits us all: a study in patients with non-affective psychosis, unaffected siblings and healthy controls. Schizophr Res 2023;255:24-32.
325. O'Keeffe D, Sheridan A, Kelly A et al. A qualitative study exploring personal recovery meaning and the potential influence of clinical recovery status on this meaning 20 years after a first-episode psychosis. Soc Psychiatry Psychiatr Epidemiol 2022;57:473-83.
326. Charles A, Korde P, Newby C et al. Proportionate translation of study materials and measures in a multinational global health trial: methodology development and implementation. BMJ Open 2022;12:e058083.

DOI:10.1002/wps.21338

The validity, reliability and clinical utility of the Alternative DSM-5 Model for Personality Disorders (AMPD) according to DSM-5 revision criteria

Carla Sharp[1], Lee Anna Clark[2], Kennedy M. Balzen[1], Tom Widiger[3], Stephanie Stepp[4], Mark Zimmerman[5], Robert F. Krueger[6]

[1]Department of Psychology, University of Houston, Houston, TX, USA; [2]Department of Psychology, University of Notre Dame, Notre Dame, IN, USA; [3]Department of Psychology, University of Kentucky, Lexington, KY, USA; [4]Department of Psychiatry, University of Pittsburgh, Pittsburgh, MA, USA; [5]Department of Psychiatry and Human Behavior, Brown University School of Medicine, South County Psychiatry, North Kingstown, RI, USA; [6]Department of Psychology, University of Minnesota, Minneapolis, MN, USA

A substantial body of empirical evidence has accumulated over the last 12 years since the publication of the Alternative Model for Personality Disorders (AMPD) in the DSM-5. As yet, this evidence has not been organized and reported using the criteria required by the American Psychiatric Association (APA) for proposals submitted to revise the DSM-5. These criteria are based on the Kendler-Kupfer update and expansion of the classic Robins-Guze criteria for establishing psychiatric diagnostic validity. We have been invited by the APA to undertake a review of the last decade of research on the AMPD and to propose a revised, simplified version of the model informed by this evidence. Here we present the findings of the review and our recommendations for the revision of the model. We begin with a brief reiteration of the background and rationale for the AMPD, followed by a description of the revision criteria required by the APA. We then summarize the evidence in support of the AMPD using the required framework. Our review indicates that AMPD-defined personality disorder (PD) shows similar patterns of associations as have been demonstrated for categorical PD diagnoses in terms of antecedent, concurrent and predictive validators. Head-to-head comparisons between AMPD-defined PD and categorical diagnoses suggest a more precise characterization of personality pathology by the AMPD. In addition, AMPD-defined PD appears to show higher reliability estimates than categorical PDs, and strong clinical utility, often outperforming categorical PD diagnoses. We conclude that the AMPD is ready for inclusion in the main section of the DSM. Recommendations are made for: a) further streamlining the AMPD in light of the last decade of accumulated evidence, and b) future research directions in areas where evidence is lacking or more limited.

Key words: Alternative Model for Personality Disorders (AMPD), DSM-5, categorical personality disorders, antecedent validators, concurrent validators, predictive validators, reliability, clinical utility

(World Psychiatry 2025;24:319–340)

In 2007, the DSM-5 Task Force convened the Personality and Personality Disorder Workgroup, charged with revising the DSM-IV chapter on personality disorders (PDs) and given a free hand to do so. The Workgroup members quickly came to the decision to develop a dimensional model for PD diagnosis, based on the growing consensus that the categorical diagnostic approach had stifled progress in the understanding and treatment of psychiatric disorders in general and PDs in particular[1-3].

However, despite the Workgroup's best efforts and the Task Force's support to develop a dimensional diagnostic system, the American Psychiatric Association (APA) Board of Trustees voted in December 2012 to reject the proposed model for inclusion in the DSM's main Section II. The dimensional model (named the Alternative DSM-5 Model for Personality Disorders, AMPD) was instead placed in a new Section III for "Emerging Measures and Models".

The reasons for this decision were complex, and have been described as "a story of shifting expectations, conflicting goals, and fractured alliances"[4]. However, the rejection appeared to be at least partly attributable to incompatibilities between the Workgroup's efforts and the criteria delineated by the APA's Scientific Review Committee (SRC). Indeed, although the development of the AMPD was grounded in considerable scientific evidence, accumulated over decades, it did not align readily with the type of evidence required by the SRC. According to this framework, which still governs the structure and organization of evidence for proposals submitted for revising the DSM-5, scientific evidence must be organized according to the well-established Robins and Guze criteria[5] for psychiatric nosology.

Since the publication of the AMPD in 2013, a substantial amount of research on the model has accumulated[6-19]. However, this research is yet to be organized according to the APA revision criteria to align with the SRC framework. Importantly, the DSM-IV's PD system has also not been evaluated in terms of these criteria. Against this background, in 2024, the authors of this paper, in addition to several other clinical scientists (see Acknowledgements) were invited by the APA Steering Committee to undertake a fresh review of the accumulated literature on the AMPD for evaluation of possible inclusion in Section II of the DSM-5, and to propose any necessary revisions of the model.

The goal of this paper is to present the findings of the review and to make recommendations for a revised, simplified version of the AMPD informed by the review. We begin with a brief outline of the background and rationale for the AMPD, followed by a description of the revision criteria employed by the SRC. We then summarize the evidence in support of the AMPD using the SRC framework.

We demonstrate that the AMPD is ready for inclusion in the main section of the DSM-5 according to the criteria used by the APA to adjudicate decisions over major changes in the diagnostic system. We conclude with a set of recommendations for: a) potential amendments to the current AMPD in light of the last decade of accumulated evidence, and b) future research directions in areas where we considered evidence to be lacking or of low quality.

BACKGROUND AND RATIONALE FOR THE AMPD

The DSM-5's AMPD represents a paradigm shift away from the traditional categorical model in favor of a dimensional approach in PD classification. Dimensional systems such as the AMPD were developed based on several important research findings[20-27]: a) significant heterogeneity exists within specific PDs, such that two individuals who meet criteria for a given PD may have very different clinical presentations; b) there are high levels of comorbidity (and/or overlap) among purportedly distinct PDs, such that individuals who meet criteria for one specific PD will likely meet criteria for two or more other PDs, calling into question the discrete nature of specific PDs; c) very few unique antecedents, correlates or consequences have been identified for any specific PD, while these are typically shared by other PDs and frequently co-occurring common mental disorders such as depression, anxiety and substance use disorders; d) low inter-rater reliability has consistently been demonstrated for the majority of categorically defined PDs (for instance, the median kappa for specific PD diagnoses has been shown to be 0.35, and the kappa between interview and questionnaire diagnoses is around 0.29)[28,29]; e) meta-analytic evidence of quantitative studies calls into question the structural integrity of the DSM-IV ten discrete PD syndromes[30], and emerging evidence indicates that PD manifestations may be better represented by a general factor of personality dysfunction that captures the shared variance of all PD manifestations plus trait dimensions that capture unique variance[31,32]; and f) there is no evidence supporting existing diagnostic thresholds for specific PDs – that is, diagnostic thresholds (e.g., five out of nine criteria for Borderline PD) do not actually demarcate the presence versus absence of disorder; rather, they are arbitrary thresholds along a continuum of prototypicality and level of impairment[20,22].

To address these limitations, the AMPD utilizes a more parsimonious conceptualization of personality pathology that accounts for heterogeneity within disorder and comorbidity to increase the validity, reliability and clinical utility of PD diagnosis. Accordingly, a single underlying severity continuum shared by all PDs defines core personality pathology. This underlying unidimensional severity criterion is called the Level of Personality Functioning (LPF; Criterion A of the AMPD), defined as impaired self and interpersonal functioning. LPF is rated on a 5-point scale from healthy/typical (=0) to severely impaired (=4), with a rating of 2 or more indicating personality dysfunction[33].

Apart from offering for the first time in the history of the DSM a psychiatric construct that is truly dimensional, ranging from typical to atypical, the LPF provides parsimony by eliminating the need for ten overlapping PDs which, as discussed, have been shown to have multiple problematic features. In so doing, the LPF also eliminates the need for Personality Disorder Not Otherwise Specified, which research has shown to be the most diagnosed PD, because real-life patients tend not to fit neatly into any of the ten categories[34], and clinicians have limited time for diagnosis.

After determination of the individual's LPF, the next step in the dimensional diagnostic process is evaluation of the pathological severity across five trait domains (Criterion B of the AMPD), to describe the ways in which the individual's self and interpersonal dysfunction are manifested. In the AMPD, these five trait domains include Negative Affectivity, Detachment, Antagonism, Disinhibition, and Psychoticism. A third diagnostic step in the AMPD allows the clinician to map the combination of Criterion A and B features onto traditional PD criteria. However, this step has been viewed as redundant, because the LPF and trait domains have been shown to provide full coverage of all PDs[13,35-39].

Overall, the combination of determining a patient's LPF in a first step and assignment of trait manifestation in a second step allows for the assessment of a patient's capacities for adaptive self and interpersonal functioning, and a distinction of how such functioning expresses itself in that particular patient. Two patients may, for instance, both be assigned an LPF score of 3, but one of them may show high levels of Impulsiveness and Disagreeableness (e.g., the patient frequently gets into fights), whereas another may show high levels of Detachment and low levels of Impulsiveness and Disagreeableness (e.g., he is a loner who avoids contact with others and leads a restricted life). In this way, the AMPD addresses heterogeneity within and comorbidity among different PD manifestations, thereby providing a more valid and clinically useful approach for characterizing personality pathology.

Indeed, since the publication of the AMPD in 2013, several reviews on the structural, concurrent and predictive validity, reliability and clinical utility of this model have been published[6-19]. However, the integration of the last 12 years of empirical evidence in support of the dimensional approach mainly followed conventions for evaluating construct validity[40,41], whereas APA revision criteria require a different set of standards, as described below.

THE ROBINS-GUZE / KENDLER-KUPFER CRITERIA FOR THE VALIDITY OF PSYCHIATRIC DISORDERS

By virtue of its grounding in medicine, the constructs of interest in psychiatry are necessarily *diagnostic* constructs, which immediately introduces a constraint in how they must be conceptualized in relation to the clinical utility requirement – that is, a psychiatric construct must be able to tell us whether a person suffers from a disease or not. In medicine, this is a prerequisite, because the diagnosis serves as the gateway to whether and which treatment might be indicated.

To assess whether a psychiatric construct can identify who suffers from disease, psychiatry has relied, for the last 50 years, on five principles proposed by Robins and Guze[5]. These principles reflect five types of research studies, each aimed at capturing a specific component of a disorder's nosology.

The first type are clinical description studies, aiming to demonstrate that the disorder has a particular and consistent pattern of symptoms and that these symptoms co-occur. The goal of these studies is to develop a coherent clinical picture of the disorder. Therefore, important non-psychopathological features that are common or prototypical of clinical presentations of the disorder must be identified in these studies, including for instance sociodemographic correlates such as age, sex, and age of onset. The

Table 1 American Psychiatric Association's criteria for proposed DSM-5 revisions

Validators

Antecedent validators	• Environmental risk factors • Prior psychiatric history • Familial aggregation and/or co-aggregation (i.e., family, twin or adoption studies) • Sociodemographic and cultural factors
Concurrent validators	• Cognitive, emotional, temperament and personality correlates (unrelated to the diagnostic criteria) • Biological markers, e.g., molecular genetics, neural substrates • Patterns of comorbidity • Degree or nature of the functional impairment
Predictive validators	• Diagnostic stability • Course of illness • Response to treatment

Reliability

- Inter-rater reliability
- Test-retest reliability
- Internal consistency

Clinical utility

The degree to which the proposed changes:

- do not alter caseness
- improve user acceptability
- improve clinicians' ability to apply the diagnostic criteria accurately and adherence to practice guidelines
- improve clinical outcomes
- improve the clinician's ability to select the best treatment or determine prognosis
- do not introduce unwanted negative consequences

second type are laboratory studies, that focus on identifying neurobiological and physiological substrates of the disorder, and are generally viewed as more empirically valid compared to clinical descriptive studies, given the external validity evidence that they presumably provide.

The third type are studies that delimitate the disorder from other related syndromes. These studies support the discriminant validity of the construct with the aim to establish its uniqueness relative to other psychiatric disorders with similar phenotypic presentations. The fourth type are follow-up studies that demonstrate a prototypical course and outcome of the symptoms. For example, demonstrating that individuals who were first identified with the disorder in baseline assessments present with the same disorder (as opposed to a different psychiatric disorder) in later assessments provides evidence for the original diagnostic criteria utilized at baseline. The fifth type are studies that aim to identify a familial and potentially genetic basis of the disorder. Demonstrating that the disorder displays significant heritability (through research designs that can separate genetic from environmental effects, such as twin studies) provides evidence for distinct psychopathological processes related to its phenomenology, and thus confirms the validity of the construct.

Fifty years later, the Robins and Guze criteria are still considered the gold standard approach for establishing the validity of psychiatric disorders. Indeed, these criteria informed what came to be known as the Kendler-Kupfer criteria, used by the then newly established SRC to evaluate the readiness of the AMPD for adoption into Section II. These criteria were outlined in a 2009 document entitled "Guidelines for making changes to DSM-V"[42], and still form the basis of the current APA guidelines for submitting proposals for making changes to the DSM-5 (see Table 1). In what follows, we provide an updated review of the literature evaluating the validity of the AMPD using these criteria.

The APA guidelines describe different types of proposals for changes to DSM-5 diagnostic criteria. Type 1 proposals involve changes to existing criteria to improve reliability and validity, and must demonstrate superiority of the proposed system over the existing one in head-to-head comparisons. Type 2 proposals involve addition of a new diagnostic category or specifier and require demonstration of validity, reliability and clinical utility in the absence of comparison with existing systems. The AMPD represents both Type 1 and Type 2 changes. Therefore, in our review, we provide evidence for the AMPD's validity, reliability and clinical utility according to the APA guidelines, as well as evidence in support of the AMPD's superiority over Section II's diagnostic categories in head-to-head comparisons.

VALIDITY

Table 2 summarizes Type 1 and Type 2 evidence in terms of antecedent, concurrent and predictive validators of AMPD, taking into account its strength.

Antecedent validators

Environmental antecedents

Strong evidence supports associations between childhood trauma and/or maltreatment and LPF[43-51]. Moderately strong evidence supports a link between poor parental bonding and/or closeness and LPF[52-54]. In addition, there is some evidence for an association of bullying victimization[55] and parental discord[56] with LPF.

This research is consistent with well-established developmental models for PD which indicate that the early caregiving and family environment is particularly important for the development of healthy personality functioning. Given the overlap between LPF and other measures of maladaptive self and interpersonal functioning, it is reasonable to assume that the mass of data available on environmental antecedents for the development of self-concept, self-esteem, self-appraisal, self-monitoring, self-directedness, moral decision-making, identity coherence, empathy, mentalizing, perspective-taking, and quality of relationships is also relevant[57-60].

Strong evidence across nine studies[54,61-68] supports a relation of a measure of AMPD Criterion B traits, the Personality Inventory for DSM-5 (PID-5)[69], with childhood trauma and maltreatment. Studies indicate that emotional (rather than physical or sexual) traumatic experiences are particularly associated with high levels of trait

Table 2 Summary of the evidence concerning the Alternative DSM-5 Model for Personality Disorders (AMPD) validity

	Basic evidence evaluation (Type 2)		Head-to-head superiority evaluation (Type 1)		
	LPF	**Traits**	**LPF**	**Traits**	**LPF + Traits**
Antecedent validators					
Environmental risk factors					
Poor maternal bonding or parental closeness	++	+	+		
Childhood trauma/maltreatment	+++	+++			
Bullying	+				
Parental discord	+				
Temperament	++	+++			
Psychiatric history	++	+			
Familial aggregation and/or co-aggregation		+++		++	
Concurrent validators					
Sociodemographic factors					
Invariance across genders	++	+++			
Mean differences according to sex assigned at birth	+++	++			
Mean differences for gender/sexual minorities		+			
Invariance across age groups	++	++			
Invariance across cultural groups	+	+++			
Cognitive correlates					
Social cognitive impairment	+++		+		
Cognitive distortions	++	++			
Executive functioning impairment	+	++			
Emotional correlates					
Emotion dysregulation	++	+++			
High levels of negative affect	++				
Alexithymia	+	+			
Emotional empathy	+	+			
Temperamental correlates	+	+++			
Convergent validity with similar constructs					
Self functioning	+++				
Interpersonal functioning	+++			+++	
Extreme traits		+++		+	++
Section II PD correlates	+++	+++	++	++	++
Patterns of comorbidity					
Anxiety disorders	++	+++			
Mood disorders	++	+++			
Post-traumatic stress disorder	++				
Conduct disorder	+	+++			
Substance use disorders	++	+++			
Psychosis		++			
Attention-deficit/hyperactivity disorder		+			
Psychosocial dysfunction	+++	+++	++	++	++
Pathophysiology, neurobiology		++			

Table 2 Summary of the evidence concerning the Alternative DSM-5 Model for Personality Disorders (AMPD) validity *(continued)*

	Basic evidence evaluation (Type 2)		Head-to-head superiority evaluation (Type 1)		
	LPF	Traits	LPF	Traits	LPF + Traits
Predictive validators					
Treatment response					
Treatment satisfaction and rapport	+				
Improved symptoms	+++	+++	++	+++	+++
Dropout	++	+	+		+
Diagnostic stability and course of illness	+++	+++			

$^{+++}$means that the finding was replicated in at least five studies; $^{++}$means that the finding was reported at least twice; $^{+}$means some evidence for the effect; Type 1 studies are those in which head-to-head comparisons are made between existing and new/proposed criteria sets; Type 2 studies do not require head-to-head comparisons, but simply require evidence that validity, reliability and clinical utility criteria are met. PD – personality disorder, LPF - Level of Personality Functioning.

scores, with Psychoticism and Detachment showing the strongest effect sizes. These findings fit with known research indicating that dissociation and detachment are prominent post-traumatic and trauma-coping mechanisms. In addition, one study shows a link between retrospective parental invalidation and PID-5 trait scores[54].

In terms of head-to-head comparisons, only one study reports data directly comparing antecedent environmental validators between Section II PDs and LPF[52]. In this study, maternal bonding during infancy predicted later LPF, but not Borderline PD symptom count.

Temperament as antecedent

Developmental models indicate that extreme levels of trait expression during childhood and adolescence confer vulnerability for emergence of PD. Two prospective follow-up studies demonstrate this to be true for LPF. In a combined sample of community and clinical youth, Vanwoerden et al[70] showed that high levels of temperamental trait expression in early adolescence prospectively predicted interpersonal problems in middle adolescence, culminating in maladaptive self functioning in early adulthood. Moreover, in a sample of 101 mother-child dyads recruited from obstetric units, Fleck et al[52] demonstrated that high levels of novelty seeking and low levels of harm avoidance at age 5 predicted Level of Personality Functioning Scale - Brief Form 2.0 (LPFS-BF 2.0)[71] scores at age 14.

We are not aware of studies that show a prospective association between early temperament and later AMPD-defined maladaptive trait scores *per se*. However, a robust literature exist documenting prospective links between early temperament and later similar personality traits[72].

Prior psychiatric history as antecedent

Moderate evidence supports prior psychiatric history as an antecedent validator for LPF[73,74]. Using both the Structured Clinical Interview for the DSM-5 Alternative Model for Personality Disorders (SCID-5-AMPD) Module I[75] and the Structured Interview of Personality Organization (STIPO)[76], Kampe et al[73] demonstrated that higher levels of personality dysfunction correlated with number of prior psychiatric hospitalizations, suicide attempts, and prior psychiatric diagnosis. In addition, a study using the Levels of Personality Functioning Scale (LPFS)[77] and the AMPD Clinician Rating Form (CRF)[77] found that self-reported past psychiatric hospitalization was associated with greater impairments across all four LPF domains as well as the trait domains of Detachment and Negative Affectivity[74].

Familial aggregation and/or co-aggregation

As yet, no twin or adoption studies have been conducted with measures of LPF. However, there are data using twin and adoption methodologies for evaluating the genetic basis of components of LPF, such as self-esteem, empathy, and interpersonal functioning [78-80].

Concerning AMPD traits, a study conducted in a population-based sample of Norwegian twin pairs found that the heritability estimates for Negative Affectivity, Detachment and Disinhibition ranged from 0.26 to 0.37[81].

Two studies provided a head-to-head comparison of Section II PDs vs. AMPD heritability estimates. In a Norwegian population-based sample of 1,408 adult twins, Reichborn-Kjennerud et al[82] estimated overlap in genetic and environmental risk factors for Section II PDs and maladaptive ends of trait domains assessed by the PID-5 Norwegian Brief Form 5 (PID-5-Norwegian-BF 5)[69] at two time points spanning a 10-year period. Results showed that, when measured concurrently, an average 81% of genetic variance was shared between the maladaptive-end trait domains and Paranoid, Schizotypal, Antisocial, Borderline and Avoidant PDs. For Obsessive-Compulsive PD, 43% of the genetic variance was shared. Genetic correlations between the individual maladaptive-end trait domains and PDs ranged from +0.21 (Detachment with Antisocial PD) to +0.91 (Negative Affectivity with Paranoid PD). When measured longitudinally, an average 54% of genetic vari-

ance was shared between the maladaptive-end trait domains and nine of the Section II PDs with, again, the exception being Obsessive-Compulsive PD, which shared only 28% of the genetic variance. The authors concluded that maladaptive trait domains tap, at an aggregate level, the same genetic risk factors as the DSM-5 Section II classification for most of the PDs.

A similar study measured Five-Factor Model personality traits, which reflect the healthy end of the PID-5 domain traits[83]. When assessed concurrently, the shared genetic variance of the traits to six of the Section II PDs was 58%. With a 6-to-12 year interval, the median shared variance was 36%.

Summary of antecedent validators

Moderate to strong evidence indicates that LPF and AMPD traits show similar antecedent validators as the PD categories in terms of environmental and temperamental antecedents and prior psychiatric history, as well as – in the case of traits – shared genetic variance. Given associated costs, more longitudinal and genetic studies will likely be pursued if and when the AMPD is fully implemented in Section II. We note that antecedent validators are mostly unstudied for current Section II PD categories, with the exceptions of Borderline and Antisocial PDs.

Concurrent validators

Sociodemographic correlates and cultural factors

The results of invariance testing studies for gender show no bias across sex assigned at birth, confirming that the examination of mean differences across sex is valid for both LPF[84-86] and maladaptive traits[87-90].

Studies on sex assigned at birth provide no evidence of mean differences between women and men in community or clinical samples for LPF (there were some small differences at the subcomponent level, where women were shown to have higher scores in self dysfunction compared to men[91,92]).

Sex differences have been observed for mean maladaptive trait scores, with the most consistent findings revealing higher levels of Negative Affectivity in adult females and higher scores on Antagonism in adult males[84]. Mean differences have been observed for sexual and gender minorities, who show higher mean scores across all maladaptive trait domains (with the exception of non-significant differences on Detachment in a clinical sample) than heterosexual individuals[93]. However, this should be interpreted with caution, given evidence of criterion bias that may make it easier for sexual and gender minority individuals to endorse criteria[94].

Extant research evaluating potential bias in test items for personality functioning[84,86,95] and maladaptive trait domains[96] shows no bias in terms of assessment across age groups when evaluating adolescents and adults. A recent study found differential item functioning of AMPD measures for older adults[97]. Specifically, 18 of 80 items on the Level of Personality Functioning Scale - Self-Report (LPFS-SR)[98] and all five maladaptive trait domains demonstrated large differential item functioning, with Psychoticism (100% of items) most affected. These findings need replication, and more work may be needed to understand the effects of old age on response patterns for AMPD measures.

The cross-cultural validity of the LPFS-BF-2.0[71] was evaluated in a large population-based study of adults in Canada, Spain and Belgium, demonstrating invariance among Dutch, English, French and Spanish versions of the measure[84]. Similar research has been conducted for maladaptive trait domains[18,84,99-102], suggesting the cross-cultural validity of measures of traits across different countries, cultures, languages and racial groups. Slight variations depending on the level of individualism vs. collectivism have been noted, in that some cultures encourage the development of distinct attitudes, self-definition, and striving to attain personal goals more than other cultures[102,103].

Although we are not aware of head-to-head comparisons for Section II- and III-related sociodemographic and cultural factors, studies of the ten categorical PDs have suggested bias in terms of sex, race, culture, and age in criteria[101,104-106].

Cognitive correlates

Strong evidence indicates that LPF is associated with social-cognitive impairment. Two studies have found that higher LPF scores associate with lower mentalizing capacity: one measured LPF by the Semi-Structured Interview for Personality Functioning DSM-5 (STiP-5.1[107])[108], and the other by self-report[109]. Two studies (one in an adult[109] and the other in an adolescent[86] sample) demonstrated that higher LPF is associated with lower self-reported reflective functioning. A study in adolescents found anomalies in social (but not monetary) reward processing[110].

Strong evidence supports a link between LPF and cognitive distortions, including early maladaptive schemas[111], deficits in cognitive components of empathy[112], daily-level distorted fortune telling (e.g., catastrophic predictions of the future), executive functioning and problem-solving difficulties, lower self-awareness[113,114], and lower levels of cognitive reappraisal[115].

Studies of the cognitive correlates of PID-5[69] maladaptive traits show moderate evidence for correlations with executive functioning, with one study suggesting that 73.3% of 30 PID-5 scales show correlations of 0.30 and higher for executive functioning tasks[116], and another demonstrating larger effect sizes for the domains of Negative Affectivity and Disinhibition[117].

Other cognitive factors that have been evaluated include cognitive distortions. Specifically, Detachment was shown to be associated with reduced truth bias in deception detection task, suggesting a protective role for Detachment in high-deception-frequency environments[118]. Negative Affectivity, Disinhibition and Psychoticism were significantly associated with difficulties in daily thinking (e.g., problem-solving[113]), and several maladaptive traits moderated the associations between daily-level cognitive distortions and LPF[114]. Finally, maladaptive trait domains have been linked with deficits in cognitive empathy[112].

One head-to-head comparison study of LPF versus Borderline PD demonstrated that both of them were associated with alterations in social (but not monetary) reward processing in adolescents, but that alterations in social reward were better predicted by LPF than by borderline traits[119].

Emotional correlates

At least six studies indicate strong evidence for emotion dysregulation problems associated with LPF, including distress intolerance[109,120] and maladaptive shame-coping[121]. Some evidence indicates links between LPF and higher levels of daily negative affect[120,122], alexithymia[86], and deficits in emotional empathy[112].

Studies of maladaptive traits have typically focused on emotion dysregulation as a correlate, and have primarily been conducted in non-clinical college student and online samples (though at least one study was conducted in a clinical sample[112]). In one study, emotion dysregulation showed a direct relation with the PID-5 total score, mediated by identity disturbance[123]. In another study, emotional dysregulation, as a transdiagnostic factor, was shown to mediate the relation between maladaptive traits and emotional disorders (anxiety, depression and stress)[124].

In a large ecological momentary assessment (EMA) study, Antagonism was shown to be associated with impulse-control difficulties and limited access to emotion regulation strategies, and Negative Affectivity, Detachment and Antagonism each moderated individuals' reactions to daily negative interpersonal events, such that the frequency of these events was higher in the presence of higher levels of maladaptive traits[125].

Another study of a mixed college and clinical sample (90% with a PD diagnosis) found significant associations between maladaptive trait domains and greater dysfunction in emotion dysregulation coping[126]. Finally, studies have also found significant links between maladaptive traits and both alexithymia and deficits in emotional empathy[112,127].

Temperament as correlate

Some evidence supports a negative link of LPF with effortful control and emotion regulation, and a positive link with impulsivity[128]. A vast literature provides strong evidence for a link between temperament and basic personality traits[129]. In turn, AMPD traits map onto basic personality traits[130-132], further supporting temperament as an important correlate of maladaptive trait scores.

Similar constructs as correlates

The results of 16 studies provide strong evidence that LPF is associated with other measures of maladaptive self functioning[85,95,98,133-142] and interpersonal functioning[86,133,134,139,140,142,143]. LPF has also been shown to correlate with measures of borderline personality organization[85,138,140,144]. These studies are important because they confirm the argument made earlier that other measures and constructs tapping maladaptive self and interpersonal functioning can be used as stand-ins for LPF, and provide strong evidence that LPF assesses the constructs that it was intended to capture.

Strong evidence supports the convergent and discriminant validity of the PID-5[69] with other measures of maladaptive traits[145-154]. These and other studies have been included in several meta-analyses confirming the high congruence between the PID-5 and measures of the Five-Factor Model of Personality[15,99,153].

At least three studies provide moderate support for head-to-head superiority of the AMPD compared to Section II PDs in associations with similar constructs. In a sample of 300 community-based adults, the PID-5[69] was found to be a better predictor of interview-assessed personality pathology than a Section II PD self-report measure (the Personality Diagnostic Questionnaire-4+, PDQ-4+)[155]. In a study of 200 male inmates, the AMPD outperformed Section II Antisocial PD in predicting scores on Hare's Psychopathy Checklist-Revised[156,157]. In a study of outpatients, the AMPD explained 46% of variance in psychopathy scores, whereas Section II Antisocial PD explained less than half of that variance (22%)[146].

Correlations with Section II PDs

Strong evidence from at least 27 studies supports the conclusion that LPF correlates with traditional Section II PDs, in adults and adolescents. These studies used a variety of LPF measures across clinical and community samples of adults, including the LPFS[98,157-161], the STiP 5.1[107,140], the SCID-5-AMPD[138], the LPFS-BF 2.0[141,162,163], the LPFS-SR[164,165], and the Five-Item Screening Scale for Personality Disorders (FISSPD)[95]. These findings were replicated in samples of adolescents using the STiP-5.1[142,166,167], the Levels of Personality Functioning Questionnaire 12-18 (LoPF-Q 12-18)[135,168-170], the LPFS-BF 2.0[85], and the FISSPD[137]. In addition, three studies used proxies for LPF and demonstrated similar correlations with Section II PDs[120,171,172].

Demonstrating correlations between LPF and all Section II PDs confirms that the LPF, as intended, is a measure of PD as defined by the traditional categorical system. However, it improves upon the latter by providing a more parsimonious assessment. Whereas evidence supports correlations with individual PDs, several studies also demonstrate associations with total number of PDs[162] or the severity of PD criterion count[95,98,163,164]. These studies confirm that the LPF, as intended, is a severity continuum indexing general personality dysfunction.

Importantly, studies also show that LPF increments both general psychopathology[168] and general disability[169] in predicting outcomes, indicating that it contains additional information regarding functioning that goes beyond general severity and disability – that is, maladaptive self and interpersonal functioning, the core features of personality dysfunction[173].

At least 28 studies provide strong evidence for the correlation between PID-5 maladaptive trait domains and Section II PDs[146,155,157,171,174-198]. Most of this research is summarized in six review papers[35-39,199], mirroring the finding by Morey et al[33] of mean dimen-

sional correlations between the two systems of 0.73, with a kappa of 0.54 for categorical diagnoses.

Several head-to-head comparison studies provide moderate evidence for comparability of the AMPD with Section II PDs. One study[200] examined whether the AMPD and Section II PD assessments identify the same patients in a sample of 305 psychiatric outpatients and 302 community residents who scored above threshold on a PD screen. Convergence across the two models was good to very good and demonstrated that the AMPD yields essentially the same overall prevalence of Section II PDs (~50% using each model, with a base rate difference of 5.3%) and largely identifies the same overall population (kappa = 0.74).

At least two studies examined how well interview-based ratings of Antisocial PD of both DSM-5 Section II and AMPD predicted either an interview-based measure of psychopathy in sample of prisoners[157] or self-reported measures of psychopathy in a sample of outpatients[201]. In both studies, AMPD traits and impairment predicted psychopathy more strongly than did Section II PD ratings.

Patterns of comorbidity with common mental disorders

Moderate evidence indicates correlations between self-report measures of LPF and measures of anxiety[146,202], depression[146,202,203], substance use[158,159,202], post-traumatic stress disorder[170,204], and conduct disorder (in adolescents)[170]. Strong evidence also exists for associations with general psychopathology[205]. Moderate evidence also exists for correlations between LPF and number of comorbid diagnoses[166,167], and some evidence indicates greater severity of anorexia nervosa associated with LPF in clinical samples of adolescents[206]. These patterns are consistent with what is known for traditional PDs[77].

A substantive research literature supports strong links between maladaptive trait scores and common mental disorders, which lays the foundation for a reformulation of psychopathology in terms of trait variability[21,207-210]. Consistent with this literature, evidence supports the association between PID-5 scores and mood disorders[211], psychosis[212-214], substance use disorder[211,215,216], and attention-deficit/hyperactivity disorder (ADHD)[217]. Some evidence indicates associations between maladaptive traits and proneness to migraine[218], as well as general psychiatric severity[89].

Nature and degree of psychosocial dysfunction

At least 28 studies provide strong evidence for the association of LPF with psychosocial dysfunction, including increases in general disability scores; and reductions in physical health[97,158,219], general adaptive functioning[33,140,162,169], quality of life[170,220], life satisfaction[221], well-being[222,223], and work and social adjustment[158,224,225]; as well as with pronounced loneliness[226], and discomfort and instability in relationships[227]. In a sample of community adolescents, LPF predicted social difficulties and lower well-being and life satisfaction one year later[228].

Several studies show that general psychiatric severity/impairment[73,146,165,168,217,225,229] and suicide severity and self-harm are correlated with LPF[166,230]. Notably, a study in clinical adolescents found that the score on the STiP-5.1 significantly predicted suicide attempt in the past year, and that especially the self dysfunction component of LPF explained additional variance in suicidal attempt over and above all psychiatric disorders[231]. LPF has also been shown to be associated with reduced capacity to meet developmental milestones in adolescents[232].

Some studies point to the self component of LPF as a better predictor of functioning[225] and suicide attempt[231] than the interpersonal component. With regard to degree, one study showed that, with each level of LPF, the risk for living alone, being single, being on a disability pension, and having symptom disorders increased, whereas months of working decreased[233].

Self-report measures of LPF also show correlations with general factors of personality pathology and total severity scores on personality measures[162,165], confirming that the LPF captures the general, shared features of personality dysfunction reflected in a unidimensional severity continuum.

Similar correlations have been demonstrated for maladaptive trait domain scores. For instance, one study showed that AMPD trait facets were strongly associated with the Global Assessment of Functioning (GAF) score (intraclass correlation coefficient, ICC between 0.85 and 0.92)[234], and another found that maladaptive trait domains prospectively predicted general psychological distress one year later[235]. In addition, a study of young people with first-onset psychosis demonstrated an association between GAF score and Detachment, suggesting that personality traits may be useful correlates of first-onset psychosis[236].

EMA studies have found that maladaptive traits are associated with greater daily aggression (particularly Negative Affect[237]), and greater daily levels of interpersonal tension and poorer social connectedness in adolescents[238]. A study also found associations between maladaptive trait domains and greater impairments in social and parental relationships, although associations between maladaptive traits and impairments in social relationships became non-significant after controlling for LPF (with a few exceptions)[227].

One study in a sample of 1,377 twins found that maladaptive trait domains were significantly associated with greater loneliness, with evidence of both genetic (r_g = 0.45-0.75) and unique environmental (r_e = 0.10-0.48) influences[239]. Maladaptive traits have also been linked to greater suicidal ideation (Negative Affect, Detachment) and behavior (Detachment, Disinhibition, Psychoticism)[240]. Finally, in a study of community adolescents, the overall score on the PID-5 predicted social difficulties one year later, and Psychoticism contributed to social rebuff whereas Detachment was associated with lower quality of life[228].

Head-to-head comparison studies provide moderate evidence for the superiority of the AMPD over DSM-5 Section II PDs in its association with psychosocial dysfunction.

In a study on 317 individuals, including a clinical sample of 282 patients of whom 192 were diagnosed with a PD[225], the SCID-5-AMPD Module I was a stronger predictor of scores on the Work and Social Adjustment Scale (WSAS)[241] and the GAF-Functioning (GAF-F)[242] than the sum of DSM-IV PD criteria, with the self com-

ponent providing the strongest predictive power.

A second study compared Section II PDs with the AMPD in explaining concurrent psychosocial functioning levels in 600 psychiatric outpatients and community residents screened as at risk for PD pathology[74]. The AMPD dimensions showed stronger associations with psychosocial difficulties and explained more of their variance compared with the Section II PDs.

In another study with this same sample[195], the two models were compared for their longitudinal predictive power of psychosocial functioning eight months later. Both models predicted functioning outcomes and each added significant predictive power, but the AMPD domains outpredicted the Section II PDs by 2.56%, and the AMPD facets outperformed the Section II PD criteria by 5.31%.

Pathophysiological / neurobiological correlates

A wealth of data is available on the neurobiological and pathophysiological correlates of components of LPF, including identity, self-esteem, self-appraisal, empathy, interpersonal functioning, social exclusion, rejection sensitivity, self-reflective capacity, and mentalizing ability[243]. These constructs are most often studied in the Research Domain Criteria (RDoC) systems for social processes – specifically Affiliation and Attachment, Social Communication, Perception and Understanding of Self, and Perception and Understanding of Others, each demonstrating differential convergence of associated brain areas in a meta-analysis[244].

Thus, while LPF proper has not been evaluated directly using neurobiological designs, research using these methods on its component parts suggests that, should such studies be conducted, they are likely to validate the LPF also from this perspective.

Some evidence suggests that PID-5 traits provide good coverage of biobehavioral externalizing liability[245]. A recent systematic review examining the neural correlates of maladaptive traits among individuals with Borderline and Antisocial PDs revealed that greater trait anger/hostility and aggression is associated with alterations in the interplay between subcortical (primarily the amygdala) and prefrontal regions[246]. Trait impulsivity was associated with alterations in serotonergic and endocannabinoid pathways and abnormalities in fronto-temporal-limbic regions; greater risk-taking was associated with weaker cortico-striatal connectivity.

Summary of concurrent validators

The LPF and maladaptive trait domain measures demonstrate moderate, and in some cases strong, associations with expected concurrent validators. Most studies have focused on validating LPF and maladaptive trait domains through evaluating their convergent validity with other measures of personality pathology, in addition to other measures that tap into maladaptive self and interpersonal functioning, and psychosocial dysfunction. An evidence base is developing for cognitive, emotional and pathophysiological correlates.

Since it is based in personality functioning and trait dimensions rather than categorical diagnoses, and its main components are transdiagnostic, the AMPD is much more compatible than Section II PDs with the RDoC approach to understanding psychopathology. At least six of the studies reported above have carried out head-to-head comparisons between AMPD and Section II PDs in terms of concurrent validators, all suggesting equivalence or superiority of the AMPD.

Predictive validators

Treatment response

Several prospective studies provide strong evidence for the predictive validity of LPF with regard to treatment response.

One study showed that self-reported LPF predicts treatment satisfaction and rapport with the provider in individuals on treatment for substance use[247]. In a sample of 191 patients with PD and 91 patients without PD, LPF predicted treatment dropout, with the risk being 2.3 times higher for patients with high LPF scores[160].

In a study of over 1,000 patients in outpatient settings in Norway, 57% had severe level LPF scores (LPFS-BF >18)[233]. From the start to the last phase of evidence-based treatment (mentalization based therapy, dialectical behavior therapy), 64% of the sample presented a score reduction in LPFS-BF. This rate of improvement is higher than that found for Borderline PD in a recent review of treatment response studies[248].

In a sample of patients receiving psychodynamic-based psychotherapy and matched controls, participants with higher LPF and maladaptive traits of Negative Affectivity and Psychoticism at baseline were more likely to drop out of therapy[249]. Additionally, patients' LPF scores declined significantly from baseline to follow-up (d=0.40), but were stable in the control group (d=0.10). Finally, in a naturalistic follow-up study of adolescent inpatients, a combination of LPF and trait self-report assessment predicted significant reduction in general psychiatric severity scores from admission to discharge[250].

Three studies have evaluated treatment response in relation to maladaptive traits. The first demonstrated mean differences in Negative Affectivity between admission and discharge[251]. In the second, Negative Affectivity and Detachment were related to higher admission severity in all four outcome domains (anxiety, depression, somatic symptoms, and psychosocial dysfunction)[252]. In a third study, participants with higher scores on Negative Affectivity and Psychoticism at baseline were significantly more likely to drop out of psychotherapy[249].

A meta-analysis of 207 studies investigated the extent to which personality traits changed as a result of intervention (primarily clinical interventions) and documented that "interventions were associated with marked changes in personality trait measures over an average time of 24 weeks"[253]. Negative Affectivity was the primary trait showing change, with an average effect size of 0.69, followed by Detachment, for which the average was 0.38. Changes replicated across experimental and non-experimental designs, for non-clinical interventions, and persisted in longitudinal follow-up. Type of therapy was not strongly associated with the amount of

change in personality traits.

Strong evidence from eight head-to-head comparison studies shows superiority of the AMPD compared to Section II PDs in predictive validity (e.g., for treatment dropout[160]). The self components of LPF are particularly predictive.

In two studies of adolescent inpatients, a combination of LPF and trait scores was a better predictor of treatment outcome (overall reduction in psychopathology) compared to a Borderline PD diagnosis[250,254]. In another study of a treatment-seeking sample of adult participants diagnosed with Section II Borderline PD, AMPD traits captured a more severe variant of the condition and incremented Section II diagnosis as a predictor of reduction in overall psychopathology by the end of treatment[255].

In a study of over 600 adults (50% patients) followed in a naturalistic study design, LPF in combination with AMPD Criterion B had greater power than Section II PDs in predicting twenty clinically relevant outcomes over 8 months[195]. In another study[256] of 311 patients (of whom 50% received past-year mental health treatment), Section II PDs were found to account for little variance in outcomes over and above the AMPD domains/facets, whereas the AMPD facets were generally more predictive of outcomes than the Section II PDs.

A study of 185 psychiatric outpatients found that the PID-5 trait domains predicted more variance and provided significant incremental prediction compared to Section II PDs for three of five areas of clinical dysfunction, and had better model fit for four of the five[257]. A study of 63 patients with PD found that the two systems yielded comparable one-year prediction of various clinical symptoms (e.g., depression, anxiety), but the PID-5 had stronger one-year predictive validity with respect to naturalistically observed EMA variables and informant reports of interpersonal functioning[258].

Diagnostic stability and course of illness

Strong evidence indicates that the PD features defined by LPF become recognizable during adolescence or early adult life, with several studies demonstrating the validity of LPF measures in adolescents, including the LPFS-BF 2.0[85,86], the LoPF-Q 12-18[135,170,259-261], and the STiP 5.1[142].

The natural course of LPF over adolescent development was documented in a study of 1,477 adolescents in the community, which showed that levels of LPF assessed at baseline were maintained through adolescence, and that rate of change in LPF was predicted by high levels of psychopathology at baseline[115].

Course of illness in treatment-seeking adults was evaluated in a Norwegian study, which showed that, although there is significant improvement in LPF symptoms over time, psychosocial dysfunction remains relatively high[233]. This finding mirrors the general consensus for the course of Borderline PD[262-264].

Two studies speak to the relative stability of LPF over medium-term intervals. In a study of 93 outpatients followed up approximately 1.5 years later, Wright et al[265] evaluated rank-order stability of proxy measures for LPF across eight indices of impairment, yielding coefficients of 0.17-0.65, with a mean rank-order stability of 0.37. This estimate was lower than for maladaptive trait domains, which showed a coefficient of 0.71, suggesting stability of PD traits, as conceptualized in most models of trait-based PD. In another study, Clark et al[195] re-tested individuals eight months after baseline, and demonstrated moderate stability for LPF domains (ranging from 0.43 to 0.53). The mean PID-5 rank order stability was 0.79.

Summary of predictive validators

Overall, the studies reported here indicate strong predictive validity of LPF and AMPD maladaptive traits for treatment response, with a superiority of the AMPD compared to Section II PDs (e.g., for treatment dropout[160]). Several studies show that LPF is a malleable treatment target. Extant research suggests that treatment response is similar for LPF-defined personality disorder as demonstrated for Section II PDs, particularly Borderline PD. We note that, although treatment response is relatively well studied for Borderline PD, it is much less so for other traditional PDs. For instance, the most recent meta-analytic review for Cluster C treatment response was conducted in 2009, listing 15 studies between 1982 and 2006[266].

Available research on course of LPF and AMPD maladaptive traits confirms that the onset of PD is in adolescence or young adulthood. Research also indicates that personality functioning is less stable than traits, and more susceptible to change.

RELIABILITY

Table 3 summarizes the Type 1 and 2 evidence in terms of reliability of the AMPD, taking into account its strength.

Inter-rater reliability

Strong evidence from ten studies[114,139,158,161,229,234,267-270] supports the inter-rater reliability (IRR) for interview-based LPFS[77], averaging an excellent coefficient of 0.87. Keeping in mind that

Table 3 Summary of the evidence concerning the Alternative DSM-5 Model for Personality Disorders (AMPD) reliability

	Basic evidence evaluation (Type 2)		Head-to-head superiority evaluation (Type 1)		
Type of reliability	**LPF**	**Traits**	**LPF**	**Traits**	**LPF + Traits**
Inter-rater reliability	+++	+++	++		
Test-retest reliability	+++	+++	+	+	+
Internal consistency	+++	+++	++		

$^{+++}$means that the finding was replicated in at least five studies; $^{++}$means that the finding was reported at least twice; $^{+}$means some evidence for the effect; Type 1 studies are those in which head-to-head comparisons are made between existing and new/proposed criteria sets; Type 2 studies do not require head-to-head comparisons, but simply require evidence that validity, reliability and clinical utility criteria are met. LPF - Level of Personality Functioning.

the LPFS is not a (semi-)structured interview, and that seven of the studies involved lay persons' ratings, this IRR value is notable, and significantly higher than typical IRRs for any categorical PDs.

Strong evidence is also available when using the STiP-5.1[107], with an average IRR of 0.89[107,140,142,144,271], and the SCID-5-AMPD Module I, with an average IRR of 0.87. At least one study has reported an excellent IRR estimate for the STIPO, at a value of 0.81[163].

These data are consistent with a recent meta-analysis[272] of seventeen IRR scores across fourteen studies using single-rater ICC or equivalent for total LPFS score. This resulted in a pooled ICC of 0.75, which is above the DSM-5 cutoff for acceptable IRR and indicates good reliability under ICC reporting guidelines[273,274].

Moderate evidence from four head-to-head comparisons with Section II PDs further establishes the superiority of the LPF. Morey[270] reported an ICC of 0.50 compared to values ranging from 0.11 to 0.49 for individual Section II PDs. Cruitt et al[158] found an ICC of 0.80 for LPF, outperforming the average ICC for the PD categories (0.67).

Strong evidence based on seven studies on maladaptive trait measures yielded mean/median ICCs of 0.73/0.75 for domains and 0.55/0.59 for facets[99,130,146,200,234,270,275]. Thus, although generally lower than for LPF, trait-domain level IRR was still good, with facets in the fair range.

Test-retest reliability

Strong evidence supports the test-retest reliability of LPF. In two studies in adults[164,221], self-report versions of the LPF (the LPFS-SR[98] and the Self and Interpersonal Functioning Scale, SIFS[221]) had an average r=0.90 over a 2-week interval. A study that evaluated the self-reported LPF in adolescents using the LoPF-Q 12-18[276] found a 2-week value of r=0.76[259]. suggesting that test-retest reliability may be lower in adolescents, although more studies are needed to assess this question fully.

This level of test-retest reliability compares favorably with values for Section II PDs. An early review of PD temporal stability reported short-term (~2 weeks) mean/median kappa values of 0.55/0.56 for any PD and 0.56/0.59 for specific PDs. Recent studies confirm these low-to-moderate test-retest reliability for Section II PDs[277,278].

In a study of 93 outpatients who had a PD diagnosis at Time 1, and completed a Time 2 assessment an average of 1.4 years later[32], AMPD domain mean/median trait test-retest reliability was 0.78/0.79, whereas that for facet traits was 0.80/0.81. In a 100-day daily-diary study of 101 individuals diagnosed with any PD[279], the mean/median test-retest reliability of five trait domain scores averaged over the 100 days was 0.86, whereas the mean/median day-to-day variability was 0.68/0.69. Thomadakis et al[89] evaluated test-retest reliability across 4 weeks for the PID-5 and reported an excellent range from r=0.82 to r=0.89 across the five domains. Fossati et al[280] found a 2-month test-retest reliability for PID-5-BF ranging from 0.78 (Negative Affectivity) to 0.97 (Detachment).

Longer-term stability – which, of course, confounds true change with measurement error – has not been widely assessed for AMPD Criterion A, but one longitudinal study[195] found moderately high stability for a fully dimensional assessment and moderate stability when this value was dichotomized. The AMPD traits were highly stable, but some stability was lost when these were configured as diagnoses based on the traits' continuous dimensions, and dichotomous diagnoses were even less stable[195]. Thus, the AMPD provides somewhat more stable values when dichotomized PDs are compared across models and considerably more stability when the models are compared as they are intended to be used – the AMPD dimensionally and the Section II PDs categorically.

Internal consistency

Strong evidence supports the internal consistency of LPF across various assessment tools. Seven studies[33,139,158,160,161,268,281] evaluated the internal consistency of the interview-based LPFS[77], mostly relying on clinician raters, with an average alpha of 0.80 across studies. Eight studies[91,92,107,108,140,271,282,283] evaluated internal consistency of the semi-structured interview-based STiP-5.1[107], averaging 0.95 across studies.

Eleven studies[71,84-86,128,134,162,222,284-286] across both patient and community samples evaluated the internal consistency of the 12-item self-report LPFS-BF 2.0, averaging 0.82 across studies. Two of these studies provided evidence in support of the internal consistency of the LPFS-BF 2.0 in adolescents, with one showing an omega (reliability coefficient) value of 0.93[85], and the other an omega of 0.83[86].

Moderate support for the superiority of LPF over Section II PDs is provided by three studies that carried out head-to-head comparisons. In the first[287], the alpha for the LPFS-SR[98] was 0.93 in an inpatient sample and 0.94 in a college sample, outperforming Section II PDs, which showed alpha values between 0.38 (Obsessive-Compulsive PD) and 0.80 (Antisocial PD) using the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II). In a second head-to-head comparison study, Weekers et al[283] showed that the STiP[107] outperformed the SCID-5-PD, with the former demonstrating an internal consistency alpha of 0.97 and the latter an average alpha of 0.76 across five discrete diagnostic categories. The third head-to-head comparison used Structured Interview for DSM Personality (SIDP) responses to score both the LPF and the Section II PDs. The LPF alpha was 0.85, whereas the mean/median alphas of the Section II PDs were 0.65/0.66 (range: 0.50-0.73), indicating considerable variability in internal consistency across the Section II PDs[195].

Strong evidence in support of the internal consistency of the PID-5 is provided by a meta-analytic review[153] including 10 studies for domains and 24 studies for facets, which reported a mean domain range of 0.88-0.95 and a mean facet range of 0.70-0.95, which exceeds the internal consistency of most categorical PD assessment tools. This conclusion mirrors the findings of a systematic review including 40 studies, also demonstrating strong internal consistency for the PID-5[288]. In a study conducted after the publication of those reviews, Clark et al[195] found a mean item-level alpha of 0.93 across trait domains for the PID-5, and a mean alpha of 0.72 for interview-based trait domain scores. For facets, the PID-5 mean

item-level alpha was 0.85.

More recent studies have generally found satisfactory internal consistency across self-reported maladaptive trait domains, with alphas generally above 0.70 for each of the domains, and average alphas across domains above 0.75[88,89,120,147,227,289].

Summary of reliability

IRR is excellent for LPF and good for trait-domain measures. IRR values are generally better for LPF than those previously reported for Section II PDs. Indeed, low IRR for categorical PDs was cited as one of the major motivations for the development of AMPD prior to 2013[290].

Short-term (~2 week) test-retest reliability for continuous measures of the LPF is high in adults, whereas in adolescents it is slightly lower, and daily variability (i.e., assessed in the context of EMA) is somewhat lower still.

The internal consistency of LPF dimensional measures is very strong, regardless of format (unstructured interview, semi-structured interview, or self-report) and sample. That for PID-5 traits is similarly strong. The internal consistency for traditional PDs ranges widely from poor to relatively strong, with significant heterogeneity within many of the categories.

We also note that, although factor analytic studies are not considered as part of the APA validator criteria set, they provide an important window into the construct validity of a diagnostic construct. In this regard, several systematic reviews point to factor-analytic evidence in support of the LPF construct as either unidimensional or a general factor in a bifactor structure[7,14,31]. This evidence stands in contrast to the failure of factor-analytic support for the ten discrete PDs[210,291-293].

CLINICAL UTILITY

Table 4 summarizes the evidence in support of the clinical utility of the AMPD, taking into account its strength.

Identification of patients with PD

Strong evidence from 13 studies demonstrates that LPF can distinguish patients with vs. without PD. This aspect of clinical utility has been documented for clinician-rated LPFS[163,193,294], LPFS rated by lay persons[281], LPFS-SR[295,296], LPFS-BF 2.0[71,226,297], STiP-5.1[91,107,226,282], and LoPF-Q 12-18 in adolescents[135].

Strong evidence from five studies in adults has shown the ability of the PID-5 to distinguish between individuals with vs. without PD[174,294,298-300].

In a head-to-head comparison study, Clark et al[195] compared the AMPD model with Section II in identifying individuals with PD, using a sample of 305 outpatients and 302 community adults screened for high risk of personality pathology[195]. There was a small (5.3%) difference in overall prevalence of PD between the two models, and they identified largely the same individuals.

Sensitivity and specificity

Strong evidence suggests that LPF measures have excellent sensitivity and specificity. Two studies[33,281] of the LPFS found an average sensitivity of 0.85 and specificity of 0.80. Three studies[135,259,276] of the self-report LoPF-Q 12-18[276] in adolescents showed an average sensitivity of 0.80 and specificity of 0.76. The SCID-5-AMPD Module I similarly showed an area under the curve (AUC) value of 0.84, suggesting a high degree of precision in detecting PD[225]. In a study of 772 inpatients, the PID-5 Borderline PD algorithm provided a good balance of sensitivity, specificity and odds ratio in identifying people with PD from those with bipolar disorder[300].

Acceptability among clinicians

Strong evidence reported in six studies has demonstrated acceptability among clinicians (across differing levels of experience) for use of the LPF on a wide range of indices of clinical utility, such as communicating with patients, formulating interventions, comprehensiveness, and global descriptive utility[234,281,301-304]. Similarly, strong evidence provided support for trait models in describing individuals' personality problems[305-313].

Strong evidence across six studies conducting head-to-head comparisons of Section II vs. AMPD acceptability among clinicians indicate superiority of the latter model. In a sample of 361 PD experts, Morey et al[301] demonstrated that they preferred a dimensional (73.4%) over a categorical (26.6%) approach. In another study of 337 clinicians, the AMPD was favored over the DSM-IV-TR PDs with respect to communicating with patients, formulating interventions, comprehensiveness, and global descriptive utility, but clinicians found the categorical system easier to use in professional communication[302]. In a different paper, the AMPD predicted clinicians' decisions better than the DSM-IV PDs in 10 of 11 clinical judgements[175].

Consistent with these findings, Maffly-Kipp and Morey[314] asked 136 mental health professionals to provide clinical judgments on a random subset of four (out of a possible 12) case vignettes. For each case, clinicians made a variety of diagnostic judgments corresponding to each model, as well as prognostic judgments. Results showed that the AMPD predictors consistently added unique variance beyond the Section II predictors, whereas the Section II predictors were rarely incremental over the AMPD. Further, the AMPD judgments predicted outcome judgments more consistently (98.3% of regressions) than the Section II predictors (70% of regressions), and the single Criterion A judgment (LPF) was the strongest overall predictor.

In another study among 20 clinicians[303], the SCID-5-AMPD interview was deemed more capable of describing patients' problems than the SCID-II, but required orientation to its less familiar theoret-

Table 4 Summary of the evidence concerning the Alternative DSM-5 Model for Personality Disorders (AMPD) clinical utility

	Basic evidence evaluation (Type 2)		Head-to-head superiority evaluation (Type 1)		
Type of clinical utility	**LPF**	**Traits**	**LPF**	**Traits**	**LPF + Traits**
Identification of patients with PD	+++	+++	+	+	+
Sensitivity and specificity	+++	+			
Acceptability among clinicians	+++	+++	+++	+++	+++
Acceptability among patients	+	+	+	+	+
Utility in identifying treatment targets and planning	++	++	++	++	++

$^{+++}$means that the finding was replicated in at least five studies; $^{++}$means that the finding was reported at least twice; $^{+}$means some evidence for the effect; Type 1 studies are those in which head-to-head comparisons are made between existing and new/proposed criteria sets; Type 2 studies do not require head-to-head comparisons, but simply require evidence that validity, reliability and clinical utility criteria are met. PD – personality disorder, LPF - Level of Personality Functioning.

ical basis. No concerns regarding SCID-5-AMPD complexity were noted. In addition, the SCID-5 AMPD outperformed the SCID-II in terms of ease of use with regard to clinical decision making.

These results were confirmed in a large study during the DSM-5 field trials, inclusive of 621 mental health professionals providing data for 1,269 patients, in which the AMPD received favorable clinical utility ratings, and was considered to be better than DSM-IV-TR[315]. Two studies[234,304] demonstrated that trainees favored the AMPD over Section II PDs on most clinical utility criteria.

Acceptability among patients

Moderate evidence indicates acceptability of the AMPD among patients. Cano and Sharp[316] compared the full AMPD model with Section II (Borderline PD specifically) among patients and their families (N=154). Participants rated mock diagnostic reports on six indices of clinical utility. The AMPD was favored over Section II on all six indices.

Utility in identifying treatment targets and planning

Six studies used clinical case study designs to evaluate the LPF in terms of identifying treatment targets and planning.

Bliton et al[285] showed that LPF was more useful in identifying core deficits of personality pathology for treatment planning than Section II diagnoses, whereas Pincus et al[317] found that the LPF provided an important severity dimension for treatment planning. In addition, the latter study showed that the AMPD could accommodate both narcissistic grandiosity and vulnerability, whereas the Section II narcissistic PD diagnosis could not.

Skodol et al[318] found that the LPF provides clinicians with a clear, consistent and coherent system for identifying personality pathology, quantifying its severity and characterizing clinical manifestations in terms of personality impairment. Waugh et al[11] showed that the full AMPD model facilitates case conceptualization, is easy to learn and use, and assists in patient feedback. Schmeck et al[319] documented the clinical utility of the AMPD for work with adolescents, showing that the LPF provides a useful index for conceptualizing adolescent personality pathology. A head-to-head comparison study[314] found that cross-validated LPF ratings outpredicted combined Section II PDs in determining clinical judgments about required level of care, risk and prognosis.

Summary of clinical utility

Many studies have converged on the conclusion that the LPF meets criteria for clinical utility. Of note, the brevity associated with assessment of only one severity continuum of maladaptive self and interpersonal functioning offers a clear advantage over the ten discrete PDs in the assessment of personality pathology. No information is lost using LPF, but much is gained by the more parsimonious evaluation of personality pathology that it offers.

The studies reviewed here are corroborated by a number of previous reviews and case-studies that have focused on the AMPD's utility for case conceptualization and treatment planning[12,251,285,318,320-324]; psychological assessment[11]; and use in forensic settings[325], as well as two meta-analytic reviews[9,326] concluding that the AMPD is generally perceived by clinicians as more useful than the current categorical approach.

SUMMARY AND LIMITATIONS OF THE CURRENT EVIDENCE BASE

We evaluated the empirical support for the AMPD organized according to the criteria outlined by the APA guidelines for proposing changes to the DSM-5, which are based on the Kendler-Kupfer update and expansion of the classic Robins-Guze criteria for establishing psychiatric diagnostic validity.

Overall, the research presented here suggests similar patterns of associations for AMPD-defined PD as demonstrated for categorical PD diagnoses in terms of antecedent, concurrent and predictive validators. Specifically, AMPD-defined PD shows associations with similar environmental risk factors, psychiatric history, and sociodemographic, temperament and personality correlates. AMPD-

defined PD also shows similar patterns of comorbidity and functional impairment as Section II PDs. A moderately stable course has been demonstrated, similar to that of studies of Section II PDs. In addition, AMPD-defined PD appears to be as responsive to treatment in naturalistic treatment studies.

Important to note is that, when we state similarities between AMPD-defined PD and Section II PDs for these validators, we refer mostly to the literature base on Borderline PD (and to a lesser extent Antisocial, Schizotypal, Narcissistic and Avoidant PD). This is because the literature base in terms of antecedent, concurrent and predictive validators for the other five PDs is so sparse that it precludes meta-analytic or even systematic review. This means that the current categorical system for the majority of PDs falls short of the evidence required to be considered valid based on Robins-Guze / Kendler-Kupfer criteria. In addition, only Borderline PD has a treatment literature strong enough to meet criteria upon which APA Treatment Guidelines could be developed[327].

The convergence between the AMPD and existing Section II PD categories was intentional; that is, the AMPD was not designed to have more predictive validity than traditional PDs[328]. Rather, it was designed to cover the same information (and identify the same patients) covered by the traditional categorical disorders, but reorganized into more conceptually coherent dimensions that demonstrate better structural validity. Despite this, some head-to-head comparisons did in fact demonstrate superiority for the AMPD over Section II diagnosis in terms of antecedent, concurrent and predictive validators, suggesting that a dimensional characterization of the personality pathology leads to larger effect sizes in correlates, resulting in incremental predictive validity of the AMPD over Section II (mostly Borderline PD) diagnosis.

The tendency of AMPD-defined PD to outperform Section II diagnoses was even more pronounced for reliability, with inter-rater reliability and internal consistency coefficients consistently higher than those previously reported for categorical diagnoses. This pattern of findings was also evident for clinical utility, with most head-to-head comparisons of Section II vs. AMPD demonstrating improved clinical utility for the latter.

This review, however, also identified areas where research is lacking for the AMPD, most notably that evaluating familial aggregation and/or co-aggregation and biological markers, as well as randomized controlled trials evaluating existing evidence-based treatment approaches (e.g., dialectical behavior therapy and mentalization-based therapy) for use in AMPD-defined PD. More large-scale epidemiological and long-term follow-up studies using AMPD-defined PD are also needed. These are studies that will not be conducted without large-scale funding, which, in turn, may be facilitated by the inclusion of the AMPD in DSM-5 Section II.

The incremental validity demonstrated in the studies reviewed here confirms that the AMPD provides more precise information about what it is to have a PD and how it manifests itself, with full acknowledgement of its heterogeneity and degree of severity. If we have better and more precise information about patients that better predict outcomes, we will also be able to better select which outcomes to pursue and when[328].

RECOMMENDATIONS FOR FURTHER IMPROVEMENT OF THE CURRENT VERSION OF THE AMPD

The research reviewed here also identified areas for improvement of the AMPD.

First, the three-step diagnostic process, especially the hybrid diagnostic aspect, adds unnecessary complexity and redundancy. As shown here, the combination of LPF and maladaptive trait expression more than adequately covers the traditional categories. Thus, a third step in which a hybrid diagnosis is determined is redundant and continues to reify categories whose validity is not documented.

The second major issue is consistency with the ICD-11[329]. This diagnostic system, which is now the official one for PD diagnosis in most of the world, mirrors the AMPD in defining a general level of personality functioning in a first step, followed by an option to describe individuals' unique trait manifestations. We recommend that the APA aligns the DSM-5 with the ICD-11 without losing any important aspects of the AMPD that have been empirically validated over the last decade or of the current categorical model.

Against this background, we propose a simplified version of the AMPD that retains its essential features but eliminates some details to reduce complexity. Specifically, we propose:

- Removal of step 3 in the current AMPD diagnostic process, the hybrid diagnosis, based on research demonstrating the coverage of PD constructs by the AMPD reviewed herein. Instead, we propose offering optional specifiers for trait combinations that cover traditional categorical PDs. More precisely, we propose adding optional specification on how maladaptive traits combine to create trait patterns that resemble traditional categorical diagnoses, based on research showing that certain trait and facet combinations provide coverage of the traditional categorically defined PDs in conjunction with the level of personality functioning. A similar step was made in the ICD-11 for Borderline PD. Although such pattern specifiers are ultimately redundant, they may provide important cross-walk information while health systems continue to transition to a dimensional system, with the ultimate goal of eventual removal of these remnants of the traditional PDs over time.
- Removal of the requirement that at least two of the four elements of LPF need to be present, based on the fact that an overall level of 2 on the 5-point scale from healthy/typical to severely impaired has strong sensitivity (0.85), specificity (0.73) and AUC (0.83) for predicting PD[33]. In addition, impairment in two of the four components was the empirically derived algorithm for the hybrid PD types, which, as we are proposing, should be removed.
- Providing guidance to users for evaluating LPF severity by placing the levels on a linear severity scale, so that severity increases as self and interpersonal functioning impairment increases in intensity, chronicity, pervasiveness, and impact on psychosocial functioning. This proposed change is motivated by feedback from clinicians using the LPFS asking how features in per-

sonality dysfunction change as a function of severity. By providing this guidance, the levels are not fundamentally changed, but simply operationalized in a more accessible form, with familiar and well-used benchmarks of severity for most psychiatric disorders.

- Changing LPF-level labels to remove the "extreme" label and incorporate a "mild" label. This change is motivated by: a) harmonization with the ICD-11; b) concern in clinical settings that there is no AMPD option for diagnosing "mild" PD in the current LPF; and c) feedback from clinicians that the distinction between "severe" and "extreme" PD does not match the level descriptions in the LPF well. The proposed levels are: little or no impairment (=0), subthreshold impairment (=1), mild impairment (=2), moderate impairment (=3), and severe impairment (=4). This terminology is also consistent with that widely used in the medical field.
- Providing an optional rating scale for Criterion B. In keeping with a dimensional approach, a 4-point rating scale is proposed for specification of trait manifestation: 0= "very little or not at all descriptive", 1= "mildly descriptive", 2= moderately descriptive", 3= "very descriptive". Justification for this 4-point scale is provided by ample data on the PID-5, which uses this same 4-point scale when evaluating trait manifestation in individuals.
- Adding Compulsivity as a sixth trait domain. This proposed change is based on the fact that this domain was included in a pre-final DSM-5 proposal and acknowledged as late as 2011[27]. In addition, adding Compulsivity facilitates alignment with the ICD-11, which includes the trait domain "Anankastia". Adding Compulsivity will also increase trait coverage of DSM-5 Section II Obsessive-Compulsive PD[39,330]. Indeed, there is growing evidence that Compulsivity is a trait domain not completely captured by the current AMPD five domain approach[39,331-333].

CONCLUSIONS

This review was undertaken at the invitation of the APA Steering Committee to assess the readiness of the AMPD to be included in Section II of the DSM-5. To this end, we used the criteria required by the APA for proposals submitted for further revising the DSM-5. Moreover, we aimed to propose a revised, simplified version of the AMPD informed by the last decade of research.

We conclude that the accumulated evidence supports the inclusion of the AMPD in Section II, with recommended amendments aimed at further simplifying the model and bringing it in closer alignment with the ICD-11. With these proposed changes, we address the call for a simplified version of the AMPD that is nonetheless able to leverage the empirical support to its current version, whose essential features remain unchanged. Thus, the proposed simplified version does not necessitate any further development of measures or assessment tools.

Further, we contend that the proposed changes do not introduce any unwanted negative consequences for providers or patients. The proposed changes do not imply removal of LPF or trait domain features, except for the removal of the hybrid diagnosis, which has been proven to be redundant. Concerns over the removal of the hybrid diagnosis are mitigated by the pattern specifiers in the proposed simplified version. Moreover, the data we presented on head-to-head comparisons with Section II PDs indicate that users of the simplified AMPD will be able to describe the behavioral patterns, as well as the phenomenology, associated features, correlates and outcomes of traditional PDs more precisely than the categorical diagnosis itself through the inclusion of LPF and trait domains.

Another concern may relate to the argument that, if the simplified AMPD is to be consistent with the ICD-11 diagnostic system for PD, an option may be to just use the ICD-11. We decided against this option given the wealth of empirical research that exist for the AMPD – a much larger literature base than that for the ICD-11 system.

Much research has accumulated since the publication of the AMPD in the DSM-5, and our review offers clear evidence that the criteria for its inclusion as the main DSM PD model have been met. The ICD system has already shifted to align with the burgeoning empirical literature. Our recommendation is that the DSM follow suit and be brought into alignment with the scientific literature reviewed here, as well as with the general international consensus supporting an empirically based nosology for PDs.

ACKNOWLEDGEMENTS

The authors wish to thank the clinical scientists who served on the initial committee invited to undertake a review of the AMPD literature, including L. Choi-Kain, R. Dudley, L. Morey, J. Oldham, C. Seijas-Rodriguez and M. Zanarini. They also thank D. Clarke and L. Yousif from the American Psychiatric Association.

REFERENCES

1. Kupfer DJ, First MB, Regier DA. A research agenda for DSM-V. Washington: American Psychiatric Association, 2002.
2. Regier DA. Dimensional approaches to psychiatric classification: refining the research agenda for DSM-V: an introduction. Int J Methods Psychiatr Res 2007;16(Suppl. 1):S1-5.
3. Regier DA, Narrow WE, Kuhl EA et al. The conceptual development of DSM-V. Am J Psychiatry 2010;166:645-55.
4. Zachar P, Krueger RF, Kendler KS. Personality disorder in DSM-5: an oral history. Psychol Med 2016;46:1-10.
5. Robins E, Guze SB. Establishment of diagnostic validity in psychiatric illness – Its application to schizophrenia. Am J Psychiatry 1970;126:983-7.
6. Zimmermann J, Kerber A, Rek K et al. A brief but comprehensive review of research on the Alternative DSM-5 Model for Personality Disorders. Curr Psychiatry Rep 2019;21:92.
7. Sharp C, Wall K. DSM-5 Level of Personality Functioning: refocusing personality disorder on what it means to be human. Annu Rev Clin Psychol 2021; 17:313-37.
8. Krueger RF, Hobbs KA. An overview of the DSM-5 Alternative Model of Personality Disorders. Psychopathology 2020;53:126-32.
9. Milinkovic MS, Tiliopoulos N. A systematic review of the clinical utility of the DSM-5 section III alternative model of personality disorder. Personal Disord 2020;11:377-97.
10. Morey LC, Benson KT, Busch AJ et al. Personality disorders in DSM-5: emerging research on the alternative model. Curr Psychiatry Rep 2015;17:558.
11. Waugh MH, Hopwood CJ, Krueger RF et al. Psychological assessment with the DSM-5 Alternative Model for Personality Disorders: tradition and innovation. Prof Psychol Res Pr 2017;48:79-89.
12. Rodriguez-Seijas C, Ruggero C, Eaton NR et al. The DSM-5 Alternative Model for Personality Disorders and clinical treatment: a review. Curr Treat Options Psychiatry 2019;6:284-98.

13. Sharp C, Miller JD. Ten-year retrospective on the DSM-5 Alternative Model of Personality Disorder: seeing the forest for the trees. Personal Disord 2022; 13:301-4.
14. Morey LC, McCredie MN, Bender DS et al. Criterion A: Level of Personality Functioning in the Alternative DSM-5 Model for Personality Disorders. Personal Disord 2022;13:305-15.
15. Clark LA, Watson D. The trait model of the DSM-5 Alternative Model of Personality Disorder (AMPD): a structural review. Personal Disord 2022;13:328-36.
16. Widiger TA, Hines A. The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition alternative model of personality disorder. Personal Disord 2022;13:347-55.
17. Bach B, Tracy M. Clinical utility of the alternative model of personality disorders: a 10th year anniversary review. Personal Disord 2022;13:369-79.
18. Watters CA, Bagby RM. A meta-analysis of the five-factor internal structure of the Personality Inventory for DSM-5. Psychol Assess 2018;30:1255-60.
19. Zachar P, First MB. Transitioning to a dimensional model of personality disorder in DSM 5.1 and beyond. Curr Opin Psychiatry 2015;28:66-72.
20. Clark LA. Assessment and diagnosis of personality disorder: perennial issues and an emerging reconceptualization. Annu Rev Psychol 2007;58:227-57.
21. Krueger RF, Kotov R, Watson D et al. Progress in achieving quantitative classification of psychopathology. World Psychiatry 2018;17:282-93.
22. Widiger TA, Trull TJ. Plate tectonics in the classification of personality disorder: shifting to a dimensional model. Am Psychol 2007;62:71-83.
23. Livesley WJ, West M, Tanney A. Historical comment on DSM-III schizoid and avoidant personality disorders. Am J Psychiatry 1985;142:1344-7.
24. Tyrer P, Alexander J. Classification of personality disorder. Br J Psychiatry 1979;135:163-7.
25. Widiger TA, Frances A. The DSM-III personality disorders. Perspectives from psychology. Arch Gen Psychiatry 1985;42:615-23.
26. Morey LC, Waugh MH, Blashfield RK. MMPI scales for DSM-III personality disorders: their derivation and correlates. J Pers Assess 1985;49:245-51.
27. Skodol AE, Clark LA, Bender DS et al. Proposed changes in personality and personality disorder assessment and diagnosis for DSM-5 Part I: description and rationale. Personal Disord 2011;2:4-22.
28. Clark LA. A critique of Gunderson's views of DSM-5: commentary on "Seeking clarity for future revisions of the personality disorders in DSM-5". Personal Disord 2013;4:379-80.
29. Clark LA, Livesley WJ, Morey L. Personality disorder assessment: the challenge of construct validity. J Pers Disord 1997;11:205-31.
30. Sheets E, Craighead WE. Toward an empirically based classification of personality pathology. Clin Psychol Sci Pract 2007;14:77-93.
31. Sharp C, Wright AGC, Fowler JC et al. The structure of personality pathology: both general ('g') and specific ('s') factors? J Abnorm Psychol 2015;124:387-98.
32. Wright AGC, Hopwood CJ, Skodol AE et al. Longitudinal validation of general and specific structural features of personality pathology. J Abnorm Psychol 2016;125:1120-34.
33. Morey LC, Bender DS, Skodol AE. Validating the proposed diagnostic and statistical manual of mental disorders, 5th edition, severity indicator for personality disorder. J Nerv Ment Dis 2013;201:729-35.
34. Gunderson J. Understanding the speed of change: rejoinder for "Seeking clarity for future revisions of the personality disorders in DSM-5". Personal Disord 2013;4:381.
35. Kwapil T, Barrantes-Vidal N. Schizotypal personality disorder in the alternative model for personality disorders. Personal Disord 2022;13:392-6.
36. Vanwoerden S, Stepp SD. The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, alternative model conceptualization of borderline personality disorder: a review of the evidence. Personal Disord 2022;13:402-6.
37. Miller JD, Crowe ML, Sharpe BM. Narcissism and the DSM-5 alternative model of personality disorder. Personal Disord 2022;13:407-11.
38. Hummelen B, Ulltveit-Moe Eikenaes I, Wilberg T. Is the alternative model for personality disorders able to capture avoidant personality disorder according to Section II of the DSM-5? A systematic review. Personal Disord 2022;13:412-7.
39. Samuel DB, Balling CE, Bucher MA. The alternative model of personality disorder is inadequate for capturing obsessive-compulsive personality disorder. Personal Disord 2022;13:418-21.
40. Cronbach LJ, Meehl PE. Construct validity in psychological tests. Psychol Bull 1955;52:281-302.
41. Loevinger J. Objective tests as instruments of psychological theory. Psychol Rep 1957;3:635-94.
42. Kendler KS, Kupfer DJ, Narrow W et al. Guidelines for making changes to DSM-5. Washington: American Psychiatric Association, 2009.
43. Gaudiešiūtė E, Skabeikytė-Norkienė G, Barkauskienė R. Personality functioning in adolescents: exploring the links with childhood maltreatment types and current psychological difficulties. Child Adolesc Ment Health 2025; doi: 10.1111/camh.12781.
44. d'Huart D, Hutsebaut J, Seker S et al. Personality functioning and the pathogenic effect of childhood maltreatment in a high-risk sample. Child Adolesc Psychiatry Ment Health 2022;16:95.
45. Krakau L, Tibubos AN, Beutel ME et al. Personality functioning as a mediator of adult mental health following child maltreatment. J Affect Disord 2021;291:126-34.
46. Ernst M, Brahler E, Kampling H et al. Is the end in the beginning? Child maltreatment increases the risk of non-suicidal self-injury and suicide attempts through impaired personality functioning. Child Abuse Negl 2022;133:105870.
47. Gander M, Buchheim A, Bock A et al. Unresolved attachment mediates the relationship between childhood trauma and impaired personality functioning in adolescence. J Pers Disord 2020;34(Suppl. B):84-103.
48. Back SN, Flechsenhar A, Bertsch K et al. Childhood traumatic experiences and dimensional models of personality disorder in DSM-5 and ICD-11: opportunities and challenges. Curr Psychiatry Rep 2021;23:60.
49. Bertsch K, Gohre I, Cottin M et al. Traumatic childhood experiences and personality functioning: effect of body connection in a cross-sectional German and Chilean sample. Borderline Personal Disord Emot Dysregul 2024;11:20.
50. Penner F, Gambin M, Sharp C. Childhood maltreatment and identity diffusion among inpatient adolescents: the role of reflective function. J Adolesc 2019;76:65-74.
51. Kerber A, Gewehr E, Zimmermann J et al. Adverse childhood experiences and personality functioning interact substantially in predicting depression, anxiety, and somatization. Personal Ment Health 2023;17:246-58.
52. Fleck L, Fuchs A, Moehler E et al. Maternal bonding impairment predicts personality disorder features in adolescence: the moderating role of child temperament and sex. Personal Disord 2021;12:475-83.
53. Cervantes BR, Kerr S, Vanwoerden S et al. Operationalizing intimacy and identity aspects of personality functioning in relation to personality disorder in adolescents. Front Psychiatry 2023;14:1153274.
54. Li Z, Duan Y, Liu Y et al. Retrospective reports of perceived parental invalidation and borderline personality traits: the indirect effect of personality functioning. Personal Disord 2023;14:584-9.
55. Seiffert N, Cavelti M, Schmidt SJ et al. The relationship between bullying victimization and impairment in personality functioning in a clinical adolescent sample. Personal Disord 2024;15:173-80.
56. Skabeikyte-Norkiene G, Sharp C, Kulesz PA et al. Personality pathology in adolescence: relationship quality with parents and peers as predictors of the level of personality functioning. Borderline Personal Disord Emot Dysregul 2022;9:31.
57. Aldrich NJ, Chen J, Alfieri L. Evaluating associations between parental mind-mindedness and children's developmental capacities through meta-analysis. Dev Rev 2021;60:100946.
58. Deneault AA, Hammond SI, Madigan S. A meta-analysis of child-parent attachment in early childhood and prosociality. Dev Psychol 2023;59:236-55.
59. Zeegers MAJ, Meins E, Stams G-JJM et al. Does attachment security predict children's thinking-about-thinking and thinking-about-feeling? A meta-analytic review. Dev Rev 2019;54:100885.
60. Zhang H, Wang W, Liu S et al. A meta-analytic review of the impact of child maltreatment on self-esteem: 1981 to 2021. Trauma Violence Abuse 2023;24:3398-411.
61. Veith AC, Russell TD, King AR. PID-5 trait mediation of childhood maltreatment effects. Pers Indiv Differ 2017;104:58-63.
62. King AR. PID-5 trait indicators of emotional instability and childhood adversity antecedents. Psychol Rep 2022;125:310-27.
63. Granieri A, Guglielmucci F, Costanzo A et al. Trauma-related dissociation is linked with maladaptive personality functioning. Front Psychiatry 2018;9:206.
64. Borroni S, Somma A, Krueger RF et al. Assessing the relationships between self-reports of childhood adverse experiences and DSM-5 alternative model of personality disorder traits and domains: a study on Italian community-dwelling adults. Personal Ment Health 2019;13:180-9.
65. Bach B, Fjeldsted R. The role of DSM-5 borderline personality symptomatology and traits in the link between childhood trauma and suicidal risk in psychiatric patients. Borderline Personal Disord Emot Dysregul 2017;4:12.
66. Boland J, Rock R, Johnson AK et al. Pathways to incarceration: an examination of childhood maltreatment and personality psychopathology in incarcerated adults. Psychol Crime Law 2021;27:253-64.
67. Hemmati A, Newton-Howes G, Falahi S et al. Personality pathology among adults with history of childhood sexual abuse: study of the relevance of DSM-5

proposed traits and psychobiological features of temperament and character. Indian J Psychol Med 2021;43:135-43.
68. Back SN, Zettl M, Bertsch K et al. Personality functioning, maladaptive traits, and childhood trauma. Psychotherapeut 2020;65:374-82.
69. Krueger RF, Derringer J, Markon KE et al. Initial construction of a maladaptive personality trait model and inventory for DSM-5. Psychol Med 2012;42:1879-90.
70. Vanwoerden S, Franssens R, Sharp C et al. The development of Criterion A personality pathology: the relevance of childhood social functioning for young adult daily self-functioning. Child Psychiatry Hum Dev 2022;53:1148-60.
71. Weekers LC, Hutsebaut J, Kamphuis JH. The Level of Personality Functioning Scale - Brief Form 2.0: update of a brief instrument for assessing level of personality functioning. Personal Ment Health 2019;13:3-14.
72. Rothbart MK. Temperament, development, and personality. Curr Dir Psychol Sci 2007;16:207-12.
73. Kampe L, Zimmermann J, Bender D et al. Comparison of the Structured DSM-5 Clinical Interview for the Level of Personality Functioning Scale with the Structured Interview of Personality Organization. J Pers Assess 2018;100:642-9.
74. Ro E, Nuzum H, Clark LA. Competing models of personality disorder: relations with psychosocial functioning. Assessment 2025;32:321-34.
75. Bender DS, Skodol AE, First MB. Structured Clinical Interview for the DSM-5 Alternative Model for Personality Disorders (SCID-5-AMPD) Module I: Level of Personality Functioning Scale. Washington: American Psychiatric Association, 2018.
76. Clarkin JF, Caligor E, Stern BL et al. Structured Interview of Personality Organization (STIPO). New York: Personality Disorders Institute, Weill Medical College of Cornell University, 2004.
77. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Washington: American Psychiatric Association, 2013.
78. Abramson L, Uzefovsky F, Toccaceli V et al. The genetic and environmental origins of emotional and cognitive empathy: review and meta-analyses of twin studies. Neurosci Biobehav Rev 2020;114:113-33.
79. Luo YLL, Cai H. Self in the mirror of behavioural genetics: reflections from twin studies on self-esteem and self-concept. In: Kovas Y, Malykh S, Gaysina D (eds). Behavioural genetics for education. London: Palgrave Macmillan, 2016:96-112.
80. Kendler KS, Gardner CO, Prescott CA. A population-based twin study of self-esteem and gender. Psychol Med 1998;28:1403-9.
81. Wright ZE, Pahlen S, Krueger RF. Genetic and environmental influences on Diagnostic and Statistical Manual of Mental Disorders - Fifth Edition (DSM-5) maladaptive personality traits and their connections with normative personality traits. J Abnorm Psychol 2017;126:416-28.
82. Reichborn-Kjennerud T, Krueger RF, Ystrom E et al. Do DSM-5 Section II personality disorders and Section III personality trait domains reflect the same genetic and environmental risk factors? Psychol Med 2017;47:2205-15.
83. Czajkowski N, Aggen SH, Krueger RF et al. A twin study of normative personality and DSM-IV personality disorder criterion counts: evidence for separate genetic influences. Am J Psychiatry 2018;175:649-56.
84. Le Corff Y, Aluja A, Rossi G et al. Construct validity of the Dutch, English, French, and Spanish LPFS-BF 2.0: measurement invariance across language and gender and criterion validity. J Pers Disord 2022;36:662-79.
85. Wu J, Allman M, Balzen K et al. First psychometric evaluation of the Level of Personality Functioning Scale - Brief Form 2.0 in adolescents. Personal Disord 2024;15:492-503.
86. Balzen KM, Sharp C, Unzurruzaga E et al. The Spanish version of the Level of Personality Functioning Scale - Brief Form 2.0 (LPFS-BF 2.0): psychometric evaluation in adolescents who have suffered from parental abuse and neglect and a community sample. Scand J Child Adolesc Psychiatr Psychol 2024;12:20240012.
87. Suzuki T, Samuel DB, Pahlen S et al. DSM-5 alternative personality disorder model traits as maladaptive extreme variants of the five-factor model: an item-response theory analysis. J Abnorm Psychol 2015;124:343-54.
88. Rivard G, Le Corff Y, Lapalme M et al. Measurement invariance of the Personality Inventory for DSM-5 across sex. Front Psychiatry 2024;15:1328937.
89. Thomadakis C, Tsaousis I, Douzenis A et al. Development and psychometric evaluation of the Greek Personality Psychopathology Inventory (GPPI-5): an ESEM application approach following the DSM-5 AMPD. Psychological Test Adaptation and Development 2023;4:178-94.
90. Gomez R, Watson S, Brown T et al. Personality Inventory for DSM-5 - Brief Form (PID-5-BF): measurement invariance across men and women. Personal Disord 2023;14:334-8.
91. Oitsalu ML, Kreegipuu M, Hutsebaut J. Psychometric evaluation of the Estonian version of the Semi-Structured Interview for Personality Functioning DSM-5 (STiP-5.1). Borderline Personal Disord Emot Dysregul 2022;9:28.
92. Heissler R, Doubkova N, Hutsebaut J et al. Semi-Structured Interview for Personality Functioning DSM-5 (STiP-5.1): psychometric evaluation of the Czech version. Personal Ment Health 2021;15:198-207.
93. Asadi S, Cunningham TJ, Morgan TA et al. Examining measurement invariance in the Personality Inventory for DSM-5 Brief Form across sexual and gender minority status. Assessment 2024;31:678-97.
94. Rodriguez-Seijas C. Sociocultural context and the DSM-5 Alternative Model of Personality Disorder: Comment on Widiger and Hines (2022). Personal Disord 2022;13:356-9.
95. Eggermont K, Smits D, Bogaerts A et al. A brief screener for impairment in personality functioning: psychometric validation of the Five-Item Screening Scale for Personality Disorders in a Dutch-speaking clinical sample. Personal Disord 2024;15:269-74.
96. Vergauwe J, Rouco V, Franssens R et al. Age and sex differences of the PID-5-100 maladaptive personality traits throughout adulthood. Personal Disord 2023;14:512-26.
97. Stone-Bury L, Premovich A, Segal D. Evidence of age bias on the Alternative Model of Personality Disorders. Innov Aging 2023;7(Suppl. 1):940.
98. Morey LC. Development and initial evaluation of a self-report form of the DSM-5 Level of Personality Functioning Scale. Psychol Assess 2017;29:1302-8.
99. Somma A, Krueger RF, Markon KE et al. The replicability of the Personality Inventory for DSM-5 domain scale factor structure in U.S. and non-U.S. samples: a quantitative review of the published literature. Psychol Assess 2019;31:861-77.
100. Freilich CD, Palumbo IM, Latzman RD et al. Assessing the measurement invariance of the Personality Inventory for DSM-5 across Black and White Americans. Psychol Assess 2023;35:721-8.
101. Becker LG, Asadi S, Zimmerman M et al. Is there a bias in the diagnosis of borderline personality disorder among racially minoritized patients? Personal Disord 2023;14:339-46.
102. Hricovec MM, Su CC, Bart TA et al. Measurement invariance of the Personality Inventory for the DSM-5 across U.S. East Asian, Southeast Asian, and White participants. Personal Disord 2024;15:1-10.
103. Lau SCL, Sellbom M, Bagby RM. An examination of the cross-cultural equivalence of the Personality Inventory for DSM-5 across Chinese and U.S. samples. Psychol Assess 2024;36:102-13.
104. Widiger TA. Invited essay: Sex biases in the diagnosis of personality disorders. J Pers Disord 1998;12:95-118.
105. Garb HN. Race bias and gender bias in the diagnosis of psychological disorders. Clin Psychol Rev 2021;90:102087.
106. Eaton NR, Greene AL. Personality disorders: community prevalence and socio-demographic correlates. Curr Opin Psychol 2018;21:28-32.
107. Hutsebaut J, Kamphuis JH, Feenstra DJ et al. Assessing DSM-5-oriented Level of Personality Functioning: development and psychometric evaluation of the Semi-Structured Interview for Personality Functioning DSM-5 (STiP-5.1). Personal Disord 2017;8:94-101.
108. Zettl M, Volkert J, Vogele C et al. Mentalization and criterion A of the alternative model for personality disorders: results from a clinical and nonclinical sample. Personal Disord 2020;11:191-201.
109. Borroni S, Ruotolo G, Geminiani C et al. DSM-5 personality dysfunction in a mentalization perspective: the interplay among personality dysfunction, emotion dysregulation and mentalization ability in a sample of Italian community dwelling-adults. Mediterr J Clin Psychol 2024;12:1.
110. Babinski DE, Kujawa A, Pegg S et al. Social and monetary reward processing in youth with early emerging personality pathology: an RDoC-informed study. Res Child Adolesc Psychopathol 2024;52:567-78.
111. Bach B, Bernstein DP. Schema therapy conceptualization of personality functioning and traits in ICD-11 and DSM-5. Curr Opin Psychiatry 2019;32:38-49.
112. Ohse L, Zimmermann J, Kerber A et al. Impairments in cognitive and emotional empathy as markers of general versus specific personality pathology. Psychopathology 2024;57:136-48.
113. Roche MJ. Examining the alternative model for personality disorder in daily life: evidence for incremental validity. Personal Disord 2018;9:574-83.
114. Roche MJ, Jacobson NC, Pincus AL. Using repeated daily assessments to uncover oscillating patterns and temporally-dynamic triggers in structures of psychopathology: applications to the Alternative Model of Personality Disorders. J Abnorm Psychol 2016;125:1090-102.
115. Eggermont K, Raymaekers K, Claes L et al. Impairment in personality functioning throughout adolescence and co-development with personality traits, emotion regulation strategies, and psychopathology. J Res Pers 2023;104:104380.
116. Fossati A, Somma A, Borroni S et al. Executive functioning correlates of maladaptive personality traits: initial evidence from an Italian sample of consecu-

tively admitted adult outpatients. J Psychopathol Behav 2018;40:484-96.
117. Roye S, Calamia M, Castagna PJ et al. Normative and maladaptive personality traits and self-reported executive functioning. Assessment 2022;29:499-507.
118. Wissing BG, Reinhard MA. The Dark Triad and the PID-5 maladaptive personality traits: accuracy, confidence and response bias in judgments of veracity. Front Psychol 2017;8:1549.
119. Boone K, Babinski D, Kujawa A et al. The incremental validity of level of personality functioning over borderline personality features in associations with early adolescent social reward processing. Personal Ment Health 2025;19:e70000.
120. Heiland AM, Veilleux JC. Severity of personality dysfunction predicts affect and self-efficacy in daily life. Pers Disord Theory Res Treat 2021;12:560-9.
121. Asgarizadeh A, Sharp C, Ghanbari S. Shame-coping clusters: comparisons regarding attachment insecurities, mentalizing deficits, and personality pathology, controlling for general emotion dysregulation. Borderline Personal Disord Emot Dysregul 2023;10:25.
122. Ringwald WR, Hopwood CJ, Pilkonis PA et al. Dynamic features of affect and interpersonal behavior in relation to general and specific personality pathology. Pers Disord Theory Res Treat 2021;12:365-76.
123. Shalala N, Tan J, Biberdzic M. The mediating role of identity disturbance in the relationship between emotion dysregulation, executive function deficits, and maladaptive personality traits. Pers Indiv Differ 2020;162:10004.
124. Abdi R, Pak R. The mediating role of emotion dysregulation as a transdiagnostic factor in the relationship between pathological personality dimensions and emotional disorders symptoms severity. Pers Indiv Differ 2019;142:282-7.
125. Pollock NC, McCabe GA, Southard AC et al. Pathological personality traits and emotion regulation difficulties. Pers Indiv Differ 2016;95:168-77.
126. Tracy M, Sharpe L, Bach B et al. Connecting DSM-5 and ICD-11 trait domains with schema therapy and dialectical behavior therapy constructs. Personal Ment Health 2023;17:208-19.
127. Winterstein S, Mazzucchelli TG, Gross JJ et al. Alexithymia and personality pathology. J Pers Disord 2024;38:435-54.
128. Biberdzic M, Tang J, Tan JH. Beyond difficulties in self-regulation: the role of identity integration and personality functioning in young women with disordered eating behaviours. J Eat Disord 2021;9:93.
129. Clark LA. Temperament as a unifying basis for personality and psychopathology. J Abnorm Psychol 2005;114:505-21.
130. Emery LT, Evans CM, Dimitrova J et al. Understanding the association between normal and maladaptive personality traits: replication and extension of Morey et al. (2020). J Pers 2024;92:1477-91.
131. Morey LC, Good EW, Hopwood CJ. Global personality dysfunction and the relationship of pathological and normal trait domains in the DSM-5 Alternative Model for Personality Disorders. J Pers 2022;90:34-46.
132. Widiger TA, McCabe GA. The Alternative Model of Personality Disorders (AMPD) from the perspective of the Five-Factor Model. Psychopathology 2020; 53:149-56.
133. Gamache D, Savard C, Leclerc P et al. Latent profiles of patients with borderline pathology based on the Alternative DSM-5 Model for Personality Disorders. Borderline Personal Disord Emot Dysregul 2021;8:4.
134. McCabe GA, Oltmanns JR, Widiger TA. Criterion A scales: convergent, discriminant, and structural relationships. Assessment 2021;28:813-28.
135. Kerr S, McLaren V, Cano K et al. Levels of Personality Functioning Questionnaire 12-18 (LoPF-Q 12-18): factor structure, validity, and clinical cut-offs. Assessment 2023;30:1764-76.
136. Paap MCS, Hummelen B, Braeken J et al. A multi-center psychometric evaluation of the Severity Indices of Personality Problems 118 (SIPP-118): do we really need all those facets? Qual Life Res 2021;30:567-75.
137. Eggermont K, Luyckx K, Smits D et al. The validation of a Five-Item Screening Scale for Personality Disorders in Dutch-speaking community adolescents and adults. J Psychopathol Behav 2022;44:418-31.
138. Ohse L, Zimmermann J, Kerber A et al. Reliability, structure, and validity of Module I (personality functioning) of the Structured Clinical Interview for the Alternative DSM-5 Model for Personality Disorders (SCID-5-AMPD-I). Personal Disord 2023;14:287-99.
139. Roche MJ, Jacobson NC, Phillips JJ. Expanding the validity of the Level of Personality Functioning Scale Observer Report and Self-Report Versions across psychodynamic and interpersonal paradigms. J Pers Assess 2018;100:571-80.
140. Balzen KM, Kerr S, Gecha T et al. First psychometric evaluation of the English version of the Semi-Structured Interview for Personality Functioning (STiP-5.1). J Pers Assess 2025;107:1-11.
141. Bogaerts A, de Moor EL, Lind M. Identity disturbance in dimensional and categorical models of personality disorder: the incremental value of self-rated identity and narrative identity. Personal Disord 2024;15:479-91.
142. Weekers LC, Verhoeff SCE, Kamphuis JH et al. Assessing Criterion A in adolescents using the Semi-Structured Interview for Personality Functioning DSM-5. Personal Disord 2021;12:312-9.
143. Dowgwillo EA, Roche MJ, Pincus AL. Examining the interpersonal nature of Criterion A of the DSM-5 Section III Alternative Model for Personality Disorders using bootstrapped confidence intervals for the interpersonal circumplex. J Pers Assess 2018;100:581-92.
144. Zettl M, Taubner S, Hutsebaut J et al. Psychometric evaluation of the German version of the Semi-Structured Interview for Personality Functioning DSM-5 (STiP-5.1). Psychother Psychosom Med Psychol 2019;69:499-504.
145. De Fruyt F, De Clercq B, De Bolle M et al. General and maladaptive traits in a five-factor framework for DSM-5 in a university student sample. Assessment 2013;20:295-307.
146. Few LR, Miller JD, Rothbaum AO et al. Examination of the Section III DSM-5 Diagnostic System for Personality Disorders in an outpatient clinical sample. J Abnorm Psychol 2013;122:1057-69.
147. Zinchuk M, Kustov G, Bach B et al. Evaluation of a 36-item measure of ICD-11 and DSM-5 personality disorder trait domains and facets in Russian inpatients. Psychol Assess 2023;35:e22-30.
148. Aluja A, García LF, Cuevas L et al. Dimensional pathological personality predicting personality disorders: comparison of the DAPP-BQ and PID-5 shortened versions in a Spanish community sample. J Psychopathol Behav 2019;41: 160-73.
149. Zimmermann J, Altenstein D, Krieger T et al. The structure and correlates of self-reported DSM-5 maladaptive personality traits: findings from two German-speaking samples. J Pers Disord 2014;28:518-40.
150. Wright AGC, Simms LJ. On the structure of personality disorder traits: conjoint analyses of the CAT-PD, PID-5, and NEO-PI-3 trait models. Personal Disord 2014;5:43-54.
151. Watters CA, Sellbom M, Uliaszek AA et al. Clarifying the interstitial nature of facets from the Personality Inventory for DSM-5 using the five factor model of personality. Personal Disord 2019;10:330-9.
152. Debast I, Rossi G, van Alphen SPJ. Construct validity of the Section III maladaptive trait domains in older adults. J Pers Disord 2017;31:671-88.
153. Al-Dajani N, Gralnick TM, Bagby RM. A psychometric review of the Personality Inventory for DSM-5 (PID-5): current status and future directions. J Pers Assess 2016;98:62-81.
154. De Caluwe E, Verbeke L, Aken MV et al. The DSM-5 trait measure in a psychiatric sample of late adolescents and emerging adults: structure, reliability, and validity. J Pers Disord 2019;33:101-18.
155. Fossati A, Somma A, Borroni S et al. A head-to-head comparison of the Personality Inventory for DSM-5 (PID-5) with the Personality Diagnostic Questionnaire-4 (PDQ-4) in predicting the general level of personality pathology among community dwelling subjects. J Pers Disord 2016;30:82-94.
156. Hare RD. The Hare Psychopathy Checklist Revised, 2nd ed. Toronto: Multi-Health Systems, 2003.
157. Wygant DB, Sellbom M, Sleep CE et al. Examining the DSM-5 Alternative Personality Disorder Model operationalization of antisocial personality disorder and psychopathy in a male correctional sample. Personal Disord 2016;7:229-39.
158. Cruitt PJ, Boudreaux MJ, King HR et al. Examining Criterion A: DSM-5 Level of Personality Functioning as assessed through life story interviews. Personal Disord 2019;10:224-34.
159. Amini M, Pourshahbaz A, Mohammadkhani P et al. The DSM-5 levels of personality functioning and severity of Iranian patients with antisocial and borderline personality disorders. Iran Red Crescent Med J 2015;17:e19885.
160. Busmann M, Wrege J, Meyer AH et al. Alternative Model for Personality Disorders (DSM-5) predicts dropout in inpatient psychotherapy for patients with personality disorder. Front Psychol 2019;10:952.
161. Dereboy F, Dereboy C, Eskin M. Validation of the DSM-5 Alternative Model for Personality Disorders diagnoses in Turkey, Part 1: LEAD validity and reliability of the personality functioning ratings. J Pers Assess 2018;100:603-11.
162. Paap MCS, Pedersen G, Kvarstein E et al. Evaluating the construct validity of the Norwegian Version of the Level of Personality Functioning Scale - Brief Form 2.0 in a large clinical sample. J Pers Assess 2023;106:49-59.
163. Di Pierro R, Gargiulo I, Poggi A et al. The Level of Personality Functioning Scale applied to clinical material from the Structured Interview of Personality Organization (STIPO): utility in detecting personality pathology. J Pers Disord 2020;34:62-76.
164. Hopwood CJ, Good EW, Morey LC. Validity of the DSM-5 Levels of Personality Functioning Scale - Self Report. J Pers Assess 2018;100:650-9.
165. Bliton CF, Roche MJ, Pincus AL et al. Examining the structure and validity of self-report measures of DSM-5 Alternative Model for Personality Disorders Criterion A. J Pers Disord 2022;36:157-82.

166. Thomson M, Cavelti M, Lerch S et al. Clinical profiles of adolescent personality pathology: a latent structure examination of the Semi-Structured Interview for Personality Functioning DSM-5 (STiP-5.1) in a help-seeking sample. Borderline Personal Disord Emot Dysregul 2024;11:9.
167. Diondet S, Weekers LC, Hutsebaut J. Associations between borderline personality disorder symptoms and personality functioning in adolescents: a brief report. Personal Disord 2024;15:264-8.
168. Sharp C, Cervantes BR. Maladaptive self- and interpersonal functioning increments general psychiatric severity in the association with adolescent personality pathology. Children 2023;10:120.
169. Sharp C, Kerr S, Barkauskiene R. The incremental utility of maladaptive self and identity functioning over general functioning for borderline personality disorder features in adolescents. Personal Disord 2022;13:474-81.
170. Barkauskiene R, Sharp C, Kerr S et al. Examining the validity of the levels of Personality Functioning Questionnaire for Adolescents Aged 12-18 (LOPF-Q 12-18): a replication and extension with a sample of Lithuanian adolescents. J Pers Disord 2024;38:330-49.
171. Liggett J, Carmichael KL, Smith A et al. Validation of self-report impairment measures for Section III obsessive-compulsive and avoidant personality disorders. J Pers Assess 2017;99:1-14.
172. Vizgaitis AL, Lenzenweger MF. Identity and personality pathology: a convergence across the DSM-5 Personality Disorder Model and the Alternative Model for Personality Disorders. J Pers Disord 2022;36:537-58.
173. Bender DS, Morey LC, Skodol AE. Toward a model for assessing level of personality functioning in DSM-5, part I: a review of theory and methods. J Pers Assess 2011;93:332-46.
174. Fowler JC, Madan A, Allen JG et al. Clinical utility of the DSM-5 alternative model for borderline personality disorder: differential diagnostic accuracy of the BFI, SCID-II-PQ, and PID-5. Compr Psychiatry 2018;80:97-103.
175. Morey LC, Benson KT. Relating DSM-5 Section II and Section III personality disorder diagnostic classification systems to treatment planning. Compr Psychiatry 2016;68:48-55.
176. Liggett J, Sellbom M. Examining the DSM-5 alternative model of personality disorders operationalization of obsessive-compulsive personality disorder in a mental health sample. Personal Disord 2018;9:397-407.
177. Liggett J, Sellbom M, Bach B. Continuity between DSM-5 Section II and Section III personality traits for obsessive-compulsive personality disorder. Clin Psychol Psychother 2018;25:144-51.
178. Liggett J, Sellbom M, Carmichael KLC. Examining the DSM-5 Section III criteria for obsessive-compulsive personality disorder in a community sample. J Pers Disord 2017;31:790-809.
179. Hopwood CJ, Thomas KM, Markon KE et al. DSM-5 personality traits and DSM-IV personality disorders. J Abnorm Psychol 2012;121:424-32.
180. Fossati A, Krueger RF, Markon KE et al. Reliability and validity of the Personality Inventory for DSM-5 (PID-5): predicting DSM-IV personality disorders and psychopathy in community-dwelling Italian adults. Assessment 2013;20:689-708.
181. Anderson JL, Snider S, Sellbom M et al. A comparison of the DSM-5 Section II and Section III personality disorder structures. Psychiatry Res 2014;216:363-72.
182. Anderson JL, Sellbom M, Wygant DB et al. Examining the associations between DSM-5 Section III antisocial personality disorder traits and psychopathy in community and university samples. J Pers Disord 2014;28:675-97.
183. Samuel DB, Hopwood CJ, Krueger RF et al. Comparing methods for scoring personality disorder types using maladaptive traits in DSM-5. Assessment 2013;20:353-61.
184. Sellbom M, Sansone RA, Songer DA et al. Convergence between DSM-5 Section II and Section III diagnostic criteria for borderline personality disorder. Aust N Z J Psychiatry 2014;48:325-32.
185. Yam WH, Simms LJ. Comparing criterion- and trait-based personality disorder diagnoses in DSM-5. J Abnorm Psychol 2014;123:802-8.
186. Crego C, Gore WL, Rojas SL et al. The discriminant (and convergent) validity of the Personality Inventory for DSM-5. Personal Disord 2015;6:321-35.
187. Jopp AM, South SC. Investigating the Personality Inventory for DSM-5 using self and spouse reports. J Pers Disord 2015;29:193-214.
188. Bastiaens T, Smits D, De Hert M et al. DSM-5 Section III personality traits and Section II personality disorders in a Flemish community sample. Psychiatry Res 2016;238:290-8.
189. Rojas SL, Widiger TA. Coverage of the DSM-IV-TR/DSM-5 Section II personality disorders with the DSM-5 Dimensional Trait Model. J Pers Disord 2017;31:462-82.
190. Watters CA, Bagby RM, Sellbom M. Meta-analysis to derive an empirically based set of personality facet criteria for the Alternative DSM-5 Model for Personality Disorders. Personal Disord 2019;10:97-104.
191. McCabe GA, Widiger TA. Discriminant validity of the Alternative Model of Personality Disorder. Psychol Assess 2020;32:1158-71.
192. Komasi S, Hemmati A, Rezaei F et al. Comparison of the relative sensitivity of two dimensional personality models to the psychopathological symptoms: the Section III DSM-5 maladaptive traits versus affective temperaments. BMC Psychiatry 2022;22:503.
193. Nysaeter TE, Hummelen B, Christensen TB et al. The incremental utility of Criteria A and B of the DSM-5 Alternative Model for Personality Disorders for predicting DSM-IV/DSM-5 Section II personality disorders. J Pers Assess 2023;105:111-20.
194. Miller JD, Few LR, Lynam DR et al. Pathological personality traits can capture DSM-IV personality disorder types. Personal Disord 2015;6:32-40.
195. Clark LA, Ro E, Vittengl JR et al. Longitudinal prediction of psychosocial functioning outcomes: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Section-II personality disorders versus alternative model personality dysfunction and traits. Personal Disord 2024;15:341-51.
196. Sellbom M, Carmichael KLC, Liggett J. Examination of DSM-5 Section III avoidant personality disorder in a community sample. Personal Ment Health 2017;11:299-313.
197. Orbons IMJ, Rossi G, Verheul R et al. Continuity between DSM-5 Section II and III personality disorders in a Dutch clinical sample. J Pers Assess 2019;101:274-83.
198. Morey LC, Benson KT, Skodol AE. Relating DSM-5 Section III personality traits to Section II personality disorder diagnoses. Psychol Med 2016;46:647-55.
199. Anderson J, Kelley SE. Antisocial personality disorder and psychopathy: the AMPD in review. Personal Disord 2022;13:397-401.
200. Clark LA, Ro E, Nuzum H et al. Personality disorder coverage, prevalence, and convergence: do the DSM-5's two models of personality disorder identify the same patients? Psychol Med 2024;54:2210-21.
201. Few LR, Lynam DR, Maples JL et al. Comparing the utility of DSM-5 Section II and III antisocial personality disorder diagnostic approaches for capturing psychopathic traits. Personal Disord 2015;6:64-74.
202. Zimmermann J, Bohnke JR, Eschstruth R et al. The latent structure of personality functioning: investigating criterion A from the Alternative Model for Personality Disorders in DSM-5. J Abnorm Psychol 2015;124:532-48.
203. Vittengl JR, Jarrett RB, Ro E et al. How can the DSM-5 Alternative Model of Personality Disorders advance understanding of depression? J Affect Disord 2023;320:254-62.
204. Pimentel FU, Oliveira SES. Personality functioning, positive outlook for the future, and simple and complex post-traumatic stress disorder. Acta Psychol 2024;244:104165.
205. Sleep CE, Weiss B, Lynam DR et al. The DSM-5 Section III personality disorder criterion A in relation to both pathological and general personality traits. Personal Disord 2020;11:202-12.
206. Schumacher AM, Tschitsaz A, Lerch S et al. Impairments in personality functioning in adolescents with anorexia nervosa. Eur Eat Disord Rev 2025;33:360-73.
207. Kotov R, Krueger RF, Watson D et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): a quantitative nosology based on consensus of evidence. Annu Rev Clin Psychol 2021;17:83-108.
208. Krueger B, Markon KE. The role of the DSM-5 personality trait model in moving toward a quantitative and empirically based approach to classifying personality and psychopathology. Annu Rev Clin Psychol 2014;10:1-25.
209. Krueger RF, Eaton NR. Personality traits and the classification of mental disorders: toward a more complete integration in DSM-5 and an empirical model of psychopathology. Personal Disord 2010;1:97-118.
210. Krueger RF, Hobbs KA, Conway CC et al. Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): II. Externalizing superspectrum. World Psychiatry 2021;20:171-93.
211. Kudinova AY, MacPherson HA, Musella K et al. Maladaptive personality traits and the course of suicidal ideation in young adults with bipolar disorder: cross-sectional and prospective approaches. Suicide Life Threat Behav 2021;51:394-402.
212. Miralles C, Alonso Y, Algora MJ et al. Maladaptive personality traits in patients with recent-onset psychosis: a case-control study using the Personality Inventory for the DSM-5 (PID-5). Schizophr Res 2023;252:216-24.
213. Longenecker JM, Krueger RF, Sponheim SR. Personality traits across the psychosis spectrum: a Hierarchical Taxonomy of Psychopathology conceptualization of clinical symptomatology. Personal Ment Health 2020;14:88-105.
214. Bastiaens T, Smits D, De Hert M et al. The relationship between the Personality Inventory for the DSM-5 (PID-5) and the psychotic disorder in a clinical sample. Assessment 2019;26:315-23.

215. Moraleda-Barreno E, Diaz-Batanero C, Perez-Moreno PJ et al. Relations between facets and personality domains with impulsivity: new evidence using the DSM-5 Section III framework in patients with substance use disorders. Personal Disord 2018;9:490-5.
216. Lau C, Bagby RM, Pollock BG et al. Five-Factor Model and DSM-5 Alternative Model of Personality Disorder profile construction: associations with cognitive ability and clinical symptoms. J Intell 2023;11:71.
217. Jacobsson P, Hopwood CJ, Krueger RF et al. Conceptualizing adult ADHD with the DSM Alternative Model of Personality Disorder. Personal Ment Health 2024; 18:369-86.
218. Muscogiuri G, Dimaggio L, Giani L et al. Personality traits in migraineurs: a case-control study by Personality Inventory for DSM-5 (PID-5). Neurol Sci 2018; 39(Suppl. 1):129-30.
219. Nelson SM, Huprich SK, Meehan KB et al. Convergent and discriminant validity and utility of the DSM-5 Levels of Personality Functioning Questionnaire (DLOPFQ): associations with medical health care provider ratings and measures of physical health. J Pers Assess 2018;100:671-9.
220. Siefert CJ, Sexton J, Meehan K et al. Development of a short form for the DSM-5 Levels of Personality Functioning Questionnaire. J Pers Assess 2020;102:516-26.
221. Gamache D, Savard C, Leclerc P et al. Introducing a short self-report for the assessment of DSM-5 Level of Personality Functioning for personality disorders: the Self and Interpersonal Functioning Scale. Personal Disord 2019;10:438-47.
222. Bach B, Hutsebaut J. Level of Personality Functioning Scale - Brief Form 2.0: utility in capturing personality problems in psychiatric outpatients and incarcerated addicts. J Pers Assess 2018;100:660-70.
223. Natoli AP, Rodriguez CM. A new performance-based measure of personality functioning impairment: development and preliminary evaluation of reliability and validity. Discov Ment Health 2024;4:6.
224. Huprich SK, Nelson SM, Meehan KB et al. Introduction of the DSM-5 levels of Personality Functioning Questionnaire. Personal Disord 2018;9:553-63.
225. Christensen TB, Eikenaes I, Hummelen B et al. Level of Personality Functioning as a predictor of psychosocial functioning – concurrent validity of Criterion A. Personal Disord 2020;11:79-90.
226. Kunz JI, Frey A, Bertsch K et al. Loneliness is associated with lower self- and clinician-rated levels of personality functioning. J Pers Disord 2023;37:724-40.
227. Kwak S, Kwon J. General and distinct correspondence between personality disorder and interpersonal functioning. Acta Psychol 2024;249:104454.
228. Barkauskiene R, Gaudiesiute E, Grigaite A et al. Criteria A and B of the Alternative Model for Personality Disorders (AMPD) as prospective predictors of psychosocial functioning in community-based adolescents: a one-year follow-up study. Personal Disord 2025;16:286-96.
229. Garcia DJ, Waugh MH, Skadberg RM et al. Deconstructing criterion A of the Alternative Model for Personality Disorders. Personal Disord 2021;12:320-30.
230. Sekowski M, Gambin M, Sharp C. The relations between identity disturbances, borderline features, internalizing disorders, and suicidality in inpatient adolescents. J Pers Disord 2021;35(Suppl. B):29-47.
231. Hertel C, Cavelti M, Lerch S et al. Does personality dysfunction add incremental utility over general psychopathology when modeling previous suicide attempts in adolescent patients? J Affect Disord 2024;361:465-71.
232. Iannattone S, Schuiringa HD, Aleva A et al. Unravelling the longitudinal relations between developmental milestones, general psychopathology, and personality functioning in a youth clinical sample. J Youth Adolesc 2024;53:1887-902.
233. Kvarstein EH, Froyhaug M, Pettersen MS et al. Improvement of personality functioning among people treated within personality disorder mental health services. A longitudinal, observational study. Front Psychiatry 2023;14: 1163347.
234. Garcia DJ, Skadberg RM, Schmidt M et al. It's not that difficult: an interrater reliability study of the DSM-5 Section III Alternative Model for Personality Disorders. J Pers Assess 2018;100:612-20.
235. Chauhan K, Donahue J, Thompson R. The predictive validity of the DSM-5 alternative model for borderline personality disorder: associations with coping strategies, general distress, rumination, and suicidal ideation across one year. Personal Ment Health 2023;17:272-81.
236. Meliante M, Rossi C, Malvini L et al. The relationship between PID-5 personality traits and mental states. A study on a group of young adults at risk of psychotic onset. Medicina 2021;57:33.
237. Roche MJ, Natoli AP, Moore J. Personality dysfunction linked to future aggression in daily life: findings from two experience sampling studies. J Threat Assess Manag 2023;10:281-300.
238. Kaurin A, Do QB, Ladouceur CD et al. Daily manifestations of caregiver- and self-reported maladaptive personality traits in adolescent girls. Personal Disord 2023;14:490-500.
239. Freilich CD, McGue M, South SC et al. Connecting loneliness with pathological personality traits: evidence for genetic and environmental mediation from a study of older twins. Personal Disord 2024;15:34-45.
240. Aboul-Ata MA, Qonsua FT, Saadi IAA. Personality pathology and suicide risk: examining the relationship between DSM-5 alternative model traits and suicidal ideation and behavior in college-aged individuals. Psychol Rep 2023; doi: 10.1177/00332941231218940.
241. Mundt JC, Marks IM, Shear MK et al. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. Br J Psychiatry 2002;180:461-4.
242. American Psychiatric Association. Global Assessment of Functioning Scale. Diagnostic and Statistical Manual of Mental Disorders (4th ed). Washington: American Psychiatric Association, 1994.
243. Herpertz SC, Bertsch K, Jeung H. Neurobiology of Criterion A: self and interpersonal personality functioning. Curr Opin Psychol 2018;21:23-7.
244. Pintos Lobo R, Bottenhorn KL, Riedel MC et al. Neural systems underlying RDoC social constructs: an activation likelihood estimation meta-analysis. Neurosci Biobehav Rev 2023;144:104971.
245. Bowyer CB, Joyner KJ, Latzman RD et al. A model-based strategy for interfacing traits of the DSM-5 AMPD with neurobiology. J Pers Disord 2020;34:586-608.
246. Kolla NJ, Tully J, Bertsch K. Neural correlates of aggression in personality disorders from the perspective of DSM-5 maladaptive traits: a systematic review. Transl Psychiatry 2023;13:330.
247. Papamalis FE, Kalyva E, Teare MD et al. The role of personality functioning in drug misuse treatment engagement. Addiction 2020;115:726-39.
248. Woodbridge J, Townsend M, Reis S et al. Non-response to psychotherapy for borderline personality disorder: a systematic review. Aust N Z J Psychiatry 2022; 56:771-87.
249. Kiel L, Hopwood CJ, Lind M. Changes in personality functioning and pathological personality traits as a function of treatment: a feasibility study. J Psychopathol Behav 2024;46:695-704.
250. Sharp C, Kulesz P, Kerr S. Prospective prediction of treatment outcomes in adolescents: a head-to-head comparison of alternative model for personality disorder versus borderline personality disorder. Personal Disord 2024;15:379-85.
251. Torres-Soto JF, Iborra-Giner P, Giner-Alegría CA et al. Results of a day-hospital program for personality disorders. Application of the PID-5 and DSM-5 dimensional model. Clínica y Salud 2021;32:41-7.
252. Fowler JC, Lamkin J, Allen JG et al. Personality trait domains predict psychiatric symptom and functional outcomes. Psychotherapy 2022;59:38-47.
253. Roberts BW, Luo J, Briley DA et al. A systematic review of personality trait change through intervention. Psychol Bull 2017;143:117-41.
254. Hessels CJ, de Moor EL, Deutz MHF et al. Personality pathology in youth: a comparison of the categorical and alternative model in relation to internalizing and externalizing pathology and age-adequate psychosocial functioning. Personal Disord 2024;15:293-303.
255. Hines A, Kushner ML, Stumpp N et al. Different routes to the same destination? Comparing Diagnostic and Statistical Manual of Mental Disorders, fifth edition Section II- and Alternative Model of Personality Disorder-defined borderline personality disorder. Personal Disord 2024;15:352-60.
256. Ringwald WR, Woods WC, Wright AGC. Comparing the Diagnostic and Statistical Manual of Mental Disorders, fifth edition, personality disorder models scored from the same interview. Personal Disord 2024;15:371-8.
257. Bagby RM, Lau SCL, Watters CA et al. A comparison of the associations of the Diagnostic and Statistical Manual of Mental Disorders, fifth edition, Section II personality disorders and Section III personality domains with clinical dysfunction in a psychiatric patient sample. Personal Disord 2024;15:322-31.
258. Calabrese WR, Emery LT, Evans CM et al. Diagnostic and Statistical Manual of Mental Disorders, fifth edition, personality disorders and the alternative model: prediction of naturalistically observed behavior, interpersonal functioning, and psychiatric symptoms, 1 year later. Personal Disord 2024;15:361-70.
259. Cosgun S, Goth K, Cakiroglu S. Levels of Personality Functioning Questionnaire (LoPF-Q) 12-18 Turkish version: reliability, validity, factor structure and relationship with comorbid psychopathology in a Turkish adolescent sample. J Psychopathol Behav 2021;43:620-31.
260. Goth K, Birkhölzer M, Schmeck K. LoPF-Q 12-18 (Levels of Personality Functioning Questionnaire) German version: a self-report questionnaire for measuring personality functioning in adolescence – short description. www.academic-tests.com.
261. Zimmermann R, Steppan M, Zimmermann J et al. A DSM-5 AMPD and ICD-11 compatible measure for an early identification of personality disorders in adolescence – LoPF-Q 12-18 latent structure and short form. PLoS One

2022;17:e0269327.
262. Leichsenring F, Fonagy P, Heim N et al. Borderline personality disorder: a comprehensive review of diagnosis and clinical presentation, etiology, treatment, and current controversies. World Psychiatry 2024;23:4-25.
263. Sharp C. Personality disorders. N Engl J Med 2022;387:916-23.
264. Winsper C. Borderline personality disorder: course and outcomes across the lifespan. Curr Opin Psychol 2021;37:94-7.
265. Wright AG, Calabrese WR, Rudick MM et al. Stability of the DSM-5 Section III pathological personality traits and their longitudinal associations with psychosocial functioning in personality disordered individuals. J Abnorm Psychol 2015;124:199-207.
266. Simon W. Follow-up psychotherapy outcome of patients with dependent, avoidant and obsessive-compulsive personality disorders: a meta-analytic review. Int J Psychiatry Clin Pract 2009;13:153-65.
267. Thylstrup B, Simonsen S, Nemery C et al. Assessment of personality-related levels of functioning: a pilot study of clinical assessment of the DSM-5 level of personality functioning based on a semi-structured interview. BMC Psychiatry 2016;16:298.
268. Zimmermann J, Benecke C, Bender DS et al. Assessing DSM-5 level of personality functioning from videotaped clinical interviews: a pilot study with untrained and clinically inexperienced students. J Pers Assess 2014;96:397-409.
269. Preti E, Di Pierro R, Costantini G et al. Using the Structured Interview of Personality Organization for DSM-5 Level of Personality Functioning rating performed by inexperienced raters. J Pers Assess 2018;100:621-9.
270. Morey LC. Interdiagnostician reliability of the DSM-5 Section II and Section III Alternative Model criteria for borderline personality disorder. J Pers Disord 2019;33:1-33.
271. Hutsebaut J, Weekers LC, Tuin N et al. Assessment of ICD-11 personality disorder severity in forensic patients using the Semi-Structured Interview for Personality Functioning DSM-5 (STiP-5.1): preliminary findings. Front Psychiatry 2021;12:617702.
272. Young SE, Beazley P. Interrater reliability of criterion A of the Alternative Model for Personality Disorder (Diagnostic and Statistical Manual for Mental Disorders, Fifth Edition - Section III): a meta-analysis. Personal Disord 2023;14:613-24.
273. Kraemer HC, Kupfer DJ, Clarke DE et al. DSM-5: how reliable is reliable enough? Am J Psychiatry 2012;169:13-5.
274. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155-63.
275. Regier DA, Narrow WE, Clarke DE et al. DSM-5 field trials in the United States and Canada, Part II: Test-retest reliability of selected categorical diagnoses. Am J Psychiatry 2013;170:59-70.
276. Goth K, Birkholzer M, Schmeck K. Assessment of personality functioning in adolescents with the LoPF-Q 12-18 Self-Report Questionnaire. J Pers Assess 2018; 100:680-90.
277. Ottosson H, Grann M, Kullgren G. Test-retest reliability of a self-report questionnaire for DSM-IV and ICD-10 personality disorders. Eur J Psychol Assess 2000;16:53-8.
278. Osone A, Takahashi S. Twelve month test-retest reliability of a Japanese version of the Structured Clinical Interview for DSM-IV Personality Disorders. Psychiatry Clin Neurosci 2003;57:532-8.
279. Wright AGC, Simms LJ. Stability and fluctuation of personality disorder features in daily life. J Abnorm Psychol 2016;125:641-56.
280. Fossati A, Somma A, Borroni S et al. The Personality Inventory for DSM-5 Brief Form: evidence for reliability and construct validity in a sample of community-dwelling Italian adolescents. Assessment 2017;24:615-31.
281. Morey LC. Application of the DSM-5 Level of Personality Functioning Scale by lay raters. J Pers Disord 2018;32:709-20.
282. Doubkova N, Heissler R, Preiss M et al. Differences in personality functioning impairment in mood, anxiety, and personality disorders: a cluster analysis. BMC Psychiatry 2022;22:315.
283. Weekers LC, Hutsebaut J, Zimmermann J et al. Changes in the classification of personality disorders: comparing the DSM-5 Section II personality disorder model to the Alternative Model for Personality Disorders using structured clinical interviews. Personal Disord 2022;13:527-35.
284. Spitzer C, Muller S, Kerber A et al. The German version of the Level of Personality Functioning Scale - Brief Form 2.0 (LPFS-BF): latent structure, convergent validity and norm values in the general population. Psychother Psychosom Med Psychol 2021;71:284-93.
285. Bliton CF, Rosenstein LK, Pincus AL. Trading patients: applying the Alternative Model for Personality Disorders to two cases of borderline personality disorder over time and across therapists. Front Psychol 2022;13:794624.
286. Ran T, Balzen KM, Kerr S et al. Psychometric evaluation and clinical utility of the Level of Personality Functioning Scale - Brief Form 2.0. J Pers Disord 2025; 39:185-205.
287. Hemmati A, Morey LC, McCredie MN et al. Validation of the Persian translation of the Level of Personality Functioning Scale - Self-Report (LPFS-SR): comparison of college students and patients with personality disorders. J Psychopathol Behav 2020;42:546-59.
288. Barchi-Ferreira Bel AM, Osorio FL. The Personality Inventory for DSM-5: psychometric evidence of validity and reliability-updates. Harv Rev Psychiatry 2020;28:225-37.
289. Miller JD, Bagby RM, Hopwood CJ et al. Normative data for PID-5 domains, facets, and personality disorder composites from a representative sample and comparison to community and clinical samples. Personal Disord 2022;13:536-41.
290. Morey LC, Berghuis H, Bender DS et al. Toward a model for assessing level of personality functioning in DSM-5, part II: empirical articulation of a core dimension of personality pathology. J Pers Assess 2011;93:347-53.
291. O'Connor BP. A search for consensus on the dimensional structure of personality disorders. J Clin Psychol 2005;61:323-45.
292. Trull TJ, Verges A, Wood PK et al. The structure of DSM-IV-TR personality disorder diagnoses in NESARC: a reanalysis. J Pers Disord 2013;27:727-34.
293. Wright AGC, Zimmermann J. At the nexus of science and practice: answering basic clinical questions in personality disorder assessment and diagnosis with quantitative modeling techniques. In: Huprich S (ed). Personality disorders: assessment, diagnosis, and research. Washington: American Psychological Association, 2015:109-44.
294. Bach B, Markon K, Simonsen E et al. Clinical utility of the DSM-5 Alternative Model of Personality Disorders: six cases from practice. J Psychiatr Pract 2015;21:3-25.
295. Le Corff Y, Lapalme M, Rivard G et al. Prevalence of the Alternative Model of Personality Disorders diagnoses in populational and at-risk samples, gender and age groups comparisons, and normative data for the LPFS-SR and PID-5. Personal Disord 2023;14:591-602.
296. Pires R, Henriques-Calado J, Sousa Ferreira A et al. The utility of ICD-11 and DSM-5 traits for differentiating patients with personality disorders from other clinical groups. Front Psychiatry 2021;12:633882.
297. Hutsebaut J, Feenstra DJ, Kamphuis JH. Development and preliminary psychometric evaluation of a brief self-report questionnaire for the assessment of the DSM-5 Level of Personality Functioning Scale: the LPFS Brief Form (LPFS-BF). Personal Disord 2016;7:192-7.
298. Bach B, Sellbom M, Bo S et al. Utility of DSM-5 Section III personality traits in differentiating borderline personality disorder from comparison groups. Eur Psychiatry 2016;37:22-7.
299. Calvo N, Valero S, Saez-Francas N et al. Borderline personality disorder and Personality Inventory for DSM-5 (PID-5): dimensional personality assessment with DSM-5. Compr Psychiatry 2016;70:105-11.
300. Fowler JC, Madan A, Allen JG et al. Differentiating bipolar disorder from borderline personality disorder: diagnostic accuracy of the Difficulty in Emotion Regulation Scale and Personality Inventory for DSM-5. J Affect Disord 2019;245:856-60.
301. Morey LC, Hopwood CJ. Expert preferences for categorical, dimensional, and mixed/hybrid approaches to personality disorder diagnosis. J Pers Disord 2020;34(Suppl. C):124-31.
302. Morey LC, Skodol AE, Oldham JM. Clinician judgments of clinical utility: a comparison of DSM-IV-TR personality disorders and the alternative model for DSM-5 personality disorders. J Abnorm Psychol 2014;123:398-405.
303. Heltne A, Bode C, Hummelen B et al. Norwegian clinicians' experiences of learnability and usability of SCID-II, SCID-5-PD and SCID-5-AMPD-I interviews: a sequential multi-group qualitative approach. J Pers Assess 2022;104: 599-612.
304. Nelson SM, Huprich SK, Shankar S et al. A quantitative and qualitative evaluation of trainee opinions of four methods of personality disorder diagnosis. Personal Disord 2017;8:217-27.
305. Glover NG, Crego C, Widiger TA. The clinical utility of the Five Factor Model of personality disorder. Personal Disord 2012;3:176-84.
306. Mullins-Sweatt SN, Widiger TA. Clinician's judgments of the utility of the DSM-IV and five-factor models for personality disordered patients. J Pers Disord 2011;25:463-77.
307. Rottman BM, Kim NS, Ahn WK et al. Can personality disorder experts recognize DSM-IV personality disorders from five-factor model descriptions of patient cases? J Clin Psychiatry 2011;72:630-9.
308. Rottman BM, Ahn WK, Sanislow CA et al. Can clinicians recognize DSM-IV personality disorders from five-factor model descriptions of patient cases? Am J Psychiatry 2009;166:427-33.

309. Samuel DB, Widiger TA. Clinicians' judgments of clinical utility: a comparison of the DSM-IV and five-factor models. J Abnorm Psychol 2006;115:298-308.
310. Samuel DB, Widiger TA. Clinicians' use of personality disorder models within a particular treatment setting: a longitudinal comparison of temporal consistency and clinical utility. Personal Ment Health 2011;5:10.1002/pmh.152.
311. Sprock J. A comparative study of the dimensions and facets of the five-factor model in the diagnosis of cases of personality disorder. J Pers Disord 2002;16: 402-23.
312. Lowe JR, Widiger TA. Clinicians' judgments of clinical utility: a comparison of the DSM-IV with dimensional models of general personality. J Pers Disord 2009;23:211-29.
313. Blais MA. Clinician ratings of the five-factor model of personality and the DSM-IV personality disorders. J Nerv Ment Dis 1997;185:388-93.
314. Maffly-Kipp J, Morey LC. Comparing the DSM-5 categorical model of personality disorders and the alternative model of personality disorders regarding clinician judgments of risk and outcome. Personal Disord 2024;15:315-21.
315. Moscicki EK, Clarke DE, Kuramoto SJ et al. Testing DSM-5 in routine clinical practice settings: feasibility and clinical utility. Psychiatr Serv 2013;64:952-60.
316. Cano K, Sharp C. A consumer perspective on personality diagnostic systems: one size does not fit all. J Pers Disord 2023;37:263-84.
317. Pincus AL, Dowgwillo EA, Greenberg LS. Three cases of narcissistic personality disorder through the lens of the DSM-5 alternative model for personality disorders. Pract Innov 2016;1:164-77.
318. Skodol AE, Morey LC, Bender DS et al. The Alternative DSM-5 Model for Personality Disorders: a clinical application. Am J Psychiatry 2015;172:606-13.
319. Schmeck K, Schluter-Muller S, Foelsch PA et al. The role of identity in the DSM-5 classification of personality disorders. Child Adolesc Psychiatry Ment Health 2013;7:27.
320. Bach B, Simonsen S. How does level of personality functioning inform clinical management and treatment? Implications for ICD-11 classification of personality disorder severity. Curr Opin Psychiatry 2021;34:54-63.
321. Hopwood CJ, Kotov R, Krueger RF et al. The time has come for dimensional personality disorder diagnosis. Personal Ment Health 2018;12:82-6.
322. Waugh MH. Clinical utility and application of the DSM-5 Alternative Model for Personality Disorders. In: Hopwood C, Mulay AL, Waugh MH (eds). The DSM-5 Alternative Model for Personality Disorders: integrating multiple paradigms of personality assessment. London: Routledge, 2019:116-75.
323. Weekers LC, Hutsebaut J, Bach B et al. Scripting the DSM-5 Alternative Model for Personality Disorders assessment procedure: a clinically feasible multi-informant multi-method approach. Personal Ment Health 2020;14:304-18.
324. Luo X, Hopwood CJ, Good EW et al. Using interpersonal dimensions of personality and personality pathology to examine momentary and idiographic patterns of alliance rupture. Front Psychol 2021;12:711109.
325. Hopwood CJ, Sellbom M. Implications of DSM-5 personality traits for forensic psychology. Psychol Inj Law 2013;6:314-23.
326. Bornstein RF, Natoli AP. Clinical utility of categorical and dimensional perspectives on personality pathology: a meta-analytic review. Personal Disord 2019; 10:479-90.
327. Keepers GA, Fochtmann LJ, Anzia JM et al. The American Psychiatric Association practice guideline for the treatment of patients with borderline personality disorder. Am J Psychiatry 2024;181:1024-8.
328. Sharp C, Miller JD. Head-to-head comparisons of DSM-5 Section II and Section III personality disorder in predicting clinical outcomes. Personal Disord 2024;15:275-81.
329. World Health Organization. ICD-11 Clinical Descriptions and Diagnostic Requirements for Mental, Behavioural and Neurodevelopmental Disorders. Geneva: World Health Organization, 2024.
330. Bach B, Anderson JL. Patient-reported ICD-11 personality disorder severity and DSM-5 Level of Personality Functioning. J Pers Disord 2020;34:231-49.
331. Bach B, Christensen S, Kongerslev MT et al. Structure of clinician-reported ICD-11 personality disorder trait qualifiers. Psychol Assess 2020;32:50-9.
332. Clark LA, Ro E. Three-pronged assessment and diagnosis of personality disorder and its consequences: personality functioning, pathological traits, and psychosocial disability. Personal Disord 2014;5:55-69.
333. Ringwald WR, Emery L, Khoo S et al. Structure of pathological personality traits through the lens of the CAT-PD model. Assessment 2023;30:2276-95.

DOI:10.1002/wps.21339

Cultural competence in psychotherapy

Despite increasing investment in neurobiological approaches, psychotherapy remains a crucial intervention in psychiatry with wide application for diverse patients, conditions and contexts. However, to have maximum accessibility and impact, psychotherapy must be delivered by culturally competent practitioners and health systems.

Cultural competence involves three broad sets of issues: a) *pragmatic* – recognizing and addressing cultural and linguistic differences is essential to guide clinical assessment and negotiation of the goals, methods, process and progress of psychotherapy; b) *conceptual* – psychotherapy needs to mobilize changes in psychological functioning and adaptive strategies that fit the resources of individual patients and their social world; this may require rethinking the mechanisms and modalities of therapy; and c) *ethical* – psychotherapy conveys particular concepts of the person that may be at odds with the values or ways of life of particular cultures and communities. This cultural proselytization may be intentional or inadvertent for the therapist, and desired or unwanted by the patient. In either case, it may be liberatory and empowering or disruptive and undermining of individual's adaptation and social integration.

Cultural competence in psychotherapy requires knowledge, attitudes and skills for context-sensitive assessment and intervention[1]. This includes a general framework for thinking about cultural identity and difference, language and communication skills, and specific knowledge about the cultural background, lifeworld and communities of patients. Older notions of culture as constituting distinct, homogeneous and self-contained social systems have given way to more dynamic views of culture as hybrid local worlds and extended transnational networks that afford individuals multiple strands to their identities and multiple niches they can inhabit to realize their capabilities and adapt to challenges. This requires a more dynamic approach to assessing and integrating cultural dimensions into clinical case formulation and psychotherapeutic intervention[2].

Generic cultural competence makes use of tools such as the DSM-5 Outline for Cultural Formulation and the Cultural Formulation Interview, which provide a place to start in basic assessment[3]. However, psychotherapy requires more nuanced understanding of local idioms, social contexts, life predicaments and possibilities. Given the great variation within and across cultural communities, clinicians need a general attitude of humility and openness to foster a collaborative process of inquiry and mutual learning. In addition to competence and humility, the complementary construct of *cultural safety* recognizes that the effort to mobilize psychological resources to heal must consider the ongoing structures of inequity in which patient and therapist are embedded.

There is a long tradition of research on cultural variations in healing practices and psychotherapy[4]. This has identified processes of change and sharpened our understanding of how the mechanisms of psychopathology and psychotherapy are influenced by culture. Putatively universal processes in healing and psychotherapy include expectancy effects, cognitive reframing and restructuring, meta-cognitive self-regulation, and relational learning[5]. While these processes can be described in abstract terms, they take specific forms in each culture, and require appropriate models and metaphors to evoke or mobilize. Moreover, cultures may favor particular mechanisms of coping and adaptation.

Culture can then influence psychotherapy in multiple ways: a) shaping the cognitive-emotional loops that constitute or contribute to mental disorders; b) articulating shared and divergent experiences and modes of expression and communication; c) determining the ways in which modes of self-construal, coping and adaptation are enacted to yield positive effects for the sufferer and others in their social world; d) establishing norms of behavior and expression that set thresholds for pathology; and e) determining the social niches or ways of life that provide pathways for recovery[6].

The narratives of suffering and healing that are the basis for self-understanding and the medium of therapeutic transformation in psychotherapy are rooted in cultural concepts of the person. Much of current psychotherapeutic theory and practice rests on individualistic notions of the person that emphasize autonomy and independence. Alternate cultural concepts of the person may be characterized as *sociocentric*, emphasizing the embedding of the individual in interdependent social relationships; *ecocentric*, linking the person to the environment; or *cosmocentric*, recognizing relationships with ancestors or a spirit world[7]. Each of these versions of self and personhood is elaborated in indigenous or ethno-psychologies, along with specific notions of health and illness. Each provides ways of structuring the self and potential targets for intervention. The pattern theory of self suggests that embodied experience, narratives of the self, and modes of active engagement with the social world are configured in ways that may result in pathology but that also offer opportunities for transformation[8].

The metaphors of psychological theory – both its formal constructs and everyday uses to explain or rationalize behavior – draw from these cultural concepts of the person and cultural ontologies. So too do our notions of distress, dysfunction or disorder. But other metaphors also come into play. Thus, illness may be seen as a breakdown in biological machinery ("broken brain"), a dysfunction in information processing or computation, deficient learning or lack of skill, traumatic memory, weak will, or a snarl of psychological conflicts that reflect internal struggles or interpersonal tensions and contradictions. Other metaphors drawn from different ontologies may provide ways to rethink the form and content of psychotherapy to better fit cultural systems of meaning and social contexts.

Culturally adapting existing therapies can increase their acceptability, uptake and effectiveness[9]. Adaptation ranges from minor changes in language or framing of activities to more fundamental shifts in the ways problems are conceptualized and processes are mobilized to achieve a desired outcome. Recognizing cultural

difference can lead us to enlarge our models of psychotherapy to include different frameworks and practices. In recent years, we have seen this with the emergence of therapies that borrow from Buddhist or other contemplative practices. There is great potential for creative development of therapeutic methods that build on shared and distinctive features of the cultural constitution of the self. To do this, we need to continue to develop our understanding of cultural variations in the self and the poetics of illness and healing[4].

Laurence J. Kirmayer
Division of Social & Transcultural Psychiatry, McGill University, Montreal, QC, Canada

1. Kirmayer LJ. Transcult Psychiatry 2012;49:149-64.
2. Kirmayer LJ, Guzder J, Rousseau C (eds). Cultural consultation: encountering the other in mental health care. New York: Springer, 2014.
3. Aggarwal NK, Lewis-Fernández R. In: Schooler-Ocak M, Kastrup MC (eds). Intercultural psychotherapy: for immigrants, refugees, asylum seekers and ethnic minority patients. Berlin: Springer, 2020:47-57.
4. Kirmayer LJ. Healing and the invention of metaphor: toward a poetics of illness experience. Cambridge: Cambridge University Press, 2025.
5. Kohrt BA, Ottman K, Panter-Brick C et al. Clin Psychol Rev 2020;82:101920.
6. Gómez-Carrillo A, Kirmayer LJ. Front Psychiatry 2023;14:1031390.
7. Kirmayer LJ, Adeponle A, Dzokoto VAA. In: Fernando S, Moodley R (eds). Global psychologies: mental health and the global south. Berlin: Springer, 2018:21-37.
8. Daly A, Ritunnano R, Gallagher S et al. Front Psychol 2024;15:1390885.
9. Lee E, Greenblatt A, Hu R. J Cross Cult Psychol 2021;52:511-32.

DOI:10.1002/wps.21340

The role of case formulation in the current practice of psychotherapy

There is broad consensus among experts across theoretical orientations that case formulation is a core clinical skill in psychotherapy. It has been described as a "linchpin", a "map", and a "north star", among other metaphors, that guides the delivery of effective psychotherapy. Competence or training in case formulation is required by numerous accreditation, credentialing and certifying bodies.

Despite this widespread agreement, there is relatively little empirical research on case formulation in psychotherapy, including the extent to which therapists apply it in practice, nor is there consensus on how best to develop a formulation, although many models have been proposed. Additionally, multiple studies demonstrate that formulation skills are challenging to acquire, and that they may be lacking in many clinicians.

Case formulation in psychotherapy has been defined as "a process of developing a hypothesis about the causes, precipitants, and maintaining influences of a person's psychological, interpersonal and behavioral problems, in the context of the individual's culture and environment, as well as a plan to address these problems"[1]. Expert clinicians appear to agree that case formulation should include coping strategies to manage problems; factors that perpetuate problems, including cycles or patterns; protective factors, such as interpersonal support; potential organic causes; relationship and cultural considerations; and origins, including potential childhood traumas and/or more recent events. There is also agreement that a formulation should be tailored to the individual client, constructed collaboratively, empirically based, and that it should resonate with the client's experience, avoid stigmatization, and restore agency, meaning and hope[2]. The content of a case formulation will vary depending on the theory that guides its development, although theoretically integrative and transdiagnostic models of case formulation have increasingly been developed[3].

Much research on case formulation addresses reliability and validity. Results indicate greater reliability for formulation components that require less inference, such as the identification of problems, and less reliability for aspects that require more inference, particularly a core hypothesis offered to explain a client's problems. Research further suggests greater reliability when the formulation is structured and is framed in non-technical language[4].

Research comparing outcomes between formulation-driven and manual-based therapy shows either no difference or a slight advantage for the former. However, several studies have demonstrated that competence in case formulation appears to predict improved outcome and process, including symptom reduction, lower dropout rates, sudden gains among individuals with treatment-resistant depression, and more emotional processing[5-7]. The limited research on training in psychotherapy case formulation demonstrates improvement in the skill, even when the training duration is very short. Several methods are available to evaluate formulation competence, which vary in how formulation quality is defined and measured, and the settings for which they have been designed and assessed[8].

Relatively little is known about how practitioners employ case formulation in routine clinical practice. A survey of cognitive-behavior therapists found a high level of agreement on the importance of identifying problems and maintaining factors, developing explanatory hypotheses, and identifying goals, but lower ratings on the importance of consulting theory or external empirical evidence to develop the formulation[9]. Therapists also gave lower ratings to the value of self-report measures to identify problems, to the use of structured case formulation templates, and to the importance of evaluating hypotheses during therapy. However, using a structured case formulation template, evaluating hypotheses during therapy, and consulting theory related to the presenting problems were rated as significantly more important by more experienced clinicians compared to less experienced ones. These findings are consistent with prior research showing that clinicians may rely more on personal experiences, existing training, and consultation with col-

leagues when conceptualizing and treating clients than on staying abreast of and drawing upon theory and empirical research.

At least four trends in case formulation research and scholarship can be identified. One is greater attention to the importance of incorporating cultural considerations. Culturally responsive formulations may help address treatment disparities and reduce the stigma of psychotherapy in some communities. Evidence suggests that culturally adapted psychotherapy reduces dropout rates and improves outcome.

A second trend is the application of new and sophisticated technologies to explore the impact of formulation-based interventions on psychotherapy processes and outcomes. These methods are likely to advance case formulation research and the understanding of effective psychotherapy processes. Recent major advances in artificial intelligence (AI) will inevitably have an impact on the case formulation process. Efforts are already underway to apply AI tools to evaluate case formulations as well as to construct them based on interview data.

A third trend appears to be the development of team-based formulations, which have been employed in outpatient settings to improve treatment of complex cases and challenging behavior problems. If implemented systematically, rigorously and based on evidence, these formulations hold the promise of improving reliability and validity, as well as outcomes when treating complex problems.

A fourth trend is the proliferation of systematic, case-formulation-guided case studies and the establishment of journals that publish them. The publication of detailed case studies of evidence-based, formulation-guided psychotherapy has tremendous teaching potential and may create a database of cases that clinicians can consult for treatment ideas, and the exploration of which may contribute to better understanding of effective psychotherapy processes and improving outcomes.

It is hoped that the discrepancy between the consensus view that case formulation is a core clinical competence in psychotherapy and the relative lack of research on this topic, including on training practitioners to a point of competency, will be remedied. The advent of new tools, the expansion of opportunities for publishing evidence-based, formulation-guided case studies, and recent research showing that case formulation competence predicts improved processes and outcomes, bodes well for this prospect.

Tracy D. Eells
Department of Psychiatry and Behavioral Sciences, University of Louisville, Louisville, KY, USA

1. Eells TD. Psychotherapy case formulation, 2nd ed. Washington: American Psychological Association, in press.
2. Thrower NE, Bucci S, Morris L et al. Br J Clin Psychol 2024;63:213-26.
3. Eells TD (ed). Handbook of psychotherapy case formulation, 3rd ed. New York: Guilford, 2022.
4. Easden MH, Kazantzis N. J Clin Psychol 2018;74:356-84.
5. Abel A, Hayes AM, Henley W et al. J Consult Clin Psychol 2016;84:726-37.
6. Easden MH, Fletcher RB. Psychother Res 2020;30:151-69.
7. Gates V, Hsiao M, Zieve GG et al. Behav Res Ther 2021;142:103874.
8. Bucci S, French L, Berry K. J Clin Psychol 2016;72:517-33.
9. Huisman P, Kangas M. Behav Change 2018;35:1-21.

DOI:10.1002/wps.21341

Advances in personalization of psychological interventions

Tailoring treatment to a patient's unique circumstances and needs is a central tenet of mental health care. In routine practice, this tailoring process is most often based on clinical theory and intuition, with known biases such as confirmation bias (i.e., the tendency to interpret information in a way that supports one's prior beliefs) and the use of heuristics (such as representativeness heuristic, i.e. estimating the probability of an event by comparing it to a stereotype that one already has in mind), which limits its reliability.

A new research paradigm has emerged in the last decade, aiming to achieve evidence-based personalization and data-informed decision making[1]. The core idea is that it is possible to predict a patient's treatment outcome by observing how other similar patients have responded to this treatment in the past. This approach, therefore, involves collecting comprehensive data from patients at the beginning and throughout treatment to support personalized treatment selection and continuous adaptation.

Before treatment begins, multidomain patient data can be used to inform treatment selection. Using algorithms from the field of machine learning, which can analyze complex and non-linear relationships in big data, it is possible to generate predictions about a patient's probable treatment outcome. If outcome predictions are available for different treatment approaches or strategies, the most promising approach for an individual patient can be recommended. Data sources that are commonly used for this endeavor include sociodemographic and psychopathological variables, personality traits, and sometimes digital phenotypes[2]. In recent years, prediction algorithms have also focused on matching patients to therapists, and selecting specific treatment strategies and target processes in the context of modularized and personalized therapies[1].

Several model-development studies have emerged in the field of psychotherapy. These studies typically develop clinical prediction algorithms using cross-validation methods which involve partitioning clinical datasets into training, validation and test samples to develop, fine-tune and evaluate the generalizability of these models[2]. A recent meta-analysis of clinical trials that apply clinical prediction models indicates that algorithm-driven personalized psychological interventions are more effective than standard psychological treatments[3].

However, despite such promising results, the practical implementation and generalizability of such prediction algorithms is still a matter of investigation[3]. Only a few studies have attempted to validate treatment selection algorithms in statistically independent external datasets, and these attempts have yet to be convinc-

ing[4]. Important challenges include understanding the necessary sample sizes and data sources (i.e., number and types of variables) required to optimize prediction accuracy, and to identify and correct potential biases. Therefore, despite the advanced statistical approaches applying corrections against overfitting (such as shrinkage and cross-validation), the generalizability to new data and samples as well as real-life implementation require further refinement and testing.

To demonstrate a prediction algorithm's external validity, prospective studies must be conducted in which trained prediction models are applied to new incoming patients, which are then treated with the recommended strategy to test whether data- and algorithm-informed treatment selection can actually improve outcomes. Only a few of these prospective studies have been conducted to date, with promising results indicating that data-informed treatment selection improves clinical outcomes by comparison to usual psychological care[1,5].

Furthermore, a shift towards personalization of psychological interventions requires trial designs that allow for a better understanding of variance at the individual level, for example pragmatic or adaptive trials. Pragmatic trials allow a better understanding of the real-world effects of data-informed psychological therapy by assessing the effectiveness of personalized systems within naturalistic, heterogeneous routine care settings. Adaptive trials (e.g., sequential multiple assignment randomized trials, SMART) use multiple treatment groups and interim outcome assessments, which are utilized as triggers to start or terminate specific subtrials. Such subtrials are integrated in the adaptive trial framework to simultaneously evaluate different interventions and reduce required sample sizes in multi-intervention studies[6].

Besides the need for suitable designs to test treatment selection capabilities, new data layers and assessment strategies at the beginning of as well as during treatment have also recently been applied to further refine and improve the personalization of psychological therapies. For example, routine outcome monitoring in combination with feedback has been a key component of evidence-based treatment personalization, particularly for patients at risk of adverse outcomes. Algorithm-based visual feedback is provided to therapists (and patients) to guide clinical decision-making during treatment[6]. This is especially important as long as predictive treatment variables and recommendations are limited.

Strong empirical support from multiple meta-analyses confirms outcome monitoring and feedback to be an effective, resource-efficient enhancement to psychological therapy[7]. Such personalization during treatment allows therapists to identify patient-specific trends, deviations, or stagnation in patient progress. If significant deviations occur, alerts are generated that signal a potential risk of treatment failure. This enables early adaptations, especially when combined with personalized clinical recommendations that guide therapists in overcoming specific challenges.

Most recently, to advance the personalization and precision of psychological therapies, treatment selection and monitoring concepts have been combined in comprehensive decision-support systems[1]. Such systems include treatment strategy recommendations at the beginning of the intervention, and treatment progress recommendations to facilitate adaptive clinical decisions throughout treatment. A recent randomized clinical trial that prospectively evaluated such a system in a large outpatient sample found that outcomes improved when therapists followed the recommended treatment strategy, and that therapist-rated usefulness of the recommendations moderated the decision support system's effect. The combination of accurate treatment selection and monitoring led to improved outcomes, particularly for therapists who found it useful, suggesting that the use of such systems requires adequate therapist training[6].

Despite these positive developments, there is a need to further improve measurement and to address implementation and generalizability issues to make treatment personalization more feasible in clinical practice. For example, integrating novel assessment methods such as ecological momentary assessments (including behavioral markers such as physical activity levels, sleep patterns, or physiological signals) and video or audio recordings, and evaluating these novel data layers with time series analysis, emotion detection algorithms, and large language models (LLMs) holds considerable potential to enhance measurement precision[1]. For example, speech-to-text models enable the automatic transcription of session recordings, providing a rich data source for analysis. LLMs, with their advanced natural language processing capabilities, can process these transcripts at scale and extract linguistic and/or psychological markers to predict outcome or dropout[8].

The key advantages of these technologies include the reduction of retrospective biases, the identification of micro patterns and subtle therapy processes, and the integration of multimodal data sources. For example, rather than receiving generic intervention recommendations, therapists could receive personalized insights through LLMs that highlight the most relevant treatment adjustments based on the patient's expressed needs. However, the validity and reliability of these new measures and the best practices to implement them need further investigation, including the trustworthiness and interpretability of outputs[9].

Overall, we can be optimistic about the opportunities and benefits associated with advances in the personalization of psychological interventions, especially given the new technological possibilities. There is great potential in developing this paradigm through the analysis of big data from electronic health records, especially in heterogeneous populations that are typical of routine care. Infrastructure must be established for research centers to implement and test standard and novel assessment tools and personalization options with trained therapists in diverse samples. However, further developments require suitable designs and larger and more diverse databases.

Wolfgang Lutz[1], Brian Schwartz[1], Antonia Vehlen[1], Steffen T. Eberhardt[1], Jaime Delgadillo[2]
[1]Department of Psychology, Trier University, Trier, Germany; [2]Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

1. Lutz W, Schwartz B, Delgadillo J. Annu Rev Clin Psychol 2022;18:71-98.
2. Cohen ZD, Delgadillo J, DeRubeis R. In: Barkham M, Lutz W, Castonguay LG (eds). Bergin and Garfield's handbook of psychotherapy and behavior change. Hoboken: Wiley, 2021:667-700.
3. Nye A, Delgadillo J, Barkham M. J Consult Clin Psychol 2023;91:389-97.

4. Chekroud AM, Hawrilenko M, Loho H et al. Science 2024;383:164-7.
5. Delgadillo J, Ali S, Fleck K et al. JAMA Psychiatry 2022;79:101-8.
6. Lutz W, Vehlen A, Schwartz B. J Consult Clin Psychol 2024;92:671-3.
7. Barkham M, de Jong K, Delgadillo J et al. Psychother Res 2023;33:841-55.
8. Eberhardt ST, Schaffrath J, Moggia D et al. Psychother Res 2025;35:174-89.
9. Deisenhofer AK, Barkham M, Beierl ET et al. Behav Res Ther 2024;172:104443.

DOI:10.1002/wps.21342

Rethinking the therapeutic alliance in digital mental health interventions

Over the past two decades, research on digital mental health interventions has grown rapidly. They hold promise for addressing existing gaps in mental health care, with numerous studies demonstrating their effectiveness across various psychiatric and somatic conditions[1]. Their formats vary, ranging from video-based therapy to text-based approaches delivered via secure email or chat. However, most research has focused on either fully automated, self-guided, or therapist-guided smartphone- and web-based applications, typically providing structured cognitive behavioral therapy content through self-help modules. More recently, conversational agents (i.e., chatbots) have gained attention as innovative tools for providing therapeutic support.

Despite the growing evidence of the effectiveness of digital mental health interventions, several challenges remain. A significant problem with self-guided interventions is low user engagement and high dropout rates, often leading to poor treatment outcomes. In contrast, therapist-guided interventions typically show higher adherence and improved effectiveness[1,2]. Reflecting this, the role of therapist support and the therapeutic alliance in these interventions have emerged as central themes, and indeed as top research priorities[3].

A substantial body of research has investigated the therapeutic alliance in digital mental health interventions, frequently measuring it through self-report questionnaires, most commonly the Working Alliance Inventory (WAI). This tool assesses the alliance through three dimensions: agreement on therapeutic goals, agreement on tasks to reach these goals, and the emotional bond. The WAI is often adapted to reflect the digital context. For instance, in guided interventions, the emotional bond can refer to the patient's relationship with the therapist, while agreement on goals and tasks primarily relates to the digital self-help program[4].

Evidence suggests that patients can form therapeutic alliances in digital environments – regardless of the communication modality (e.g., videoconferencing or text-based), diagnosis, or even minimal-contact formats such as guided self-help interventions – with alliance ratings often comparable to those observed in traditional face-to-face therapies[5]. Furthermore, although findings vary across studies, meta-analyses show that the strength of the therapeutic alliance in digital contexts positively correlates with outcomes, with effect sizes that are likely somewhat smaller, but still quite similar, to those observed in face-to-face settings[6,7].

Beyond therapist-patient relationships, research increasingly recognizes that even the emotional bond can extend to self-help apps and, more recently, conversational agents. This insight is not entirely new. As early as 1966, it was observed that people interacting with the chatbot ELIZA, a simple conversational agent mimicking Rogerian psychotherapy, attributed human-like empathy and understanding to the program – a phenomenon now known as the "ELIZA effect". This effect highlights the users' tendency to connect emotionally with digital entities, even when these interactions lack genuine human qualities. Parallel findings have emerged in research on bibliotherapy, where relational elements such as empathically addressing emotional distress, collaboratively setting goals, and normalizing setbacks have been suggested to strengthen the readers' sense of alliance with therapeutic materials[8].

The comparability of alliance ratings with traditional therapies has helped legitimize digital mental health interventions. However, caution is warranted when directly transferring alliance concepts from face-to-face to digital settings. As the following examples illustrate, similar alliance scores across treatment formats do not necessarily indicate equivalent experiences or underlying processes.

Reflecting on the therapeutic bond in guided self-help programs, some users commented that the bond with the therapist was surprisingly strong, given that it was a digital intervention with minimal contact[5]. Such statements underline that expectations shape therapeutic experiences, and highlight the importance of context when interpreting self-report ratings. Just as a five-star rating for a budget hotel has a different meaning to one for a luxury hotel, similar alliance ratings across treatment formats should be interpreted with caution.

The same caution applies to ratings of goal and task agreement. While similar scores suggest comparability, they may result from fundamentally different processes. Imagine a mental health system in which, instead of broadly trained psychotherapists who collaboratively define goals and adjust treatment over time, patients choose from a list of specialists offering pre-defined interventions for specific disorders. This model – typical of many digital mental health interventions – relies on matching diagnosis to the intervention from the outset. In such cases, high levels of agreement may indicate an initial fit rather than a co-created therapeutic process. While this structured approach can effectively align treatment with disorder-specific goals, its lack of flexibility may contribute to the engagement challenges seen in digital interventions when they fail to address a patient's broader or evolving needs.

Furthermore, the rise of chatbots and their increasingly sophisticated "empathic" responses raise questions about forming meaningful relationships with digital tools. Chatbots can convincingly simulate empathy and respond appropriately to users' emotional states. However, they cannot share another person's feelings and

genuinely care about their well-being. As one of our students once put it, engaging with a chatbot is like hearing a warm, reassuring voice in an empty room. The words may be comforting, the tone just right – but in the end, no one is truly there. Human empathy requires real presence and investment (in terms of time, effort, and emotional energy), which are factors that also encourage mutual engagement. The lack of such investment in automated treatments may partly explain persistent challenges with user engagement.

Overall, the therapeutic alliance in digital mental health interventions tends to be less rich and reciprocal than in traditional therapies, also because certain digital communication formats – such as text-based or asynchronous communication – inherently lack nonverbal and paraverbal cues needed for nuanced emotional attunement. However, digital formats also offer distinct advantages. For some therapeutic tasks, such as psychoeducation, face-to-face interactions may be overly complex and overwhelming. In these cases, the simplified nature of digital communication can reduce relational demands, increase focus, and promote more effective learning. Similarly, digital communication can facilitate quicker openness when discussing sensitive issues. This aligns with the intimacy equilibrium model[9], which proposes that physical distance and fewer nonverbal cues can encourage verbal openness, unlike situations with close physical proximity to strangers – like, for example, in an elevator – where people tend to avoid intimate conversations.

Our current understanding of the therapeutic alliance in digital mental health interventions is rooted in models developed for face-to-face therapies. Applying these models to digital interventions may obscure more than it reveals. Rather than asking whether the "digital alliance" is "as good as" the traditional one, it may be more fruitful to ask how it differs. Digital interventions may even help to broaden our understanding of what constitutes the therapeutic alliance. Is it the subjective sense of being understood that matters most, or the actual presence of an empathic mind? If a chatbot – despite lacking genuine empathy – can still promote symptom relief, this also raises questions about the underlying mechanisms through which the alliance facilitates therapeutic change. For instance, can a chatbot reduce a patient's emotional dysregulation as effectively as a human therapist?

Furthermore, not all clients value a therapeutic relationship. Some prefer more autonomy and only minimal interpersonal interaction. This highlights the need to identify moderators of the alliance-outcome relationship in digital mental health interventions. Are there patient characteristics or intervention types for which a strong alliance or therapist presence is more important?

Ultimately, digital mental health interventions encourage us to rethink not only how we deliver therapy, but also what fundamentally makes therapy effective and for whom, providing unique insights into the nature of therapeutic alliances and the nature of healing itself.

Thomas Berger
University of Bern, Bern, Switzerland

1. Hedman-Lagerlöf E, Carlbring P, Svärdman F et al. World Psychiatry 2023;22:305-14.
2. Krieger T, Bur OT, Weber L et al. Internet Interv 2023;32:100617.
3. Hollis C, Sampson S, Simons L et al. Lancet Psychiatry 2018;5:845-54.
4. Gómez Penedo JM, Berger T, Grosse Holtforth M et al. J Clin Psychol 2019;76:973-86.
5. Berger T. Psychother Res 2017;27:511-24.
6. Aafjes-van Doorn K, Spina DS, Horne SJ et al. Clin Psychol Rev 2024;110:102430.
7. Probst GH, Berger T, Flückiger C. Verhaltenstherapie 2019;32:135-46.
8. Richardson R, Richards DA, Barkham M. Behav Cogn Psychother 2010;38:67-81.
9. Argyle M, Dean J. Sociometry 1965;28:289-304.

DOI:10.1002/wps.21343

Attention-deficit/hyperactivity disorder (ADHD) in adults: evidence base, uncertainties and controversies

Samuele Cortese[1-5], Mark A. Bellgrove[6], Isabell Brikell[7], Barbara Franke[8], David W. Goodman[9], Catharina A. Hartman[10], Henrik Larsson[11], Frances R. Levin[12], Edoardo G. Ostinelli[13], Valeria Parlatini[2,3,14], Josep A. Ramos-Quiroga[15], Margaret H. Sibley[16], Anneka Tomlinson[13], Timothy E. Wilens[17], Ian C.K. Wong[18], Nina Hovén[19], Jeremy Didier[20], Christoph U. Correll[21-25], Luis A. Rohde[26], Stephen V. Faraone[27]

[1]Developmental EPI (Evidence synthesis, Prediction, Implementation) Lab, Centre for Innovation in Mental Health, School of Psychology, Faculty of Environmental and Life Sciences, University of Southampton, Southampton, UK; [2]Clinical and Experimental Sciences (CNS and Psychiatry), Faculty of Medicine, University of Southampton, Southampton, UK; [3]Hampshire and Isle of Wight Healthcare Foundation NHS Trust, Southampton, UK; [4]Hassenfeld Children's Hospital at NYU Langone, New York University Child Study Center, New York, NY, USA; [5]Department of Precision and Regenerative Medicine and Ionian Area, University of Bari "Aldo Moro", Bari, Italy; [6]School of Psychology and Psychiatry, Monash University, Clayton, VIC, Australia; [7]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; [8]Department of Cognitive Neuroscience, Radboud University Medical Center, Nijmegen, The Netherlands; [9]Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Baltimore, MD, USA; [10]Interdisciplinary Center of Psychopathology and Emotion Regulation, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; [11]School of Medical Sciences, Örebro University, Örebro, Sweden; [12]Division on Substance Use Disorders, New York State Psychiatric Institute, New York, NY, USA; [13]Oxford Precision Psychiatry Lab, NIHR Oxford Health Biomedical Research Centre, Oxford, UK; [14]School of Psychology, Faculty of Environmental and Life Sciences, University of Southampton, Southampton, UK; [15]Centro de Investigación Biomédica en Red de Salud Mental, Madrid, Spain; [16]University of Washington School of Medicine and Seattle Children's Research Institute, Seattle, WA, USA; [17]Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; [18]Centre for Safe Medication Practice and Research, Department of Pharmacology and Pharmacy, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong SAR, China; [19]ADHD Europe, Brussels, Belgium; [20]Children & Adults with ADHD (CHADD), Lanham, MD, USA; [21]Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin Berlin, Berlin, Germany; [22]Department of Psychiatry, Northwell Health, Zucker Hillside Hospital, Glen Oaks, NY, USA; [23]Department of Psychiatry and Molecular Medicine, Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, USA; [24]Center for Psychiatric Neuroscience, Feinstein Institutes for Medical Research, Manhasset, NY, USA; [25]DZPG, German Center for Mental Health, partner site Berlin, Berlin, Germany; [26]Norton College of Medicine at SUNY Upstate Medical University, Syracuse, NY, USA; [27]State University of New York Upstate Medical University, Syracuse, NY, USA

Attention-deficit/hyperactivity disorder (ADHD) was once thought to be solely a childhood condition. Now it is well established that it can persist into adulthood, with an estimated worldwide prevalence of around 2.5%. Additionally, up to 70% of individuals with childhood-onset ADHD continue to experience impairing symptoms as adults, even if they no longer meet the criteria for a formal diagnosis. The validity of adult ADHD initially faced strong criticism. Today, empirical research supports its descriptive validity (identifying characteristic signs and symptoms), predictive validity (concerning specific outcomes, courses, and responses to treatment), and concurrent validity (evidence related to its underlying causes and biological mechanisms). Despite this progress, unresolved questions and ongoing debates about adult ADHD persist. This paper summarizes current empirical evidence, alongside uncertainties and controversies, regarding the definition, epidemiology, diagnosis, etiology, neurobiology, and management of ADHD in adults. Crucially, we also include perspectives from individuals with lived experience of this condition, highlighting their views on unmet needs and priorities for improving care. Key uncertainties and controversies on adult ADHD include: a) the possibility of late-onset ADHD; b) the significance of emotional dysregulation as a core symptom; c) the definition and characterization of functional impairment; d) the persistence of comorbid psychiatric and somatic conditions after accounting for confounders; e) the relevance of executive dysfunction in the definition of the condition; f) the use of objective diagnostic measures; g) the long-term effects of treatments; and h) the role of non-pharmacological interventions. Further research on adult ADHD is urgently needed. Funding for studies on this condition lags behind that for childhood ADHD and other mental disorders in adulthood. Hopefully, efforts by clinicians, researchers and other stakeholders will ultimately help ensure that adults with ADHD are better understood, supported, and empowered to thrive.

Key words: Adult attention-deficit/hyperactivity disorder, descriptive validity, predictive validity, controversies, diagnosis, epidemiology, etiology, management, people with lived experience

(World Psychiatry 2025;24:347–371)

Attention-deficit/hyperactivity disorder (ADHD) is conceptualized as a neurodevelopmental disorder marked by developmentally inappropriate, pervasive and impairing inattention and/or hyperactivity-impulsivity[1-3]. It was initially considered a childhood disorder, with the nosological developments and clinical awareness of ADHD in adults lagging behind that of the childhood condition[3,4].

Although it is claimed that a very early depiction of a condition resembling what is nowadays referred to as ADHD can be found in the Greek texts of the philosopher Theophrastus in the 4th century BC[5], the first descriptions of children who would likely receive today this diagnosis appeared in French, German and Scottish texts in the 18th century[6]. The first report in a scientific journal was published in 1902, when the British paediatrician G. Still described 43 cases of children and adolescents who would qualify for an ADHD diagnosis today[7]. Many of these individuals were reported to struggle with sustained attention, and most were overactive[7,8].

The condition was initially referred to as "minimal brain damage", assuming that it was associated with brain lesions, which evolved into "minimal brain dysfunction", acknowledging that excessive levels of physical activity and inattention might not necessarily be associated with structural brain lesions[6]. It entered the official nosology in the DSM-II as "hyperkinetic reaction of childhood", followed by a series of nosological reconceptualizations, first as "attention-deficit disorder (with or without hyperactivity)" in the DSM-III, and later as "attention-deficit/hyperactivity disorder" in subsequent DSM editions, up to the current DSM-5-TR. The condition first appeared in the ICD-9 as "hyperkinetic syndrome of childhood", later renamed "hyperkinetic disorder" in the ICD-10, and "attention deficit hyperactivity disorder" in the ICD-11.

While all the above descriptions and diagnostic labels referred

to children and adolescents, the idea that ADHD could also affect adults is not new. Scattered reports from the 1950s and 1960s[9-11] described the symptomatic evolution of "minimal brain dysfunction" in adulthood, and a more systematic article published in 1976[12] sought to define this condition in a group of 15 adults. This definition was based on the presence of impulsivity, irritability, inattentiveness, restlessness, and emotional lability, in the absence of schizophrenia, primary affective disorder, organic brain syndrome, or intellectual disability. A history of long-standing impulsiveness, inattentiveness, restlessness, short temper, and emotional lability – based on self-report and information provided by third parties (e.g., parents) – was also required for the diagnosis of the condition in adults.

Building on this work, the Wender Utah criteria[13] were proposed in the 1990s for what we would today call ADHD in adults. These criteria required a retrospective childhood diagnosis of minimal brain dysfunction, ongoing difficulties with inattentiveness and hyperactivity, and at least two of the following five symptoms: mood lability, irritability and hot temper, impaired stress tolerance, disorganization, and impulsivity. However, the Wender Utah criteria have progressively diverged from conceptualizations of ADHD in the DSM and ICD, and have been criticized for their limited scope, excluding individuals with predominantly inattentive ADHD and those with coexisting mood disorders, as well as for conflating ADHD with conditions such as oppositional defiant disorder and bipolar disorder[14].

The importance of ADHD in adults has gained progressively more traction in the various editions of the DSM, as new data from longitudinal studies of youth and clinical studies of adults have placed the validity of the condition on a firm footing. The DSM-III introduced the category "attention deficit disorder, residual type" for adults diagnosed in childhood who continued to exhibit clinically significant levels of symptoms and impairment. The DSM-III-R recognized that ADHD could persist into adulthood in up to 30% of cases. Despite presenting a unique set of criteria across ages, the DSM-IV provided examples of how ADHD symptoms change in expression during adulthood. The DSM-IV also warned against relying solely on self-report for diagnosis and emphasized the importance of collateral information.

Based on field trials conducted specifically with adults, the DSM-5 changed the threshold of symptoms required for the diagnosis. Starting at age 17, five (rather than six) symptoms of inattention and/or hyperactivity-impulsivity were required. This change aligned with evidence showing that mandating at least six hyperactive-impulsive symptoms excludes a significant percentage (almost half) of adults who are at least 1.5 standard deviations above the population mean on a dimensional measure of hyperactivity-impulsivity[15].

Despite the progressive characterization of adult ADHD, its validity initially met with strong criticism. Beyond general arguments, such as the claim that the prevalence of ADHD has grown rapidly in some countries due to it representing an "expanding and lucrative market" for stimulants and related medications[16], specific concerns revolved around four main aspects[17]: a) the reliability of recalling childhood symptoms; b) the possibility that ADHD symptoms are accounted for by other disorders; c) uncertainty about the effectiveness of medications in adults; d) the fact that ADHD in adults can be a self-diagnosed condition and that some individuals may fake symptoms to obtain stimulants for misuse or diversion.

However, accumulating evidence has shown that: a) symptoms self-endorsed by adults with ADHD correlate with both parental rating scale scores and responses to methylphenidate[18,19]; b) there are individuals who exhibit ADHD symptoms without coexisting conditions that might otherwise explain the presentation[20]; c) medications with a proven efficacy in youth with ADHD are efficacious in clinical trials[21] and effective in real-world settings[22] for ADHD in adults; d) although fake diagnoses and stimulant misuse do occur, self-diagnosis may reflect gaps in the mental health system's ability to recognize legitimate conditions[14].

Furthermore, empirical research supports the descriptive validity (documentation of characteristic signs and symptoms), predictive validity (regarding specific courses, outcomes, and treatment responses), and concurrent validity (concerning etiology and pathophysiology) of adult ADHD[14].

However, while research has clarified many aspects, unresolved issues and controversies surrounding adult ADHD continue to emerge. In this paper, we summarize the empirical evidence, as well as the uncertainties and controversies, related to the definition, epidemiology, diagnosis, etiology, neurobiology, and management of ADHD in adults. Importantly, we also present the views of representatives from associations of individuals with lived experience, focusing on their perceptions of key unmet needs and priorities for adults with ADHD.

CLINICAL MANIFESTATIONS

Inattention, hyperactivity and impulsivity are currently considered the core symptoms of ADHD. Although most health care decisions – for example, those about who should be referred to treatment – involve categorizations (i.e., yes or no), ADHD symptoms, as those of most mental disorders and several somatic diseases, lie on a continuum.

Table 1 presents a list of the main ADHD symptoms across the lifespan. Manifestations of inattention are numerous, including mind wandering while performing a task, lack of persistence in low-motivating activities, forgetfulness, distraction by irrelevant stimuli, and disorganization. Hyperactivity manifests as excessive, inappropriate activity; fidgeting, tapping, restlessness or talkativeness. Impulsive symptoms include making decisions or actions without thinking or considering consequences, difficulty waiting turns, and social intrusiveness. In order to diagnose ADHD, according to current diagnostic systems, these symptoms must have specific characteristics in terms of age of onset, duration and pervasiveness, as detailed in Tables 2 and 3 and discussed in the section below on diagnosis.

Core ADHD symptoms may manifest differently in adults. Thus, hyperactivity in adults often manifests as inner restlessness, overscheduling, or not being able to relax properly. Impulsive behavior may manifest as acting without thinking or blurting things out,

Table 1 List of symptoms of attention-deficit/hyperactivity disorder (ADHD) across the lifespan

Inattention symptoms

- Often fails to give close attention to details or makes careless mistakes in schoolwork, at work, or with other activities.
- Often has trouble holding attention on tasks or play activities.
- Often does not seem to listen when spoken to directly.
- Often does not follow through on instructions and fails to finish schoolwork, chores, or duties in the workplace (e.g., loses focus, side-tracked).
- Often has trouble organizing tasks and activities.
- Often avoids, dislikes, or is reluctant to do tasks that require mental effort over a long period of time (such as schoolwork or homework).
- Often loses things necessary for tasks and activities (e.g., school materials, pencils, books, tools, wallets, keys, paperwork, eyeglasses, mobile telephones).
- Is often easily distracted.
- Is often forgetful in daily activities.

Hyperactivity/impulsivity symptoms

- Often fidgets with or taps hands or feet, or squirms in seat.
- Often leaves seat in situations when remaining seated is expected.
- Often runs about or climbs in situations where it is not appropriate (adolescents or adults may be limited to feeling restless).
- Often unable to play or take part in leisure activities quietly.
- Is often "on the go" acting as if "driven by a motor".
- Often talks excessively.
- Often blurts out an answer before a question has been completed.
- Often has trouble waiting his/her turn.
- Often interrupts or intrudes on others (e.g., butts into conversations or games).

spending too much money or spending it too quickly, carrying out plans immediately, resigning from jobs in a flurry, starting relationships quickly, and not being able to postpone need gratification. Older adolescents and adults frequently report "sensation-seeking", "novelty-seeking", or seeking out excitement[23].

Most investigations assessing the factor structure of ADHD symptoms in various cultures have relied on samples of children and adolescents, using different information sources (e.g., teachers and parents). These findings have suggested a two-factor model for ADHD, with inattention and hyperactivity/impulsivity as two independent but correlated dimensions[24]. Some investigations using more sophisticated mathematical modeling have suggested that a bifactor model, with one general factor and three specific factors (inattention, hyperactivity, impulsivity), might fit the data better[25]. Studies with adult samples have supported this finding, suggesting separate factors for inattentive, hyperactive and impulsive symptoms[26].

Table 2 Diagnostic criteria for attention deficit hyperactivity disorder (ADHD) according to the ICD-11

- A persistent pattern (e.g., over at least 6 months) of inattention symptoms and/or a combination of hyperactivity and impulsivity symptoms that is outside the limits of normal variation expected for age and level of intellectual development. Symptoms vary according to chronological age and disorder severity.
- Evidence of significant inattention and/or hyperactivity-impulsivity symptoms prior to age 12, though some individuals may first come to clinical attention later in adolescence or as adults, often when demands exceed the individual's capacity to compensate for limitations.
- Manifestations of inattention and/or hyperactivity-impulsivity must be evident across multiple situations or settings (e.g., home, school, work, with friends or relatives), but are likely to vary according to the structure and demands of the setting.
- Symptoms are not better accounted for by another mental disorder (e.g., an anxiety or fear-related disorder, a neurocognitive disorder such as delirium).
- Symptoms are not due to the effects of a substance (e.g., cocaine) or medication (e.g., bronchodilators, thyroid replacement medication) on the central nervous system, including withdrawal effects, and are not due to a disease of the nervous system.

Table 3 Diagnostic criteria for attention-deficit/hyperactivity disorder (ADHD) according to the DSM-5-TR

A. A persistent (at least 6 months) pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development.
B. Several inattentive or hyperactive-impulsive symptoms were present prior to age 12 years.
C. Several inattentive or hyperactive-impulsive symptoms are present in two or more settings (e.g., at home, school, or work; with friends or relatives; in other activities).
D. There is clear evidence that the symptoms interfere with, or reduce the quality of, social, academic or occupational functioning.
E. The symptoms do not occur exclusively during the course of schizophrenia or another psychotic disorder and are not better explained by another mental disorder (e.g., mood disorder, anxiety disorder, dissociative disorder, personality disorder, substance intoxication or withdrawal).

Based on the preponderance of clinical manifestations, major classification systems (DSM and ICD) propose three different clinical presentations for ADHD: *combined presentation*, when both hyperactive-impulsive and inattentive symptoms are clinically significant aspects of the current clinical picture; *predominantly inattentive presentation*, with a preponderance of inattentive symptoms; and *hyperactive/impulsive presentation*, with a predominance of hyperactivity-impulsivity symptoms. In the past, these clinical presentations were known as ADHD types or subtypes. However, for a variety of reasons, mainly the lack of developmental stability, this terminology has been abandoned[27]. The ICD-11 refers to them as "specifiers", and the DSM-5-TR as "presentations". The developmental stage of the individual must be considered when characterizing the ADHD clinical presentation. Since hyperactive/impulsive symptoms decrease more significantly than inattention in clinical and population samples[28], the most frequent presentation found in older adolescents and adults is ADHD with predominantly inattentive symptoms.

Motivation, relevance, and attractiveness of the task influence symptomatic manifestations. Individuals with ADHD may be able to remain focused when performing specific tasks such as playing videogames. Thus, overconcentration or "hyperfocus" in highly motivating situations are frequent in individuals with this condition. ADHD can thus be viewed as an attention *dysregulation* (rather than *deficit*) disorder. Many patients with ADHD can concentrate in some contexts, but they cannot deploy concentration at some ordinary moments in which it is needed.

Regarding effects of context, careful parents may provide structured environments and stimulation for their children with ADHD, creating a situation in which symptoms only manifest later in adolescence and young adulthood, when more autonomy is needed.

In addition, culture shapes the expression and expectations of behaviors and symptoms[29]. Thus, assessment of ADHD symptoms must always consider cultural aspects.

In addition to what are currently conceptualized as the core symptoms of the disorder, individuals with ADHD may often present with additional problems/dysfunctions. Problems with peers, already evident in childhood, tend to become more pronounced during adolescence and young adulthood, contributing to social rejection and fewer friendships.

Substance experimentation during adolescence is also common, and individuals with ADHD are at a significantly higher risk for substance use disorders, as well as being more vulnerable to engage in other risky behaviors such as unprotected sex, with consequent higher rates of sexually transmitted diseases and pregnancies in young adulthood[3].

Despite a large amount of research to date, there are several uncertainties and controversies regarding the clinical manifestations of adult ADHD.

First, should emotional dysregulation be part of the diagnostic construct of ADHD in adults? Although this is a transdiagnostic manifestation (i.e., with low specificity for ADHD), it is extremely frequent in adults with the disorder. Adolescents and young adults with ADHD often exhibit excessive negative and positive responses, along with outbursts of anger and irritability[3]. Indeed, emotional dysregulation was a core part of the first descriptions of adult ADHD[12], and up to 70% of adults with the disorder implement more frequently non-adaptive emotion regulation strategies compared to people without ADHD symptoms. Additionally, emotion dysregulation is clearly associated with both symptom severity and executive functioning[30].

However, emotional dysregulation is currently not part of the core symptoms of ADHD, and there is uncertainty on whether this manifestation in adults with ADHD is not simply the heir of oppositional defiant disorder, an extremely prevalent comorbid diagnosis in children with ADHD often not recognized by adult psychiatrists[31]. Additionally, emotional dysregulation is associated with controversies in the differential diagnosis between ADHD and bipolar disorder. Momentaneous outbursts of anger and irritability followed by quick return to baseline are regarded as characteristic of adults with ADHD by some clinicians, while others see this symptomatic presentation as an expression of the bipolar phenotype[32].

Second, should executive functioning deficits be more represented in diagnostic criteria for ADHD in adults? The key role of this dysfunction in adult ADHD was demonstrated by Kessler et al[33], who documented that executive deficit manifestations – such as difficulty prioritizing work or completing tasks in allotted time, and making careless mistakes – are the most important predictors of a diagnosis of adult ADHD according to the DSM-IV. However, the presence of executive dysfunction is not currently required for the diagnosis of ADHD, and some consider it as a possible associated dimension that is important to assess and address when present, rather than a core symptom[34].

An additional area of uncertainty concerns the relationship between ADHD and a cluster of symptoms comprised of lethargy, underactivity, apathy, daydreaming, slow thinking, excessive sleep, and being easily lost in thoughts, named sluggish cognitive tempo (SCT) and, more recently, cognitive disengagement syndrome[35]. The nosological status of SCT, initially considered similar to ADHD inattentive presentation and now conceptualized as a transdiagnostic specifier across many disorders, needs further refinement, especially in relation to adult ADHD.

SCREENING AND ASSESSMENT

Screening and assessment for ADHD in adults is often more complex than in children, and must take into consideration cultural norms of age- and gender-appropriate behavior, as well as family values and environmental demands. The key message is avoiding rapid evaluations based only on checklists. Relying on multiple information sources for symptom and impairment ratings (e.g., family members, close friends, co-workers) can improve diagnostic accuracy[23,36], particularly with respect to establishing the presence of symptoms before age 12[37,38].

There can be barriers to obtaining informant reports in some clinical settings, but clinicians are nonetheless encouraged to pursue these reports, given the high rates of both false positive and false negative ADHD symptom reporting that may occur. Indeed, individuals may overinterpret normative cognitive variations as ADHD symptoms, or persons with longstanding ADHD may reject their diagnosis despite continuing to display impairing symptoms [39]. When it is not feasible to integrate informant report into the diagnostic process, a structured clinical interview is advised, so that the clinician can probe for concrete examples of reported ADHD symptoms and link these symptoms to impairment, or further question symptoms denied by the person but noted in the clinical record.

There is no age limit for an ADHD diagnosis. The diagnosis is possible and reliable in children as young as three years of age[39] as well as in older adults[40]. However, ADHD diagnoses may be delayed in women, in individuals who identify as ethnic or racial minorities, and in those with high intelligence. Sociocultural factors, barriers to care, and compensatory strategies are known to produce disparities in age of first ADHD diagnosis[41,42].

Since ADHD is defined by the presence of a persistent and age-inappropriate pattern of inattention and/or hyperactivity-impulsivity interfering with normal functioning or development, it is essential to recognize the main symptoms in these two domains (see Table 1). Of note, these behaviors should not be due to defiance or lack of comprehension. Further aspects that deserve attention during the assessment process are the age of onset of symptoms, their temporal stability, their situational pervasiveness, their incongruency with expected developmental patterns, and the extent to which they lead to functional impairment.

ADHD is classified as a neurodevelopmental disorder with childhood onset and a chronic course. However, recent data show that at least some individuals present a fluctuant pattern of ADHD remission and recurrence from childhood to middle adulthood [39,43,44]. While a duration of at least 6 months is helpful to establish

symptom stability, clinicians should expect that some symptoms wax and wane depending on contextual factors.

Individuals without ADHD may experience ADHD-like symptoms that arise as short-term responses to stressors such as family problems or higher academic/occupational demands. In these cases, it is helpful to construct a developmental timeline of the onset and offset of the symptoms, including factors that the person or the clinician perceives to influence symptom severity and expression. In this assessment, it is essential to set a common understanding with the person on a culturally acceptable definition of what is considered to be frequent.

The symptoms must be inconsistent with the developmental stage of the individual under assessment, which may differ from chronological age in adults with developmental disabilities. Since ADHD symptomatology is dimensionally distributed in the population, any clinical cut-off involves a level of arbitrariness, creating uncertainty in the assessment of individuals with milder symptoms. Thus, clinicians are faced with the difficult task of defining the boundaries separating typical from pathological behavior for each individual. In this scenario, extensive knowledge of normal human development and the person's sociocultural context is crucial.

ADHD symptoms in childhood are frequently found in referred adult cases. However, clinicians should be cautious when excluding ADHD only based on onset of symptoms after 12 years of age. Indeed, in some cases, demands from the environment are lower during childhood, and higher individual cognitive resources and/or family structure and support might prevent the expression of symptoms earlier in development. Impairing symptoms may eventually hatch during adolescence and young adulthood, when the individual faces needs of more autonomy or higher demands[45].

In most cases, the origin of symptoms can be linked to childhood (e.g., subthreshold or unimpairing difficulties recollected by the patient or a parent) and their escalation over time can be related to changing environment or developmental demands. When the clinician cannot trace symptom expression back to childhood based on available information, it is critical to perform a thorough differential diagnostic assessment prior to diagnosing adult-onset ADHD. Notably, research suggests that over 90% of apparent late-onset ADHD cases are ruled out once differential diagnostic procedures are applied[36].

Pervasiveness of symptoms in different environments is another key aspect for ADHD diagnosis. The underlying rationale is to avoid diagnosis in cases where symptoms manifest only in relation to environment-specific triggers (e.g., only at home due to severe family conflicts; only at university or work due to tasks/duties inappropriate for individual's capacities). For example, one study found that up to 40% of the population will report ADHD symptoms in a single context[46]. In many cases, single-setting ADHD symptoms are not associated with impairment or risks for negative outcomes[47]. However, some individuals who display impairment in just one setting at one point in time may show impairment in multiple settings at a later point, when facing more challenging demands[45], creating another diagnostic conundrum. As a result, it may be appropriate to apply a provisional diagnosis of Unspecified ADHD to individuals showing single-setting ADHD symptoms, and to monitor whether symptom pervasiveness increases over time.

Evaluation of impairment is a further important area of assessment. Impairment domains for adults with ADHD are broader than in childhood, including risky behaviors, interpersonal difficulties, underperformance at work or in higher education, financial problems, chronic record of motor vehicle accidents or unsafe driving, and impaired parenting. Since ADHD symptoms reflect a dimensional trait in the population, failure to properly incorporate the impairment criterion as part of the diagnostic criteria for the disorder may result in an explosion of prevalence rates[48]. Two clinical challenges here are: a) how to disentangle the source of impairment (is it coming from ADHD symptoms or from the very frequently associated mental disorders?); b) how to decide if impairment is sufficiently severe to warrant diagnosis (do we threshold impairment against the average peer or the hypothetical potential of the individual?). Clinicians may meet adults with mild ADHD symptoms who show negligible impairment but report internal distress, reduced self-esteem, and self-blame as a consequence of their symptoms. Presently, the DSM-5 classification suggests that ADHD cannot be diagnosed in this scenario, but Unspecified or Otherwise Specified ADHD may be an appropriate alternative diagnosis in these cases.

Diagnostic interviews and rating scales

A structured/semi-structured clinical interview with the patient is the gold standard tool in the assessment of adult ADHD. The instrument with the greatest empirical support is the Diagnostic Interview for ADHD in Adults (DIVA-5) (http://www.divacenter.eu/DIVA.aspx), a semi-structured interview based on the DSM-5.

ADHD is often not included or is screened on a limited basis in broadband adult psychiatric interviews, given its recent recognition as a disorder that presents in adults. Several structured and semi-structured interviews can be used for differential diagnosis in adults, such as the Structured Clinical Interview for DSM-5 (SCID-5)[49]. However, their use tends to be restricted to research settings.

There is uncertainty from practitioners, reflected in varying clinical practices and protocols, on the use of rating scales. Though not recommended as a standalone tool for diagnosing ADHD[50], they are helpful in the diagnostic process in adults. Their main utility for clinicians is related to: a) initial screening, that should be followed by clinical assessment, of ADHD symptoms in targeted populations (e.g., adults seeking treatment for substance use problems or another psychiatric disorder); b) obtaining information from collaterals on ADHD symptoms; and c) monitoring symptom trajectories during treatment.

The three rating scales with the best balance among psychometric properties are the Wender Utah Rating Scale-25 (WURS-25)[51], the Conners Adult ADHD Rating Scales (CAARS)[52], and the Adult ADHD Self-Report Scale-18 (ASRS-18)[53]. We focus here on the ASRS, an open-access instrument with two versions: a) a screener version with six items developed by the World Health Organization (WHO), that is suitable for primary care settings

and for a quick screening of ADHD[54]; b) a long version with the eighteen DSM symptoms, probably more useful for specialized settings. Both versions use wording more adequate for adults. A short version adapted for the DSM-5 is also available[55], and has strong properties when applied to screening.

A systematic review and meta-analysis of studies in children and adolescents[56] – that awaits replication in adults – concluded that most included rating scales have excellent overall diagnostic accuracy, as indicated by the area under the curve. However, the use of a single reporter is unlikely to achieve sufficient sensitivity and specificity for clinical use or population screening[56]. The same is probably true in adults.

The role of biomarkers and neuropsychological tests

There are no ancillary tests or biomarkers with sufficient positive and negative predictive power for the diagnosis of ADHD. No evidence supports the inclusion of neuroimaging exams – e.g., magnetic resonance imaging (MRI), single photon emission computerized tomography (SPECT), positron emission tomography (PET) – or electroencephalography in routine clinical assessment of ADHD, although they can be useful in very specific cases for differential diagnosis[57]. A systematic review of studies in children and adolescents[58], which awaits replication in adults, after examining 780 studies across neurodevelopmental disorders (including ADHD), could not find any biomarker with evidence – from two or more studies by independent research groups, showing results in the same direction – demonstrating specificity and sensitivity of at least 80%.

There are controversies and uncertainties in relation to the use of neuropsychological tests (e.g., continuous performance tests, executive function batteries), nowadays administered mainly digitally, in the diagnostic process. Recently, the UK National Institute for Health and Care Excellence (NICE)[59], after systematically reviewing the literature (including available meta-analytic evidence [60,61]) and consulting experts and individuals with lived experience, issued its recommendations on the use of digital technologies for the diagnosis of ADHD. They suggest that the QbTest, a test combining evaluation of motor activity using an infra-red camera with continuous assessment of attention and impulsivity, could be used as an option to support the diagnosis of ADHD in children. However, for adults, none of the available neuropsychological tests (including the QbTest) were endorsed, due to insufficient methodologically sound evidence. The guidance specifically warns against using these tests, which can be costly, as a triage system to assign patients to waiting lists for assessments – a practice common in some health care centers.

Differential diagnosis

One step in the differential diagnostic process is to rule out health conditions that may mimic ADHD symptoms. For example, sleep problems and deficits in visual and/or hearing acuity may be confused for ADHD. Although the relation between ADHD and sleep disorders/problems is complex (i.e., ADHD and sleep disorders such as restless leg syndrome might co-occur; sleep problems such as long sleep onset latency might be part of the ADHD phenotype or the result of ADHD treatment), some sleep disorders such as insomnia or obstructive sleep apnea might lead to inattentive and hyperactive symptoms during the day. Thus, a good assessment of sleep conditions is mandatory in differential diagnostic assessment.

Physical and laboratory investigations can help in excluding other clinical conditions (e.g., hyperthyroidism, traumatic brain injury). Likewise, it is important to rule out the use of any medication that might cause inattentive and/or hyperactivity/impulsive symptoms. Referral for genetic examination is recommended if there is a clear developmental delay and/or if suggestive phenotypes are present (i.e., neurofibromatosis type 1; fragile X syndrome)[63].

Difficulties with attention are among the most common symptoms listed in the DSM-5-TR[64]. Thus, a number of psychiatric conditions – including mood and anxiety disorders, post-traumatic stress disorder (PTSD), psychotic disorders, neurocognitive or other neurodevelopmental disorders, impulse control disorders, and substance use intoxication or withdrawal – must be ruled out.

Some disorders that commonly co-occur with ADHD (e.g., generalized anxiety disorder, bipolar disorder, major depression, learning disorders, and PTSD) must be ruled out as the sole source of ADHD-like symptoms, which can be challenging and create diagnostic uncertainties. Of note, ADHD and autism frequently co-occur, and the presence of ADHD symptoms in patients with autism spectrum disorder generates distinct problems.

In the process of conducting a careful differential diagnosis, some clinical tips might be relevant: a) consider the age of onset of the symptoms (for example, when disentangling inattention as part of ADHD versus a mild chronic depressive disorder, the occurrence of inattentive symptoms before the onset of any mood symptoms reinforces the ADHD diagnosis); b) examine the trajectory of the symptoms (for example, the clearly episodic occurrence of hyperactivity, impulsivity and irritability can suggest bipolar disorder); and c) assess if the way symptoms manifest is better explained by another mental disorder (e.g., inattention only as a consequence of dysfunctional thoughts/rumination related to performance as in generalized anxiety disorder; inattention related to rituals of counting as in obsessive-compulsive disorder; inattention and executive deficits following abuse of cannabis without any previous history of ADHD symptoms; inattention only related to reading in dyslexia)[63].

ASSOCIATED CONDITIONS

Psychiatric comorbid conditions

ADHD is frequently comorbid with other psychiatric conditions. This comorbidity is associated with greater ADHD symptom severity[65], stronger impairments in daily functioning[66], higher health care needs[67], and higher mortality[68]. Robust knowledge on comor-

bid conditions of ADHD is potentially helpful for the prevention of their onset, and for monitoring and treatment decisions when (early signs of) comorbid conditions have already developed[69,70].

In childhood, it is well-established that oppositional defiant disorder, conduct disorder, childhood-onset anxiety disorders and autism are often comorbid with ADHD[71-73]. While none of these conditions are likely to (fully) remit in adulthood, there is an important gap in the literature, as few longitudinal studies have looked at their continuity from childhood through adolescence and into adulthood.

Studies on psychiatric conditions comorbid with ADHD in adults are mostly separate from those in childhood and have predominantly focused on the so-called "common mental disorders", i.e. anxiety, depressive and substance use disorders. A systematic review and meta-analysis of this literature[70] – including findings from general population studies based on national registers, insurance claims data, and large surveys (N>10,000) – showed strong differences in adults with ADHD compared to those without this disorder. The meta-analysis reported pooled odds ratios (ORs) of 5.0 (95% CI: 3.29-7.46) for anxiety disorders, 4.5 (95% CI: 2.44-8.34) for major depressive disorder, 8.7 (95% CI: 5.47-13.89) for bipolar disorder, and 4.6 (95% CI: 2.72-7.80) for substance use disorders. Other, less extensively studied, psychiatric conditions that are more frequent in adults with than without ADHD are (specific) personality disorders, eating disorders and schizophrenia[74,75]. These disorders, together with the potential (heterotypic) persistence of childhood comorbid conditions, need additional general population-based studies.

Somatic comorbid conditions

In children with ADHD, an elevated risk has been found for obesity and asthma[76-78]. Other somatic conditions for which an increased risk has been reported in children with ADHD include rhinitis, food allergy, dermatitis, and type 1 and 2 diabetes mellitus[78,79]. The risk of obesity may be somewhat higher in adults (OR=1.6, 95% CI: 1.3-1.8) than in children (OR=1.3, 95% CI: 1.2-1.5) with ADHD[76], while the risk of asthma seems to be similar (OR=1.4, 95% CI: 1.4-1.4 in adults; OR=1.6, 95% CI: 1.2-2.1 in children)[78,80].

For several years, few somatic conditions have been thoroughly studied in relation to adult ADHD beyond obesity and asthma[81]. More recently, there has been a surge of studies, based particularly on Swedish health registers. A study[82] reported associations of adult ADHD with obesity (OR=2.7, 95% CI: 2.6-2.8), asthma (OR= 2.4, 95% CI: 2.3-2.5), sleep disorders (OR=4.6, 95% CI: 4.4-4.8), migraine (OR=2.0, 95% CI: 1.9-2.1), epilepsy (OR= 3.0, 95% CI: 2.8-3.2), and chronic obstructive pulmonary disease (OR=3.2, 95% CI: 3.0-3.6). Studies digging deeper into specific somatic disease categories showed that adult ADHD was associated with a diagnosis of any of 13 investigated autoimmune diseases (OR=1.3, 95% CI: 1.3-1.4), with estimates ranging from 1.1 (95% CI: 1.0-1.2) for ulcerative colitis to 1.8 (95% CI:1.4-2.3) for Sjögren's syndrome[83].

Another study[84] showed an increased risk across all types of cardiovascular diseases, even when use of stimulants and other psychotropic drugs was accounted for (hazard ratio, HR=2.1, 95% CI: 2.0-2.1), with the highest risk for cardiac arrest, hemorrhagic stroke and peripheral vascular disease/arteriosclerosis. A meta-analysis on type 2 diabetes mellitus showed an OR of 2.29 (95% CI: 1.48-3.55), which was confirmed in that same paper by an HR= 2.35 (95% CI: 2.14-2.58) based on data from the Swedish registers[85]. A strong association was also found between adult ADHD and sleep disorders, with ORs ranging from 6.4 (95% CI: 6.0-6.7) in mid-to-older adulthood to 12.6 (95% CI: 12.1-13.1) in young adulthood[86]. All these findings were confirmed in a study based on insurance claims data from Germany[80], which also included data from primary care, reflecting an overall poorer somatic health in adults with ADHD. Finally, two studies[80,82] found an enhanced risk of Alzheimer's disease and Parkinson's disease (OR range: 5.2-7.1, aggregated 95% CI: 4.5-9.3).

General considerations

The associations with these psychiatric and somatic conditions have important implications for the management of ADHD (see below). However, some uncertainties remain that require additional studies or novel research strategies. One uncertainty is the extent to which comorbid psychiatric and somatic conditions persist after controlling for a series of important confounders. A study on the association between ADHD and asthma first retrieved all significant confounders with a systematic review and then assessed the association using health registers while controlling for all those confounders[87]. Second, temporal links of ADHD with psychiatric and somatic conditions are mostly unclear, with some prominent exceptions[69,75]. Assuming that ADHD may predate most comorbid conditions, prospective studies are particularly important for knowledge on the age of onset of comorbid conditions and therefore for optimal timing of preventive programs. However, cross-sectional studies on psychiatric and somatic comorbidities that report on different developmental periods from childhood to old age are a reasonable alternative[67,80].

Third, psychiatric comorbid disorders are potentially part of the pathway between ADHD and onset of somatic conditions. This is clear for alcohol-related liver disease (OR=4.7, 95% CI: 3.7-5.6)[82] and likely holds for other somatic conditions as well. Finally, the idea that we could prevent adult-onset conditions that are highly comorbid with ADHD assumes that aspects of ADHD are causal factors in these onsets. This needs more research, including study of possible mediators between ADHD and the comorbid condition, for instance through Mendelian randomization approaches. A recent example[88] focused on the pathway between ADHD, mediators including obesity, and type 2 diabetes mellitus.

EPIDEMIOLOGY

Among school-aged children, the prevalence of ADHD, based on epidemiological studies representative of the general population, is estimated at around 5.5%[89]. Although there is some varia-

tion among studies due to methodological differences, most notably the application of impairment criteria, the prevalence of the disorder is similar across geographic regions[90,91]. By contrast, the administrative prevalence – i.e., that determined based only on administrative records, e.g., billing records – varies across regions, based on multiple factors, including awareness of ADHD, training of clinicians, and conceptualization of the disorder. Only half of individuals with ADHD are diagnosed before age 14[92].

Longitudinal studies[93] have documented an age-dependent decline in symptoms from childhood through adulthood, such that most children with ADHD will no longer meet full criteria for the disorder by age 30. This decline is more marked for hyperactivity and impulsivity compared with inattention[94]. Persistence of ADHD is predicted by disorder severity, psychosocial adversity and psychiatric comorbidity[95-98].

The most recent meta-analysis points to an overall prevalence of 2.5% in adults, with a gradual decline to 1% by age 60[95]. Population studies show that, when the age at onset criterion for the disorder is ignored, prevalence increases to about 9% in early adulthood and 4% at age 60[95].

There are several challenges and uncertainties when estimating the prevalence of ADHD in adults. First, while most individuals with a diagnosis of ADHD in childhood will not meet formal diagnostic criteria in adulthood, about 71% will present with ADHD symptoms, and 65% with functional impairment[93]. Therefore, a crucial issue is whether formal criteria or functional impairment are assessed.

Second, there is an ongoing debate on the existence of cases with late onset, i.e., starting in late adolescence or later on. Indeed, although diagnostic criteria for ADHD require onset prior to age 12, some have argued that adult-onset ADHD is common and distinct from childhood-onset ADHD[38]. It is likely that most of these late-onset cases had indeed symptoms of ADHD in childhood that they were able to compensate for until adulthood. In fact, many individuals in the late-onset group exhibit some ADHD symptoms during childhood or display an externalizing disorder such as oppositional defiant disorder[29]. Moreover, current studies indicate that the majority (if not all) of late-onset ADHD cases emerge between the ages of 12 and 16, classifying them as adolescent or early adult onset ADHD[29].

However, the idea that some cases present with ADHD onset in adulthood remains an area of controversy[99]. Overall, caution should be urged in diagnosing ADHD when onset occurs in adulthood[99], although such onsets can occur due, for example, to traumatic brain injury[100].

Third, while ADHD has traditionally been considered a stable, chronic condition, more recent follow-up studies indicate that, after ADHD remits, it can recur[39]. An analysis of the Multimodal Treatment Study of ADHD, in which patients underwent eight assessments during follow-ups ranging from 2 to 16 years after baseline, showed that about 60% of them experienced a recurrence of ADHD after the initial period of remission[39]. Another study of three independent cohorts suggested that about a quarter of ADHD youth will have a fluctuating course[101].

ADHD is two to three times more common in males than females in the general population[90]. The sex difference in clinics is much larger, because girls – who tend to be inattentive and not disruptive – are less likely to be referred for treatment[102,103]. The sex ratio decreases with age, so that by adulthood it is close to 1[104].

In a meta-analysis of epidemiological population studies comprising 218,445 participants[105], no significant differences in the prevalence of ADHD were found between Black, White, Asian and Latino individuals. There was substantial heterogeneity for each minority subgroup, but meta-regression could not find the reason for it. Moreover, significant publication bias was detected. Data regarding clinical diagnoses are more consistent. Several studies suggest that underdiagnosis occurs in Black and Hispanic groups in the US[106-109]. In Europe and Israel, immigrants are less likely to be diagnosed with ADHD compared to non-immigrants[109]. Lower treatment rates have also been documented for minority groups [108-110].

ADHD treatment in children has increased rapidly in recent years[111,112]. The number of published studies in adults is far lower than in children. In a register study based on the entire adult population in Denmark, Finland, Iceland, Norway and Sweden[113], the annual prevalence of ADHD drug use increased during the study period for both genders and all age groups (from 2.4 to 5.3 per 1,000 men, and from 1.8 to 4.4 per 1,000 women). Another multi-national study using population-based databases from 14 countries[112] reported that, among adults aged 19 years or older, the prevalence of any ADHD medication use in 2010 varied between 0.003% and 1.48% (0.05% in Asia and Australia, 1.42% in North America, 0.47% in Northern Europe, and 0.03% in Western Europe). The absolute increase in ADHD medication use prevalence per year ranged from 0.0006% to 0.12%. So, the available evidence suggests that use of ADHD medications in adults is rising in many countries. This may be associated with an increased awareness and diagnosis of adult ADHD. Furthermore, many children that grow into adulthood are continuing their ADHD medication, hence increasing the prevalence of medication use in adults.

BURDEN

Functional impairment

One controversy about defining impairment in adults with ADHD is whether clinicians should compare a patient's functioning to that of the general population or to the patient's potential as indicated by measures such as IQ or aptitude tests. Using the population as a benchmark has the advantage of being objective, especially where there are standards for what constitutes impairment. For example, chronic unemployment is easier to see as impairment than a moderately successful physician struggling to maintain a practice. Assessing impairment relative to potential recognizes that ADHD symptoms may limit a person's ability to meet his/her own goals or expectations, even if he/she performs adequately compared to the population. Indeed, research has validated the diagnosis of ADHD in highly intelligent patients[114-117]. A highly intelligent adult may meet average workplace standards, yet underperform

relative to his/her potential, leading to frustration, dissatisfaction, and a lack of fulfilment. In practice, clinicians should balance these perspectives by first assessing impairment relative to the general population and then gathering information about the patient's potential, expectations, goals, and self-perceived limitations, which often reveal struggles that standardized measures miss.

It can also be difficult to define and assess functional impairment in older adults[118]. This issue is particularly challenging given the overlap between ADHD symptoms and other problems common in older age, such as cognitive decline and physical health issues, as well as the effects of long-standing compensatory strategies. Traditional definitions of impairment in ADHD focus on domains such as occupational performance, educational achievement, and parenting responsibilities. These domains do not fully capture the challenges faced by older adults, for whom impairment manifests more prominently in areas such as social functioning, health management, financial organization, or maintaining independence.

Another complicating factor is that many older adults may not recognize their struggles as impairments because they have lived with these challenges for decades. This can result in underreporting of challenges and underestimation of their impact. Furthermore, social expectations and norms for older adults may lower the perceived significance of certain impairments, such as difficulty managing time or multitasking.

When evaluating older adults for ADHD, a comprehensive assessment should include an exploration of functional difficulties across age-relevant domains. Collateral information from family members or close friends is especially valuable, as older adults may have difficulty identifying their own impairments. It is essential to evaluate the patient's quality of life and goals. For some older adults, even mild impairments in functioning may have a significant impact on their sense of well-being and autonomy. Clinicians should take these subjective experiences into account when assessing the need for a diagnosis and potential interventions.

Economic impact

Several studies have estimated the economic burden of ADHD. For instance, a study estimating the incremental costs of ADHD (i.e., excess costs over and above those of individuals without ADHD) in the US – in relation to health care, productivity and income losses, education, and justice system – reported costs of $1,137-4,100 per adult per year[119], which is similar to that of chronic complex somatic conditions[120].

Comorbidities, including somatic diseases, are common in ADHD and are important cost drivers. A study[121] that prospectively followed a cohort of 445,790 adults from ages 18 to 26 found that the annual per capita costs associated with multimorbidity were $1,223 for individuals with a childhood ADHD diagnosis. Among these, costs were higher for persisters ($1,456) compared to remitters ($837). The costs for individuals without an ADHD diagnosis were significantly lower, at $418. The main drivers of the above costs were inpatient hospital admissions, primarily due to drug abuse and injuries. Another study in Sweden[122] found that middle-aged adults (30-45 years) newly diagnosed with ADHD had significantly higher health care costs and utilization compared to those without ADHD. Data from individuals born 1966-1978 showed greater outpatient, inpatient and medication costs for psychiatric and somatic comorbidities, with females incurring higher medication costs than males.

Using data from the Danish National Registers in 5,269 adults diagnosed with ADHD in adulthood, a cross-sectional analysis[123] for the year 2010 compared costs incurred by adults with ADHD and their siblings, using data from health, education, crime, employment, and social care registers. Adults with ADHD were found to have significantly lower disposable incomes and paid less tax than their siblings. They received more state benefits and incurred higher costs related to health care, social care, and crime.

Overall, the available evidence highlights that ADHD imposes substantial economic costs on both individuals and the state. This underscores the need to consider the broader economic implications of ADHD, extending beyond income and health care-related expenses.

ETIOLOGY

The etiology of ADHD has been investigated using multiple approaches. Genetics has been the most frequently applied among these approaches, both based on family and twin studies as well as using molecular genetics methods. More recently, also other types of molecular "omics" have been explored, including analyses of the epigenome and transcriptome as well as the microbiome. In all those studies, childhood ADHD has thus far been the main topic, but studies of adult ADHD have also been conducted at least in some areas of research.

Genetics

Twin and family studies

Twin and family studies show that genetics contributes substantially to the etiology of ADHD, with heritability estimated at 70-80% [124]. Twin studies of ADHD in adults have reported lower heritability (30-40%). However, this is not the case in studies using multiple informants[125] or clinical diagnosis[126], in which the heritability of ADHD is >70% also in adults[127].

Large-scale cross-generational analyses indicate a genetic correlation of approximately 0.5 between child and adult presentations of ADHD, suggesting that developmental changes in ADHD presentation may be partly underpinned by genetics[128]. Nonetheless, few longitudinal twin studies of ADHD extend beyond young adulthood, preventing definitive conclusions of whether all forms of adult ADHD represent a continuation of childhood ADHD[29,128]. Twin and family data exploring the etiology of ADHD in later life are lacking.

Twin studies in childhood support the dimensional nature of

ADHD, with the genetic correlation between diagnosed ADHD and ADHD symptoms in the general population estimated to around 0.6[129]. While such conclusions are expected to extend to ADHD in adults, this remains to be tested. The high rates of psychiatric comorbidity in ADHD (e.g., for depression, eating disorders, bipolar disorder, and substance use disorders) are partly mediated by shared genetics, with cross-condition genetic correlations estimated at about 0.5 in both children and adults[130,131].

While much less researched, recent family studies also support genetically mediated links between adult ADHD and somatic conditions, including asthma, obesity, migraine and cardiovascular diseases[82,132], and provide tentative support for a link with neurodegenerative conditions.

Molecular genetics

In the largest available meta-analysis of genome-wide association studies (GWAS) of ADHD (comprising data on 38,691 patients and 186,843 controls), 27 significant loci were found, implicating 76 genes, many of which are upregulated in early neurodevelopment[133]. This meta-analysis primarily comprised children diagnosed with ADHD. Only two GWAS of ADHD in adults[134,135] have been conducted so far, both reporting a strong genetic correlation (>0.8) between ADHD in children and adults. Thus, current evidence from research of common genetic variants indicates a largely similar genetic background of ADHD in adults and children, particularly when adult ADHD is defined by persistence from childhood into adulthood.

No large-scale genetic studies have yet looked into well-defined adult-onset ADHD. However, some studies have reported differences in genetic profiles of common variants between individuals with persistent ADHD and those either first diagnosed with ADHD in adulthood or with symptoms first emerging in adolescence. The latter two appear to have a lower genetic burden for ADHD (as determined based on polygenic scores), stronger positive genetic links to depression and substance misuse, and stronger negative genetic links to education and cognition[29,134,136,137]. Yet, definitive conclusions cannot be drawn, as the studies on clinical diagnoses lack data on symptoms in childhood, and no longitudinal studies of sufficient sample size are yet available.

Studies using GWAS data and ADHD polygenic scores for genetic correlation analyses suggest a strong genetic link between clinically diagnosed ADHD and ADHD symptoms in the population[138], as well as a genetic link between ADHD and several psychiatric and somatic conditions, negative health behaviors (e.g., smoking initiation), psychosocial risk factors (e.g., low socioeconomic status), phenotypes in cardio-metabolic (e.g., higher body mass index and cardiovascular risk) and reproductive (e.g., lower age at first childbirth) domains, and even reduced longevity[133,138-140]. This implies that genetic variation identified in the GWAS largely based on childhood ADHD also confers risk for important health and behavioral outcomes measured across the lifespan.

In addition to common genetic risk variants, rare variants (with potentially larger effect sizes) have also been analyzed to explain ADHD etiology. Individuals with ADHD carry an increased burden of rare, protein-truncating variants in evolutionarily constrained genes[141]. One study suggested a higher rare variant burden in individuals diagnosed with ADHD in childhood, compared to those first diagnosed as adults[134]. Around 10-15% of individuals with ADHD carry rare copy number variants (CNVs)[124,142,143]. Many of these CNVs are also implicated in autism spectrum disorder and schizophrenia, and thus seem to exert effects across disorders with different ages of onset[142,144].

Epigenetics and transcriptomics

Epigenetic modification of DNA (DNA methylation) is an important factor in the regulation of gene activity, influenced by both genetic and environmental variables. Various epigenome-wide association studies (EWAS) have been performed for ADHD[145], analyzing between 450,000 and over 800,000 sites of variable DNA methylation. Most studies have focused on children with clinical diagnosis or symptoms of ADHD. However, the largest EWAS, with over 4,500 participants, was on ADHD symptoms in adults[146]. Though several interesting candidate genes were identified, no findings reproducible across the three analyzed cohorts were observed. Larger sample sizes are clearly needed before significant findings can be expected from EWAS. Moreover, epigenetic modifications are cell type-specific. Although we consider ADHD a brain disorder, epigenetic studies make use of DNA isolated from blood or buccal cells, which has to be considered in data interpretation.

The "transcriptome" represents the output of active genes – the DNA template of a gene is transcribed to produce RNA. Studies of the transcriptome in ADHD are of two types[147]: those that compare RNA isolated from the blood of people with and without ADHD[148,149], and those that integrate GWAS findings with transcriptome data from more relevant tissues/cell types (i.e., different brain areas/cell types) in so-called transcriptome-wide association studies (TWAS)[137,150-153]. All of the latter studies identified novel candidate genes for ADHD. Only one of these studies discriminated between childhood and adult-diagnosed ADHD in the GWAS they used as input information[137]. Though different genes were identified in the two resulting TWAS, the study was underpowered to test significance of differences between the two groups.

Emerging picture of the molecular biology of ADHD

Based on the results of the above-mentioned studies, we are starting to get a first glimpse of the molecular biology of ADHD. Enrichment analyses are being used to link genetic, epigenetic and transcriptomic findings to biological pathways, developmental stages of brain development, and even individual brain cell types. Gene enrichment analyses based on the results of the latest GWAS [133] link early brain development to ADHD and indicate a role for dopaminergic and GABAergic systems and glial cells in ADHD etiology. Integration of candidate genes from epigenetic studies sug-

gests involvement of neurogenesis, neuronal differentiation, cell adhesion, and axon guidance processes in ADHD[145]. TWAS reveal enrichment for several biological pathways as well, including dopaminergic neuron differentiation, noradrenaline release cycle, and triglyceride lipase activity[152]. Since the studies are heavily focused on children with ADHD, we will have to await confirmation of findings for adults.

In summary, twin and molecular studies are greatly advancing our understanding of the etiology of ADHD, and suggest that genetic factors play a substantial role and are largely shared across ages. Nonetheless, well-powered studies of ADHD in adults which contain both "omics" data and information on developmental symptom trajectories are scarce. As such, we are yet to discover if there are more nuances in the etiology of ADHD with increasing age. In addition, genetic risk interplays in complex ways with environmental exposures.

Environmental risk factors

While a large number of studies have found correlations between ADHD and environmental risk factors, such as maternal pre-pregnancy obesity and smoking during pregnancy, such associations must be interpreted with caution[154]. Alternative explanations such as familial confounding (i.e., genetic or other familial variables contributing to both the risk factor and ADHD) need to be addressed to strengthen causal inference. Twin, sibling and family studies have been used to disentangle the effects of environment from genetics, demonstrating that low birth weight, gestational age, and family income in childhood are associated with ADHD even after adjustment for familial confounding[155]. In contrast, several putative risk factors, including pregnancy-related factors (e.g., pre-pregnancy obesity and maternal smoking during pregnancy) were primarily explained by familial confounding[155]. The majority of gene-environment interplay research in ADHD has focused on early development, leaving a large knowledge gap about how environmental hits across the lifespan may interact with dynamic genetic risk to shape the expression of ADHD in adults.

A factor that has received increasing attention is the microbiome. The role of the gut microbiome in adult ADHD remains controversial and uncertain, due to inconsistencies in research findings. Studies differ in the phenotypes analyzed, sequencing methods, statistical approaches, and reported results, making comparisons difficult. A meta-analysis[156] pooling data from four adult ADHD case-control studies (N=617) found that beta diversity was associated with ADHD diagnosis. Specific microbial genera showed robust associations with ADHD: *Ruminococcus torques* was more abundant in ADHD and linked to hyperactivity/impulsivity, while *Eubacterium xylanophilum* was less abundant. These genera may influence inflammatory processes. However, significant heterogeneity between cohorts persisted despite harmonized analyses, underscoring the need for larger meta-analytic studies. Moreover, further research is essential to clarify the possible role of microbiome in ADHD pathophysiology and its potential as a therapeutic target.

NEUROPSYCHOLOGY AND NEUROBIOLOGY

Neuropsychology

Acquired damage to the frontal lobes typically gives rise to a dysexecutive syndrome, in which patients have problems with executive functioning[157] (a constellation of cognitive processes which allow humans to behave in a goal-directed manner). Dysexecutive syndromes are characterized by difficulties in planning, in inhibiting unwanted or inappropriate responses under various environmental circumstances, in using working memory to guide behavior purposively, or in maintaining consistent performance over time, particularly in routine scenarios[157].

The striking parallels between these features of the dysexecutive syndrome and the symptoms of ADHD (e.g., inattention, impulsivity) has spurred decades of research aimed at isolating patterns of cognitive difficulties in individuals with ADHD. It is hoped that discovering reliable signatures of neuropsychological difficulty in ADHD will not only inform knowledge of the underlying neural substrates, but also reveal candidate processes for remediation and rehabilitation, and even aid diagnosis.

Meta-analyses have found that ADHD is indeed associated with difficulties in executive functions – including working memory, reaction time variability, response inhibition, and planning/organization – compared to controls[158]. For adults with ADHD, meta-analyses indicate deficits in domains including decision making[159], working memory[160], focused and sustained attention[161], verbal fluency[162], set shifting[162], and verbal memory[163,164].

Different patterns of activation on functional MRI during neurocognitive tasks between individuals with ADHD and controls have been found. For example, a meta-analysis of 23 studies of response inhibition found decreased activation in the supplementary motor area, insula, caudate, and precentral gyrus, and increased activation in the postcentral gyrus, inferior frontal gyrus, and precuneus in people with ADHD[165]. There were greater decreases in children versus adults with ADHD in the right caudate.

A recent meta-analysis explored the effect of stimulant (methylphenidate) and non-stimulant (atomoxetine) medication use (minimum 3 days) on the executive functions of people with ADHD [166]. This study found a significant effect of methylphenidate for all neurocognitive domains, with the largest effect for attention and the lowest for reaction time (i.e., overall speed). The meta-analysis for atomoxetine found beneficial effects for all neurocognitive domains except for working memory. There were no significant differences in effect sizes between adults and children. An outstanding question resulting from this study is the extent to which the effects of medications on cognitive function were related to changes in symptoms and/or quality of life measures.

It is important to note that there is much heterogeneity in neuropsychological performance in individuals with ADHD[167], which may reflect multiple pathways in the brain that are relevant to the etiology of the disorder. Average effect sizes between controls and people with ADHD are much smaller for neuropsychological tests than for ADHD symptoms, indicating that differences in neuropsychological performance may be minimal in many people with

ADHD[158].

There are uncertainties and controversial views on the value of neuropsychology for clinical practice. While neuropsychological tests can differentiate people with ADHD from controls, there are no tests that can differentiate ADHD from other clinical cohorts, resulting in these tests not being useful in the diagnosis of ADHD[168], nor recommended as such by evidence-based guidelines[169]. Nevertheless, neuropsychological testing can be helpful in understanding a person's unique pattern of cognitive strengths and difficulties, thereby contributing to guide treatment, education and occupational choices.

Neuroimaging

Neuroimaging studies in ADHD have primarily focused on the paediatric population, although studies in adults are gradually increasing, highlighting age-related differences[170-172]. Structural MRI meta- and mega-analyses reported diffuse volumetric and morphometric alterations in cortico-subcortical brain regions in children, but reduced or absent case-control differences with growing age[173-175], in line with the frequently observed symptomatic improvement or remission in adulthood[176]. Conversely, a recent meta-analysis of diffusion-weighted imaging studies reported that the identified case-control differences in the corpus callosum did not survive when restricting the analyses to paediatric studies[172]. These apparently contrasting findings may be related to the distinct developmental trajectories of the gray and white matter, so that a maturational delay of gray matter may be more evident in childhood whilst that of the white matter may be more pronounced later in life. Therefore, longitudinal studies are needed to clarify the life course of brain alterations in ADHD and their relationship with its variable outcome in adulthood.

Another striking aspect of the imaging literature in ADHD is the limited convergence of results when pooling case-control studies in meta-analyses[170-172]. This has been related to both methodological and clinical heterogeneity. Suboptimal MRI acquisition and pre-processing may potentially lead to spurious results, and the substantial variation in data acquisition, pre-processing and analysis, as well as in study design and statistical procedures, limits comparisons among studies[170-172]. Ongoing methodological developments may improve reliability of findings. On the other hand, ADHD is also a highly heterogenous condition both clinically and neurobiologically[176], with preliminary evidence of brain differences between ADHD presentations, sexes, and treated vs. untreated individuals[172,177-179]. Further, more pronounced alterations may be associated with comorbidities or symptom persistence[180,181]. Most neuroanatomical investigations have focused on group comparisons with controls, yielding inconsistent or no results, especially in adults. Thus, there is an increasing need to move beyond case-control comparisons and parse neurobiological heterogeneity, perhaps especially in individual characteristics relevant to clinical practice, such as symptom persistence and treatment response[179-181].

From a functional standpoint, one of the most innovative models of ADHD is the default mode network theory[182], according to which the brain's default mode network, which is normally active during rest and self-reflection, is overly active in individuals with ADHD. This leads to difficulties in sustaining attention and regulating behavior, as the individual is distracted by internal thoughts rather than focusing on external tasks. However, two meta-analyses[171,183] aimed at testing this hypothesis in children, adolescents and adults have reached contrasting conclusions, probably due to different methodologies. As such, the default mode network theory of ADHD deserves further testing, particularly in adults.

Neuropsychopharmacology

Medications used for ADHD include stimulants (amphetamines and methylphenidate) and non-stimulants (such as atomoxetine, clonidine, guanfacine and viloxazine)[184].

The primary mechanism of action of amphetamines is to elevate extracellular dopamine and noradrenaline levels at the synapse. This occurs through the inhibition of dopamine and noradrenaline transporters, which decreases the reuptake of these neurotransmitters from the synaptic cleft[185,186]. Amphetamines also enhance vesicular dopamine release in a dose-dependent and region-specific manner, by inhibiting the vesicular monoamine transporter 2. Additionally, they inhibit monoamine oxidase activity, reducing the breakdown of cytosolic monoamines[185,188]. The striatum seems to be the primary site of action of amphetamines, although direct effects have also been observed in the cortex and the ventral tegmental area[185,186].

The direct effects of methylphenidate involve inhibiting dopamine and noradrenaline transporters, exhibiting agonist activity at the 5-HT1A receptor, and redistributing vesicular monoamine transporter 2. These effects result in increased extracellular levels of dopamine and noradrenaline[185]. Several studies also indicate that methylphenidate directly interacts with adrenergic receptors, and activation of alpha-2 adrenergic receptors has been shown to stimulate cortical excitability[185,189].

Atomoxetine selectively inhibits the noradrenaline transporter, and increases extracellular synaptic levels of noradrenaline and dopamine in the prefrontal cortex. Clonidine and guanfacine stimulate the postsynaptic alpha-2 adrenergic receptors[184], but the neural mechanisms by which they improve ADHD symptoms are still not entirely clear[181]. Viloxazine inhibits the reuptake of noradrenaline in the synapse, which increases the extracellular levels of this neurotransmitter. Additionally, the drug has a serotonin modulating activity, whose role in its effectiveness for ADHD is still being explored[190].

PHARMACOLOGICAL TREATMENT

ADHD simplex

Pharmacological treatment represents the cornerstone of the management of adult ADHD[184]. However, recommendations on

the choice of medication vary somewhat across available guidelines, reflecting not only the evidence base but also the licensed medications in the various countries. For instance, the 2018 (updated in 2019) UK NICE guidelines[59] recommend methylphenidate or lisdexamfetamine (or dexamphetamine if lisdexamfetamine is not well tolerated) as first-line, followed by atomoxetine as second-line. The 2018 guidelines of the Association of the Scientific Medical Societies in Germany[191] state that medication is the first-line treatment, without specifying the class/formulation or a ranking. The guidelines of the Canadian ADHD Resource Alliance (CADDRA)[192] provide a recommended ranking for children: long-acting stimulants are first-line treatment agents; atomoxetine, guanfacine XR, and short/intermediate-acting psychostimulants are second-line; and bupropion, clonidine, imipramine and modafinil are examples of third-line treatment agents. At the time of writing, the first US guidelines on the assessment and management of ADHD in adults are in the process of being developed[193].

In terms of the evidence base, a network meta-analysis of 113 randomized controlled trials (RCTs) examining pharmacological and non-pharmacological treatments for adult ADHD[21] reported that, with respect to efficacy (i.e., reduction in ADHD symptom severity) in the short term (i.e., at about 12 weeks of treatment), stimulants and atomoxetine were the only treatments to perform better than placebo in both clinician-rated and self-reported assessments. Effect sizes for stimulants ranged from 0.39 to 0.71, while atomoxetine showed effect sizes ranging from 0.38 to 0.51. A previous network meta-analysis in children[194] reported slightly higher effect sizes. Additionally, while the network meta-analysis concerning adults[21] found no significant differences in efficacy between methylphenidate and amphetamines, earlier findings in children[194] indicated higher effect sizes for amphetamines.

Regarding acceptability (i.e., trial dropout rates due to any cause), the network meta-analysis concerning adults[21] reported that stimulants were as acceptable as placebo, whereas atomoxetine was less acceptable. However, in terms of tolerability (i.e., dropout rates due to adverse events), all medications performed worse than placebo. Both stimulants and atomoxetine also improved emotional dysregulation in adults in the short term (up to 12 weeks), but they were not effective in improving other relevant outcomes such as executive dysfunction and quality of life, contrary to the findings of a previous meta-analysis of RCTs in children and adolescents[195].

It is important to note that the effects observed in network meta-analyses, as well as guideline recommendations, are based on averages at the group level. At the individual level, patients may respond preferentially to specific medications, but currently there are no reliable predictors of response. As a result, prescribing ADHD medications remains a trial-and-error process, highlighting a critical gap in the field.

Alongside the choice of medication, dose optimization is another crucial factor. Some guidance documents (e.g., the CADDRA guidelines[192] and the British National Formulary[196]) recommend maximum doses higher than those licensed by regulatory bodies such as the US Food and Drug Administration (FDA). This creates uncertainty regarding the effective and tolerable maximum dose. A dose-response meta-analysis[197] of 47 RCTs addressing this issue found that, for methylphenidate, increasing doses led to additional symptom reductions, though these gains diminished with higher doses and were accompanied by increased risks of adverse events and dropouts. Unlicensed doses provided slightly greater symptom reductions compared to licensed doses, but these gains were small and associated with a higher risk of dropout due to adverse events. For amphetamines, the dose-response curve plateaued, indicating no additional symptom reduction with higher doses, while the risk of adverse event dropouts continued to increase.

A further crucial aspect that remains controversial is the long-term efficacy and effectiveness of ADHD medications. Notably, the majority of the RCTs included in the above-mentioned network meta-analysis concerning adults[21] had a duration of less than 12 weeks, and it is challenging from an ethical and practical standpoint to conduct longer-term RCTs. The available evidence does not support any medication as being more efficacious than placebo in the long term when relying on both self- and clinician-rated scales. Atomoxetine was more effective than placebo at 26 weeks according to self-reported ratings, while stimulants were more efficacious than placebo on clinician-reported ratings. At about 52 weeks, no medication has supporting evidence of being more efficacious than placebo. In clinical practice, patients may report that the medication that was once effective seems to no longer work. Some studies suggest the possibility of a change in the availability of dopamine transporters, which might underlie decreased efficacy and effectiveness of medications over time[198].

A study design that may be informative regarding longer-term effects is the discontinuation-controlled trial, in which patients who have been treated with ADHD medication for years are randomized to either continue the medication or switch to placebo. Given the scarcity of such trials of ADHD medications in adults[199,200], more are needed to better estimate long-term outcomes. Furthermore, since individuals with severe ADHD may tend to decline participation in such studies, these RCTs may include only milder cases. Thus, the extent to which medications are effective in the longer term remains to be better elucidated.

ADHD with comorbidities

While limited evidence exists for adults, a larger literature in children shows that stimulant and non-stimulant agents are effective in treating ADHD in individuals with autism spectrum disorder, particularly in the context of higher intellectual functioning[201]. Generally, lower-functioning autism spectrum disorder is linked to less ADHD improvement and more adverse effects with both non-stimulants and stimulants compared to individuals with ADHD alone[201].

There is a dearth of studies of medications in adults with ADHD and prominent anxiety. Controlled data for atomoxetine in adults with ADHD and social anxiety demonstrated robust improvement in both conditions with tolerable adverse effects[202], supporting findings in youth[203]. A multisite RCT comparing paroxetine

and amphetamine, each alone and in combination, with placebo showed that paroxetine alone improved anxiety symptoms, being ineffective for ADHD[204]. In contrast, amphetamine alone worked somewhat for anxiety and worked well for ADHD, both in monotherapy and with paroxetine[204]. Reviews of stimulants in the context of anxiety in youth with ADHD have reported generally benign outcomes[205]. However, limited evidence in adults remains mixed [206,207].

Older data suggest that untreated mood symptoms result in lower ADHD response and more adverse events if using stimulants [208]. The aforementioned RCT comparing paroxetine and amphetamine indicated improvement in both depression and ADHD ratings only when paroxetine was combined with amphetamine[204]. Other trials have shown that the addition of stimulants to serotonin reuptake inhibitors and the use of bupropion monotherapy are effective in treating ADHD and depression[209-211].

Regarding bipolar disorder, data suggest that mood stabilization is essential prior to medicating ADHD[212]. Though treatment data in adults are lacking, RCTs in youth have shown successful treatment of ADHD without manic activation when amphetamine or methylphenidate is given together with mood stabilizing agents[213,214]. These clinical trial data have been supported by larger registry studies demonstrating the destabilizing tendency of stimulants over 6 months in adults with bipolar disorder not receiving mood stabilizing agents, whereas there was no manic activation when stimulants were combined with mood stabilizers[214].

Overall, there is a need for additional methodologically sound evidence to inform the treatment of ADHD comorbid with psychiatric disorders.

ADHD and substance use disorders

At present the literature is limited regarding how to best treat individuals with ADHD and a substance use disorder. Most studies that have focused on assessing the efficacy of medications for adult ADHD have either excluded or restricted those with past or current substance misuse. Given that most clinicians are more likely to feel comfortable using a non-stimulant in a patient with an active substance use, it is notable that there are so few trials assessing these medications, and that the results are mixed.

Prescribing stimulants to patients with ADHD and a substance use disorder still remains controversial, due to the risk of misuse and/or development of tolerance[219]. Extended-release formulations – particularly lisdexamfetamine and osmotic-release oral system formulations of methylphenidate (OROS-MPH) – are considered preferential over immediate-release preparations, as they contribute significantly less to the development of drug abuse or dependence[220]. A large Internet population survey[221] confirmed that extended-release methylphenidate and amphetamine formulations are less likely to be misused.

Safety concerns arise with prescribing of ADHD medication in the presence of ongoing substance use. In a meta-analysis evaluating stimulants and non-stimulants for treatment of ADHD with substance use disorders, there were no statistically significant differences in adverse events between those receiving active medication and those receiving placebo[222]. Also, all-cause treatment discontinuation was not different between the two groups. While elevations of blood pressure and heart rate have been noted in some clinical trials at a greater rate than those on placebo, untoward effects can be minimized when patients are monitored closely, and doses are adjusted or discontinued if necessary[223,224].

At present, if a patient has both ADHD and a substance use disorder, it is prudent to require the engagement of the patient in a specific treatment for the latter condition, particularly if a stimulant is being considered. In addition to using extended-release stimulant formulations, risk of misuse and diversion can be mitigated by limiting the number of pills provided, increasing the frequency of visits, monitoring for signs of misuse, discussing safe storage of medications, and obtaining urine toxicology if clinically indicated.

Naturalistic studies of medication effects

Significant questions remain regarding the representativeness of RCTs. Indeed, a recent study found that up to 70% of adults with ADHD from the Swedish registries would be ineligible for RCTs[225]. Moreover, RCTs provide limited evidence on serious and long-term outcomes such as injury and suicidality. Pharmacoepidemiology research is needed to collect better real-world evidence regarding the broader risks and benefits of ADHD medication.

Some pharmacoepidemiology studies have used a within-person design, comparing outcomes during medicated and unmedicated periods for the same individual. These studies found that, during periods on ADHD medication, patients had significantly fewer negative outcomes – such as unintentional injuries, motor vehicle accidents, substance use disorders, and criminal acts – and showed improvements in academic performance[22].

Another pharmacoepidemiology study[226] based on Swedish registries found that, among individuals diagnosed with ADHD, medication initiation was associated with significantly lower all-cause mortality, particularly for deaths due to unnatural causes.

Side effects of ADHD medications and their management

A balanced consideration of the side effects of medications is warranted. On the one hand, overlooking side effects may expose the patient to unwanted risks. On the other, excessive concerns about side effects may prevent a patient from receiving a potentially effective treatment, where side effects can be managed.

Table 4 presents a summary of the most significant adverse effects of ADHD medications, and of the recommendations for their management provided by the European ADHD Guidelines Group[227]. These recommendations were focused mainly on the use of ADHD medications in children. Although they can be extrapolated to adults, there is a need for specific guidance in this population.

Table 4 Adverse effects of medications for attention-deficit/hyperactivity disorder (ADHD) and recommendations for their management

Decreased appetite; height and weight gain deficit	• Measure height and weight every 6 months. • If weight loss is of clinical concern: take medication either with or after food, rather than before meals; take additional meals or snacks early in the morning or late in the evening when stimulant effects have worn off; obtain dietary advice; consume high-calorie foods of good nutritional value; take a planned break from treatment; or change medication.
Increased blood pressure/heart rate	• Do not offer routine blood tests or ECG unless there is a clinical indication. • Measure heart rate and blood pressure after each dose change and every 6 months. • If there is sustained resting tachycardia (>120 beats per minute), arrhythmia, or systolic blood pressure >95th percentile (or a clinically significant increase) measured on two occasions, reduce the dose and refer to hypertension specialist or adult physician. • If there are sustained orthostatic hypotension or fainting episodes during treatment with guanfacine, reduce the dose or switch to another ADHD medication.
Sleep disturbance	• Implement sleep hygiene. • If behavioral measures are insufficient and it is not convenient to stop medication, review the possible causes of sleep problems: a) treat restless legs syndrome if present; b) if there is rebound effect with stimulants, add small doses of short-acting stimulants in the evening; c) if stimulant is the current treatment, consider reducing dose, alternative classes or formulations of stimulants, or atomoxetine. • Consider adding melatonin.
Tics	• Monitor tics over a 3 months period before any decision regarding ADHD treatment. • If tics are stimulant-related, reduce the stimulant dose, or consider changing to atomoxetine or clonidine, or stopping medication, or add an antipsychotic.
Seizures	• If there are new seizures or worsening of existing seizures, review ADHD medication and stop any medication that might be contributing to the seizures. Cautiously reintroduce ADHD medication if it is unlikely to be the cause of the seizures.
Psychotic or manic symptoms	• If they occur with therapeutic doses of ADHD medications, reduce the dose or discontinue the ADHD drug. • Once the psychotic or manic symptoms resolve, consider a re-challenge with ADHD medications.

NON-PHARMACOLOGICAL TREATMENTS

Psychosocial interventions

Among psychosocial interventions, current guidelines recognize and recommend cognitive-behavioral therapy (CBT) in the treatment of adults with ADHD, either as a first-line approach in conjunction with medication, or as an alternative monotherapy when medication is not indicated for a particular patient[59,228]. However, CBT is not designed to treat the core symptoms of ADHD. Instead, its goal is to mitigate the impact of ADHD symptoms on the internal and external experiences of everyday life[229]. CBT for adults with ADHD involves a collaborative process between the clinician and the patient that can be delivered in a group, individual, or asynchronous online setting[230]. Protocols generally include the identification of environmental modifications or lifestyle changes that can support ADHD symptom management[229,231,232].

With respect to the "B" in CBT (i.e., behavioral components), adults with ADHD work with clinicians to generate coping strategies that mitigate ongoing impairments due to their ADHD symptoms (e.g., interpersonal challenges, work performance problems, difficulties managing finances, parenting). This may involve teaching executive functioning skills to support self-regulation (e.g., problem-solving, organization, time management) and planning supports to enhance preparation for everyday life situations.

With respect to the "C" in CBT (i.e., cognitive techniques), clinicians help clients identify and restructure maladaptive cognitions formed through years of negative feedback from others and subsequent self-blame (e.g., "I can't do anything right", "No one will want to be my friend", "I am inadequate"). These negative cognitions are thought to undermine self-esteem, motivation and hope in individuals with ADHD[229].

It is believed that by improving behavioral skills and thinking patterns related to oneself, the future and the world (i.e., the "cognitive triad"[233]), adults with ADHD will show greater self-regulation, which may in turn lead to an abatement of symptom severity[234]. However, the above-mentioned network meta-analysis of treatments for adult ADHD[21] showed that, in the short term (i.e., at time points close to 12 weeks), CBT was better than placebo only according to clinicians' assessments, but not to self-reported ratings of ADHD symptoms severity. Data from a very limited number of trials showed that, at about 52 weeks, CBT had supporting evidence of being more efficacious than placebo on self-reported ratings only. These discrepancies between rater assessments warrant further investigation. Furthermore, heterogeneity of ADHD symptom effects likely reflects high intervention heterogeneity with respect to content (emphasis on skills training vs. psychological change), dose/duration of care (brief vs. long-term interventions), format (group vs. individual), and involvement of family members[235].

Third-wave CBTs – including mindfulness and dialectical behavior therapy (DBT) – have also been applied for the treatment of ADHD in adults[236,237], although research evidence remains in its infancy[230]. In the above-mentioned network meta-analysis[21], mindfulness therapy was more efficacious than placebo in the short term (about 12 weeks) on clinicians' assessments but not self-ratings, and on self-report but not clinicians' ratings at 26 weeks.

Other psychosocial interventions widely marketed for adults

with ADHD in certain countries (e.g., ADHD coaching) are yet to be tested by RCTs, signalling a critical need for high-quality research on publicly available non-pharmacological interventions for ADHD in adults.

One practical issue in the application of psychosocial treatments to adult ADHD is the challenge of retaining patients in consistent clinical care. Across the lifespan, individuals with ADHD struggle to complete therapy homework assignments, demonstrate variable participation in activities during sessions, experience difficulties following through on intentions, and may show low or fluctuating self-efficacy and motivation to change behavior, which may lead to premature termination of treatment[238]. In children and adolescents, these challenges are often overcome by engaging parents as central participants in CBTs, often teaching them to reward youth engagement and participation. However, as adults with ADHD transition from a family-based to an individual care model, clinicians may struggle to cultivate client engagement.

To address engagement challenges, motivational interviewing techniques are increasingly being integrated into CBT models for adult ADHD[239]. These techniques may include exploring patients' self-awareness of their personal values and priorities, promoting personally meaningful therapy goals, and implementing a strength-based framework that bolsters self-efficacy[240]. Professionals may devote extra time in sessions to helping clients with ADHD develop specific and detailed implementation plans for intended actions, and use techniques such as advanced problem-solving to identify and address barriers to follow-through prior to their appearance[240].

Other non-pharmacological treatments

A recent pairwise meta-analysis[241] included RCTs of cognitive training across the lifespan (36 trials, eight in adults). Benefits for adults with ADHD were found only on laboratory measures of working memory (moderate effect size). There was no evidence of significant effects on ADHD core symptoms. Another pairwise meta-analysis[242], focusing on neurofeedback (38 RCTs, three in adults), could not find evidence of significant and clinically meaningful effects. While, to date, there is no support for the use of cognitive training and neurofeedback as treatment strategies for ADHD in adults, future studies may identify subgroups of patients who could benefit from these treatments.

SPECIAL GROUPS

Females

Until recently, ADHD in females has often been overlooked in clinical and research settings[243], as the disorder has historically been considered a male dominant one. Prevalence rates of ADHD favour males 3:1 in childhood, but this difference decreases during adolescence, and the rate in adulthood is nearly equal[244,245]. The diagnosis in females is often delayed compared to males[243,246,247], and females are more likely to be diagnosed later in life[248].

The reasons for the less frequent and later diagnosis of ADHD in females remain unclear. Some studies hypothesize that this is due to the difference in phenotypic expression of the disorder. Indeed, females often present with predominantly inattentive symptoms and less overt disruptive behaviors[102,249,250]. Therefore, male patients are more likely to be referred to clinical services[251]. However, recent evidence challenges these "traditional" views[252,253], suggesting that females with ADHD have comparable symptoms of hyperactivity to males. Nonetheless, other studies indicate that females must present with more severe symptoms with greater impairment compared to males to be referred for an ADHD assessment[254].

Another area of consideration is the notion of compensatory and masking behaviors reported more in women. This area of research is outlined in the autism literature, reporting that women, more often than men, utilize strategies to passively mask or hide their difficulties and actively compensate or adapt their behavior, particularly in social situations, which ultimately has detrimental long-term consequences[255,256]. There is limited evidence of this in adult ADHD[257]. Future research to explore the compensatory behaviors and masking techniques used by women with ADHD may improve diagnostic accuracy and help to reduce the harmful consequences of these strategies.

The prevalence of psychiatric comorbidities in females with ADHD has been reported to differ from males, with some studies showing that males are more likely to present with externalizing conditions (i.e., conduct disorders, substance misuse) and females having more internalizing comorbidities (i.e., anxiety and depression)[102,243,258]. This has been suggested to be a factor contributing to the lower rates of referral and diagnosis of ADHD for females compared to males, with depression and anxiety being diagnosed prior to ADHD, the so-called "diagnostic overshadowing"[249,250]. Self-harming behaviors, also more common in women with ADHD compared to males[259], have also been hypothesized to overshadow and distract from an ADHD diagnosis in women.

The recommendations for treatment of ADHD are the same in females as in males. However, females with ADHD are less likely to receive treatment with medication compared to males (independent of the severity of ADHD symptoms)[260]. A meta-analysis[261] confirmed this finding, but also showed that the difference in prescription frequency between males and females was less evident in adults compared to children.

It is unclear if there are sex-specific pharmacokinetics of ADHD medications, and whether the frequency and type of adverse events differ between males and females. Hormonal level fluctuations have been postulated to effect treatment response to stimulant medication[243,250]. A case series in a small sample of adult females (N=9)[262] reported a reduction in premenstrual worsening of depressive and ADHD symptoms when the current stimulant dose was increased. Currently clinical guidelines do not recommend different doses or treatment regimes according to sex, but developments in this area of research may lead eventually to sex-tailored treatments strategies.

Understanding the sex differences in adult ADHD is crucial

for improving timely diagnostic accuracy and clinical outcomes. There is a need for high-quality research including large numbers of female participants across different phases of life (from prepubertal all the way to post-menopausal).

Elderly

Only recently emerging research has demonstrated the presence of ADHD in adults over age 50[263,264]. In a systematic review and meta-analysis[265], the estimated prevalence of ADHD in older adults was 2.18% when diagnosed through validated scales in community samples and 0.23% when relying on clinical diagnoses in electronic health records. The prevalence of treatment for ADHD was 0.09%.

Diagnosing ADHD in older adults with cognitive complaints is complex, due to the variety of alternative diagnoses that must be considered[263], including traumatic brain injury, mild cognitive impairment, major depression, and cognitive symptoms due to medical illnesses or medications. Neuropsychological testing has been shown not to delineate ADHD from non-ADHD in this population[266].

Clinical features that can distinguish ADHD from other disorders are chronicity vs. variability over time of symptoms/impairments, age of onset (ADHD in childhood), temporal relationship to an event (traumatic brain injury, an infection, a new medication), and quality of cognitive symptoms (word finding/misspelling found in mild cognitive impairment, but not ADHD). It should be taken into account that other causes of cognitive impairment may coexist with ADHD.

There is a paucity of research concerning use of ADHD medications in older adults. A study assessing vital parameters in subjects aged 55-84 years receiving lisdexamfetamine[267] concluded that no trends in pulse and blood pressure were seen by age. However, in a minority of older adults, stimulant treatment may elevate blood pressure levels requiring clinical action. Older patients often have pre-existing somatic conditions and may be taking multiple medications concurrently[268]. The decision to treat older adults with ADHD medications involves balancing the potential improvement in quality of life, which is often substantial, against medical risks. Further research is clearly needed in this area.

ORGANIZATION OF SERVICES

Several studies have highlighted the lack of access to ADHD services as a significant issue in many countries[269-271]. In an ideal scenario, units specialized in the management of ADHD across the lifespan should serve as reference centers[269,272]. These units should provide a comprehensive assessment in complex cases and advanced treatment strategies in patients who do not respond to standard interventions. They should be coordinated with primary care physicians and community centers in order to ensure that all the aspects of patients' care are addressed.

The involvement of primary care physicians in the management of adult ADHD is essential[272]. Patients should have the opportunity of a follow-up of pharmacological treatment at the primary care level in coordination with their psychiatrist. Also, primary care physicians can improve the transition from community paediatricians to adult services. Another important role at this level of care is to screen for ADHD in patients with somatic conditions commonly associated with this disorder in adults, such as obesity, type 2 diabetes mellitus, migraine and epilepsy[269]. In general, it is necessary to improve training and education of primary care physicians on ADHD[272].

At the community mental health level, it is crucial that professionals (including psychiatrists, psychologists, nurses, and social workers) have experience in the assessment and treatment of ADHD in adults[269]. The service portfolio of community centers should regularly include the diagnostic assessment and management of adult ADHD and its psychiatric comorbidities, as well as the implementation of relevant psychological treatments[272].

The involvement of community addiction centers when adult ADHD is comorbid with a substance use disorder or behavioral addiction is pivotal[273]. It is not uncommon that the diagnosis of ADHD during childhood is missed in patients with addictions. Screening for ADHD should be included in the standard assessment of all patients with addictions, also considering that undiagnosed ADHD can impact negatively on the progression of the addictive disorder.

The transition from child community mental health services to adult services is a frequent problem in the management of ADHD in adults[274]. Specific programs are needed to improve this transition, especially in patients with comorbid conditions such as autism spectrum disorder, addictions or conduct disorders.

ADHD in adults can be a complex disorder, with multiple somatic and psychiatric comorbidities. The management of patients frequently requires the involvement of various specialists and health care levels. The coordination between professionals is essential to ensure continuity of care and avoid fragmentation[272]. Electronic health records are a very useful instrument to integrate services, share information among professionals, and optimize resources[270].

Digital health tools are playing an increasingly important role in the management of adults with ADHD[275]. They can increase access to ADHD specialists by people who have difficulties to attend in-person appointments because they live in remote areas or do not have specialized centers in their town[275].

There are several evidence-based guidelines with recommendations for assessment, treatment and monitoring of ADHD in adults[59,269]. The implementation of these guidelines at the different health care levels can improve the standard of care[272]. The care delivery in adults with ADHD needs to be person-centered, with consensus decision-making, to improve adherence to the management plan and ensure better outcomes[276]. The use of patient-reported outcome and experience measures (PROMS and PREMS) can help to empower patients.

The collaboration between professionals and patient organizations can increase awareness of adult ADHD and improve quality of care. Patient education is an integral component of management that can help improve the access to ADHD services[276].

PERSPECTIVES FROM ASSOCIATIONS OF PEOPLE WITH LIVED EXPERIENCE

Representatives of the two largest European and US associations of people with lived experience – N. Hovén (President of ADHD Europe) and J. Didier (President of Children and Adults with ADHD, CHADD) – have contributed to this section. We report verbatim their statements.

Diagnosis and access to treatment

"Access for adults to get in medical research is challenging. Adults come across a lot of understatements, even from professionals, if they have outwardly good functional capacity. That is, they have managed at work, are educated or their relationship matters are in order. Many times they are told that it can't be ADHD. Even 40-50-year-olds are told that they cannot get tested when they have survived without a diagnosis for so long. In all the discussion about overdiagnosis, it is forgotten that there is a great deal of underdiagnosis among adults. The experience of health care is that people want a diagnosis and, by applying correctly, seek it. Few people really want a diagnosis, but need it to get any support. The stigma is still associated with ADHD. Adults do not easily dare to talk about their diagnosis, for example. in the workplace, when they are afraid that attitudes towards them will change."

"As someone diagnosed with ADHD later in life, I often think about how much easier things could have been if I had known sooner. Like many women, especially those who excel outwardly in academic or professional settings, my struggles with ADHD went unnoticed by others. My unseen challenges were masked by perfectionism and an overwhelming drive to achieve. Teachers and peers praised my accomplishments, but my internal reality differed. I battled impulsivity, emotional dysregulation, and a constant sense of being 'on the edge' of chaos. When it finally came, my diagnosis provided clarity and validation, but only after years of self-doubt and self-criticism."

"This is a familiar story among adults with ADHD, particularly those diagnosed later in life. Our strengths – creativity, resourcefulness, intelligence – often camouflage the challenges, leaving us to grapple with the condition in isolation. These experiences underscore a broader issue: adults with ADHD are not just underserved – they are often invisible. Through my role as both a member of Children and Adults with ADHD (CHADD) and President of CHADD's Board of Directors, I've witnessed firsthand the profound gaps in awareness, diagnosis and treatment that adults with ADHD face and the transformative power of education, advocacy and support in addressing these needs."

"For many adults with ADHD, the journey to diagnosis is long and frustrating. Outdated stereotypes about who has ADHD – young, male, struggling academically – continue to exclude those who don't fit the mold. Women are often dismissed or misdiagnosed with anxiety or depression, while Black/Indigenous and low-income individuals face systemic biases that further delay care."

"I know what it's like to live in that gap. For years, I was told that I was 'too smart' to have ADHD or that I needed to 'try harder'. These well-meaning but harmful statements left me feeling like my struggles were a personal failing rather than a neurological difference. It wasn't until my diagnosis that I began to understand how ADHD shaped my life – and, more importantly, how to work with my brain rather than against it. This disconnect between how ADHD is perceived and how it manifests is one of the most significant barriers adults face."

Moving beyond medication: the need for comprehensive treatment

"Receiving an ADHD diagnosis is often framed as a solution, but for many, it's just the beginning of a complex process. While medication can be life-changing, it is rarely sufficient on its own. ADHD impacts nearly every aspect of life, from managing time and emotions to navigating relationships and careers. Many adults find that medication alone doesn't address their root challenges; so additional tools and support – like executive function coaching, peer support groups, disability accommodations, and occupational therapy – are needed."

"I've seen this firsthand, both personally and professionally. Cognitive behavioral therapy helped me reframe unhelpful thought patterns, while ADHD coaching and executive function skills training gave me practical strategies to manage my day-to-day life. These tools transformed how I approached everything, from prioritizing tasks to managing emotional overwhelm."

"Unfortunately, these helpful resources are not accessible to everyone. Many clinicians are not trained in ADHD-specific therapies, and insurance coverage for coaching or skills training is inconsistent at best. Even when these resources are available, individuals are often unaware of their existence. Few people diagnosed with ADHD receive education about how to create a comprehensive treatment plan, leaving them to navigate their condition with incomplete support. Organizations like CHADD help adults access resources and programs designed to fill this gap, connecting individuals with evidence-based strategies, webinars, peer support groups, and tools to help them better understand the complexities of ADHD."

The intersection of ADHD and substance use disorder

"One of the most critical yet under-addressed areas in ADHD care is the link with substance use disorder (SUD). Research shows that individuals with ADHD are significantly more likely to experience substance misuse, often to self-medicate for symptoms like impulsivity, emotional dysregulation, or restlessness."

"In my clinical work, I've seen how untreated ADHD can fuel cycles of addiction and relapse. Many individuals in recovery feel unsupported because traditional addiction programs rarely address the role ADHD plays in their behavior. Yet the research indicates that treating both ADHD and addiction helps people stay safer and sober longer. This gap in care perpetuates frustration and

prevents meaningful progress."

"*ADHD rarely travels alone. In fact, as many as 80% of adults with ADHD have at least one coexisting psychiatric disorder. The clinical community needs to continue advocating for integrated care that addresses both ADHD and co-occurring conditions like SUD. Clinicians can reshape how these conditions are treated together by promoting research, training professionals, and sharing resources tailored to these unique challenges."*

Family-centered care: a missing link

"*ADHD doesn't just affect individuals – it impacts entire families. When one person in a household is diagnosed, it often prompts a ripple effect of recognition and adjustment. Parents may realize they share similar traits, siblings may struggle to understand changing dynamics, and partners may face new challenges in communication and support."*

"*Despite this, family-centered care is rarely prioritized. Few treatment models include resources for loved ones, even though understanding ADHD as a shared experience can dramatically improve relationships and outcomes. Families need education, tools, and emotional support to navigate the complexities of ADHD together. My younger brother and sister were diagnosed with ADHD decades before I was diagnosed. How might things have been different if our entire family had been assessed back then? By addressing the whole family's needs, we can create environments where individuals with ADHD feel understood and supported at every stage."*

Addressing systemic inequities in ADHD care

"*For adults in marginalized communities, the barriers to ADHD care are even higher. Black/Indigenous and low-income individuals are often underdiagnosed or misdiagnosed, while women frequently have ADHD symptoms dismissed as stress or poor coping. These disparities perpetuate cycles of inequity, leaving many without the diagnosis and support they need to live well. These systemic issues also extend to justice-involved populations, where ADHD is disproportionately represented but rarely acknowledged and/or treated. Providing proper diagnosis and treatment in correctional settings could improve individual outcomes and reduce recidivism rates. This is a critical area for advocacy and reform."*

CONCLUDING REMARKS

Nowadays, after a period of criticism, ADHD is generally accepted as a valid nosological entity in adulthood. However, several uncertain or controversial aspects remain in the symptomatology, classification, epidemiology, comorbidity, etiology, pathophysiology, treatment, and service organization of the care for adults with ADHD, which we have highlighted in this paper. Therefore, there is a need for additional research on adult ADHD.

Notably, as of January 2023, data from the US National Institutes of Health (NIH) Reporter indicated just under $5.5 million in active funding for adult ADHD research, compared to over $42 million for paediatric ADHD research[277]. Remarkably, the NIH Reporter pointed out that funding for depression research exceeds ADHD research by at least tenfold, despite the two conditions having only slightly different population prevalence rates[277]. This highlights a critical need to expand ADHD research to develop effective public health strategies for identifying and treating its diverse presentations, using both pharmacological and non-pharmacological interventions.

Beyond the specific aspects that we discussed in each of the sections of this paper, there is an emerging potential change in ADHD conceptualization, under the influence of the neurodiversity movement. Neurodiversity, which originated as a social justice movement (rather than a clinical initiative), was initially proposed in relation to autism in the late 1990s by J. Singer[278], an Australian sociologist who identified as autistic herself. She introduced the term to describe the idea that neurological differences, such as autism, are part of natural human diversity rather than disorders to be cured. This concept is currently being extended to other nosologic entities, such as ADHD. From this perspective, impairments result not from intrinsic deficiencies but from a mismatch between the individual and a neurotypical environment. This mismatch can exacerbate challenges and undervalue the strengths of neurodiverse individuals, fostering stigma, shame, and mental health issues.

The neurodiversity movement emphasizes equality and highlights the unique strengths that neurodiverse people can contribute, such as creativity in ADHD or attention to detail in autism. Advocates encourage shifting the research and clinical focus from "fixing deficits" to understanding how environments and societal attitudes create barriers. This approach promotes adapting workplaces, schools and social settings to better suit neurodiverse needs, reducing stigma and discrimination through public education and policy changes.

Some authors in the field[279] have highlighted that, while extreme interpretations of neurodiversity that dismiss diagnosis and treatment should be avoided, integration of neurodiversity alongside traditional approaches should be explored. Combining interventions that address individual needs with societal efforts to accommodate diversity may offer a balanced path forward, enhancing both well-being and inclusion for neurodiverse individuals. We look forward to evidence-based and balanced discussions of these issues.

From a lived experience perspective, adults with ADHD need more than awareness – they need systems that actively support them. This means expanding access to affordable diagnosis and treatment, prioritizing family-centered care, and addressing co-occurring conditions such as substance use disorders through integrated models. It also means challenging outdated stereotypes and ensuring that marginalized communities have equitable access to care. Organizations of people with lived experience of ADHD play a vital role in this vision, advocating for systemic and policy-related changes while also creating peer support spaces where individuals and families can connect, learn and grow.

Ultimately, adults with ADHD require understanding, support, and the right tools to thrive. Continuing education, advocacy and collaboration are needed to create a world where adults with ADHD feel seen, supported and empowered.

ACKNOWLEDGEMENTS

S. Cortese is supported by the UK National Institute for Health and Care Research (NIHR) (grants nos. NIHR203684, NIHR203035, NIHR130077, NIHR128472 and RP-PG-0618-20003) and by the European Research Executive Agency (grant no. 101095568-HORIZONHLTH-2022-DISEASE-07-03). C.U. Correll, L.A. Rohde and S.V. Faraone are joint last authors of this paper.

REFERENCES

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 5th ed., text revision. Washington: American Psychiatric Publishing, 2022.
2. World Health Organization. International classification of diseases, 11th revision. Geneva: World Health Organization, 2022.
3. Faraone SV, Bellgrove MA, Brikell I et al. Attention-deficit/hyperactivity disorder. Nat Rev Dis Primers 2024;10:11.
4. McGough JJ, Barkley RA. Diagnostic controversies in adult attention deficit hyperactivity disorder. Am J Psychiatry 2004;161:1948-56.
5. Victor MM, da Silva BS, Kappel DB et al. Attention-deficit hyperactivity disorder in ancient Greece: the Obtuse Man of Theophrastus. Aust N Z J Psychiatry 2018;52:509-13.
6. Lange KW, Reichl S, Lange KM et al. The history of attention deficit hyperactivity disorder. Atten Defic Hyperact Disord 2010;2:241-55.
7. Still G. The Goulstonian lectures on some abnormal psychical conditions in children. Lancet 1902;159:1077-82.
8. Doyle R. The history of adult attention-deficit/hyperactivity disorder. Psychiatr Clin North Am 2004;27:203-14.
9. Morris HH Jr, Escoll PJ, Wexler R. Aggressive behavior disorders of childhood: a follow-up study. Am J Psychiatry 1956;112:991-7.
10. O'Neal P, Robins LN. The relation of childhood behavior problems to adult psychiatric status: a 30-year follow-up study of 150 subjects. Am J Psychiatry 1958;114:961-9.
11. Menkes MM, Rowe JS, Menkes JH. A twenty-five year follow-up study on the hyperkinetic child with minimal brain dysfunction. Pediatrics 1967;39:393-9.
12. Wood DR, Reimherr FW, Wender PH et al. Diagnosis and treatment of minimal brain dysfunction in adults: a preliminary report. Arch Gen Psychiatry 1976;33:1453-60.
13. Wender PH. Attention-deficit hyperactivity disorder in adults. New York: Oxford University Press, 1995.
14. Spencer T, Biederman J, Wilens TE et al. Adults with attention-deficit/hyperactivity disorder: a controversial diagnosis. J Clin Psychiatry 1998;59(Suppl. 7):59-68.
15. Solanto MV, Wasserstein J, Marks DJ et al. Diagnosis of ADHD in adults: what is the appropriate DSM-5 symptom threshold for hyperactivity-impulsivity? J Atten Disord 2012;16:631-4.
16. Moncrieff J, Timimi S. Is ADHD a valid diagnosis in adults? No. BMJ 2010; 340:c547.
17. Shaffer D. Attention deficit hyperactivity disorder in adults. Am J Psychiatry 1994;151:633-8.
18. Spencer T, Wilens T, Biederman J et al. A double-blind, crossover comparison of methylphenidate and placebo in adults with childhood-onset attention-deficit hyperactivity disorder. Arch Gen Psychiatry 1995;52:434-43.
19. Wender PH, Reimherr FW, Wood D et al. A controlled study of methylphenidate in the treatment of attention deficit disorder, residual type, in adults. Am J Psychiatry 1985;142:547-52.
20. Biederman J, Faraone SV, Spencer T et al. Patterns of psychiatric comorbidity, cognition, and psychosocial functioning in adults with attention deficit hyperactivity disorder. Am J Psychiatry 1993;150:1792-8.
21. Ostinelli EG, Schulze M, Zangani C et al. Comparative efficacy and acceptability of pharmacological, psychological, and neurostimulatory interventions for ADHD in adults: a systematic review and component network meta-analysis. Lancet Psychiatry 2025;12:32-43.
22. Chang Z, Ghirardi L, Quinn PD et al. Risks and benefits of attention-deficit/hyperactivity disorder medication on behavioral and neuropsychiatric outcomes: a qualitative review of pharmacoepidemiology studies using linked prescription databases. Biol Psychiatry 2019;86:335-43.
23. Asherson P, Buitelaar J, Faraone SV et al. Adult attention-deficit hyperactivity disorder: key conceptual issues. Lancet Psychiatry 2016;3:568-78.
24. Bauermeister JJ, Canino G, Polanczyk G et al. ADHD across cultures: is there evidence for a bidimensional organization of symptoms? J Clin Child Adolesc Psychol 2010;39:362-72.
25. Wagner F, Martel MM, Cogo-Moreira H et al. Attention-deficit/hyperactivity disorder dimensionality: the reliable 'g' and the elusive 's' dimensions. Eur Child Adolesc Psychiatry 2016;25:83-90.
26. Gomez R. ADHD and hyperkinetic disorder symptoms in Australian adults: descriptive scores, incidence rates, factor structure, and gender invariance. J Atten Disord 2016;20:325-34.
27. Willcutt EG, Nigg JT, Pennington BF et al. Validity of DSM-IV attention deficit/hyperactivity disorder symptom dimensions and subtypes. J Abnorm Psychol 2012;121:991-1010.
28. Wootton RE, Riglin L, Blakey R et al. Decline in attention-deficit hyperactivity disorder traits over the life course in the general population: trajectories across five population birth cohorts spanning ages 3 to 45 years. Int J Epidemiol 2022; 51:919-30.
29. Asherson P, Agnew-Blais J. Annual research review: Does late-onset attention-deficit/hyperactivity disorder exist? J Child Psychol Psychiatry 2019;60:333-52.
30. Soler-Gutiérrez AM, Pérez-González JC, Mayas J. Evidence of emotion dysregulation as a core symptom of adult ADHD: a systematic review. PLoS One 2023;18:e0280131.
31. Reimherr FW, Marchant BK, Olsen JL et al. Oppositional defiant disorder in adults with ADHD. J Atten Disord 2013;17:102-13.
32. Masi G, Sesso G, Pfanner C et al. An exploratory study of emotional dysregulation dimensions in youth with attention deficit hyperactivity disorder and/or bipolar spectrum disorders. Front Psychiatry 2021;12:619037.
33. Kessler RC, Green JG, Adler LA et al. Structure and diagnosis of adult attention-deficit/hyperactivity disorder: analysis of expanded symptom criteria from the Adult ADHD Clinical Diagnostic Scale. Arch Gen Psychiatry 2010; 67:1168-78.
34. Coghill DR, Seth S, Matthews K. A comprehensive assessment of memory, delay aversion, timing, inhibition, decision making and variability in attention deficit hyperactivity disorder: advancing beyond the three-pathway models. Psychol Med 2014;44:1989-2001.
35. Becker SP, Willcutt EG, Leopold DR et al. Report of a work group on sluggish cognitive tempo: key research directions and a consensus change in terminology to cognitive disengagement syndrome. J Am Acad Child Adolesc Psychiatry 2023;62:629-45.
36. Sibley MH, Rohde LA, Swanson JM et al. Late-onset ADHD reconsidered with comprehensive repeated assessments between ages 10 and 25. Am J Psychiatry 2018;175:140-9.
37. Breda V, Rovaris DL, Vitola ES et al. Does collateral retrospective information about childhood attention-deficit/hyperactivity disorder symptoms assist in the diagnosis of attention-deficit/hyperactivity disorder in adults? Findings from a large clinical sample. Aust N Z J Psychiatry 2016;50:557-65.
38. Moffitt TE, Houts R, Asherson P et al. Is adult ADHD a childhood-onset neurodevelopmental disorder? Evidence from a four-decade longitudinal cohort study. Am J Psychiatry 2015;172:967-77.
39. Sibley MH, Arnold LE, Swanson JM et al. Variable patterns of remission from ADHD in the Multimodal Treatment Study of ADHD. Am J Psychiatry 2022;179:142-51.
40. Kooij JJ, Michielsen M, Kruithof H et al. ADHD in old age: a review of the literature and proposal for assessment and treatment. Expert Rev Neurother 2016;16:1371-81.
41. London AS, Landes SD. Cohort change in the prevalence of ADHD among U.S. adults: evidence of a gender-specific historical period effect. J Atten Disord 2021;25:771-82.
42. Shi Y, Hunter Guevara LR, Dykhoff HJ et al. Racial disparities in diagnosis of attention-deficit/hyperactivity disorder in a US national birth cohort. JAMA Netw Open 2021;4:e210321.
43. Grevet EH, Bandeira CE, Vitola ES et al. The course of attention-deficit/hyperactivity disorder through midlife. Eur Arch Psychiatry Clin Neurosci 2024;274:59-70.
44. Karam RG, Rovaris DL, Breda V et al. Trajectories of attention-deficit/hyperactivity disorder dimensions in adults. Acta Psychiatr Scand 2017;136:210-9.
45. Caye A, Sibley MH, Swanson JM et al. Late-onset ADHD: understanding the evidence and building theoretical frameworks. Curr Psychiatry Rep 2017; 19:106.

46. Frick MA, Hesser H, Sonuga-Barke E. Pervasive versus situational childhood ADHD: latent classes and their clinical characteristics, based on parent and teacher ratings in a large longitudinal population sample. Eur Child Adolesc Psychiatry 2024;33:2253-66.
47. Rommelse N, Bunte T, Matthys W et al. Contextual variability of ADHD symptoms: embracement not erasement of a key moderating factor. Eur Child Adolesc Psychiatry 2015;24:1-4.
48. Chamberlain SR, Cortese S, Grant JE. Screening for adult ADHD using brief rating tools: what can we conclude from a positive screen? Some caveats. Compr Psychiatry 2021;106:152224.
49. First MB, Williams JBW, Karg RS et al. Structured Clinical Interview for DSM-5 Disorders, clinician version (SCID-5-CV). Arlington: American Psychiatric Association, 2017.
50. Caroline SS, Sudhir PM, Mehta UM et al. Assessing adult ADHD: an updated review of rating scales for adult attention deficit hyperactivity disorder (ADHD). J Atten Disord 2024;28:1045-62.
51. Ward MF, Wender PH, Reimherr FW. The Wender Utah Rating Scale: an aid in the retrospective diagnosis of childhood attention deficit hyperactivity disorder. Am J Psychiatry 1993;150:885-90.
52. Conners CK, Erhardt D, Sparrow E. Conners' Adult ADHD Rating Scales (CAARS). New York: Multihealth Systems, 1999.
53. Kessler RC, Adler L, Ames M et al. The World Health Organization Adult ADHD Self-Report Scale (ASRS): a short screening scale for use in the general population. Psychol Med 2005;35:245-56.
54. New York University and Kessler R. Adult ADHD Self-Report Scale-V1.1. 6-Question Screener. https://www.hcp.med.harvard.edu/ncs/ftpdir/adhd/6Q_ASRS_English.pdf.
55. Ustun B, Adler LA, Rudin C et al. The World Health Organization Adult Attention-Deficit/Hyperactivity Disorder Self-Report Screening Scale for DSM-5. JAMA Psychiatry 2017;74:520-7.
56. Mulraney M, Arrondo G, Musullulu H et al. Systematic review and meta-analysis: screening tools for attention-deficit/hyperactivity disorder in children and adolescents. J Am Acad Child Adolesc Psychiatry 2022;61:982-96.
57. Buitelaar J, Bölte S, Brandeis D et al. Toward precision medicine in ADHD. Front Behav Neurosci 2022;16:900981.
58. Cortese S, Solmi M, Michelini G et al. Candidate diagnostic biomarkers for neurodevelopmental disorders in children and adolescents: a systematic review. World Psychiatry 2023;22:129-49.
59. National Institute for Health and Care Excellence. Attention deficit hyperactivity disorder: diagnosis and management. www.nice.org.uk.
60. Bellato A, Hall CL, Groom MJ et al. Practitioner review: Clinical utility of the QbTest for the assessment and diagnosis of attention-deficit/hyperactivity disorder – a systematic review and meta-analysis. J Child Psychol Psychiatry 2024;65:845-61.
61. Arrondo G, Mulraney M, Iturmendi-Sabater I et al. Systematic review and meta-analysis: clinical utility of continuous performance tests for the identification of attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry 2024;63:154-71.
62. Hall CL, Valentine AZ, Walker GM et al. Study of user experience of an objective test (QbTest) to aid ADHD assessment and medication management: a multi-methods approach. BMC Psychiatry 2017;17:66.
63. Rohde LB, Buitelaar JK, Gerlach M et al. The World Federation of ADHD guide. Zurich: World Federation of ADHD, 2019.
64. Forbes MK, Neo B, Nezami OM et al. Elemental psychopathology: distilling constituent symptoms and patterns of repetition in the diagnostic criteria of the DSM-5. Psychol Med 2024;54:886-94.
65. Elwin M, Elvin T, Larsson JO. Symptoms and level of functioning related to comorbidity in children and adolescents with ADHD: a cross-sectional registry study. Child Adolesc Psychiatry Ment Health 2020;14:30.
66. Kessler RC, Adler L, Barkley R et al. The prevalence and correlates of adult ADHD in the United States: results from the National Comorbidity Survey Replication. Am J Psychiatry 2006;163:716-23.
67. Libutzki B, Ludwig S, May M et al. Direct medical costs of ADHD and its comorbid conditions on basis of a claims data analysis. Eur Psychiatry 2019;58:38-44.
68. Dalsgaard S, Østergaard SD, Leckman JF et al. Mortality in children, adolescents, and adults with attention deficit hyperactivity disorder: a nationwide cohort study. Lancet 2015;385:2190-6.
69. Galera C, Collet O, Orri M et al. Prospective associations between ADHD symptoms and physical conditions from early childhood to adolescence: a population-based longitudinal study. Lancet Child Adolesc Health 2023;7:863-74.
70. Hartman CA, Larsson H, Vos M et al. Anxiety, mood, and substance use disorders in adult men and women with and without attention-deficit/hyperactivity disorder: a substantive and methodological overview. Neurosci Biobehav Rev 2023;151:105209.
71. Antshel KM, Zhang-James Y, Wagner KE et al. An update on the comorbidity of ADHD and ASD: a focus on clinical management. Expert Rev Neurother 2016;16:279-93.
72. Schatz DB, Rostain AL. ADHD with comorbid anxiety: a review of the current literature. J Atten Disord 2006;10:141-9.
73. Jensen PS, Hinshaw SP, Kraemer HC et al. ADHD comorbidity findings from the MTA study: comparing comorbid subgroups. J Am Acad Child Adolesc Psychiatry 2001;40:147-58.
74. Solberg BS, Halmøy A, Engeland A et al. Gender differences in psychiatric comorbidity: a population-based study of 40 000 adults with attention deficit hyperactivity disorder. Acta Psychiatr Scand 2018;137:176-86.
75. Ottosen C, Larsen JT, Faraone SV et al. Sex differences in comorbidity patterns of attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry 2019;58:412-22.e3.
76. Cortese S, Moreira-Maia CR, St Fleur D et al. Association between ADHD and obesity: a systematic review and meta-analysis. Am J Psychiatry 2016;173:34-43.
77. Nigg JT, Johnstone JM, Musser ED et al. Attention-deficit/hyperactivity disorder (ADHD) and being overweight/obesity: new data and meta-analysis. Clin Psychol Rev 2016;43:67-79.
78. Arrondo G, Solmi M, Dragioti E et al. Associations between mental and physical conditions in children and adolescents: an umbrella review. Neurosci Biobehav Rev 2022;137:104662.
79. Ai Y, Zhao J, Liu H et al. The relationship between diabetes mellitus and attention deficit hyperactivity disorder: a systematic review and meta-analysis. Front Pediatr 2022;10:936813.
80. Libutzki B, Neukirch B, Reif A et al. Somatic burden of attention-deficit/hyperactivity disorder across the lifecourse. Acta Psychiatr Scand 2024;150:105-17.
81. Instanes JT, Klungsøyr K, Halmøy A et al. Adult ADHD and comorbid somatic disease: a systematic literature review. J Atten Disord 2018;22:203-28.
82. Du Rietz E, Brikell I, Butwicka A et al. Mapping phenotypic and aetiological associations between ADHD and physical conditions in adulthood in Sweden: a genetically informed register study. Lancet Psychiatry 2021;8:774-83.
83. Hegvik TA, Chen Q, Kuja-Halkola R et al. Familial co-aggregation of attention-deficit/hyperactivity disorder and autoimmune diseases: a cohort study based on Swedish population-wide registers. Int J Epidemiol 2022;51:898-909.
84. Li L, Chang Z, Sun J et al. Attention-deficit/hyperactivity disorder as a risk factor for cardiovascular diseases: a nationwide population-based cohort study. World Psychiatry 2022;21:452-9.
85. Garcia-Argibay M, Li L, Du Rietz E et al. The association between type 2 diabetes and attention-deficit/hyperactivity disorder: a systematic review, meta-analysis, and population-based sibling study. Neurosci Biobehav Rev 2023;147:105076.
86. Ahlberg R, Garcia-Argibay M, Taylor M et al. Prevalence of sleep disorder diagnoses and sleep medication prescriptions in individuals with ADHD across the lifespan: a Swedish nationwide register-based study. BMJ Ment Health 2023;26:e300809.
87. Cortese S, Sun S, Zhang J et al. Association between attention deficit hyperactivity disorder and asthma: a systematic review and meta-analysis and a Swedish population-based study. Lancet Psychiatry 2018;5:717-26.
88. Zhang J, Chen ZK, Triatin RD et al. Mediating pathways between attention deficit hyperactivity disorder and type 2 diabetes mellitus: evidence from a two-step and multivariable Mendelian randomization study. Epidemiol Psychiatr Sci 2024;33:e54.
89. Cortese S, Song M, Farhat LC et al. Incidence, prevalence, and global burden of ADHD from 1990 to 2019 across 204 countries: data, with critical re-analysis, from the Global Burden of Disease study. Mol Psychiatry 2023;28:4823-30.
90. Polanczyk G, de Lima MS, Horta BL et al. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. Am J Psychiatry 2007;164:942-8.
91. Polanczyk GV, Willcutt EG, Salum GA et al. ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. Int J Epidemiol 2014;43:434-42.
92. Solmi M, Radua J, Olivola M et al. Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. Mol Psychiatry 2022;27:281-95.
93. Faraone SV, Biederman J, Mick E. The age-dependent decline of attention deficit hyperactivity disorder: a meta-analysis of follow-up studies. Psychol Med 2006;36:159-65.
94. Vos M, Rommelse NNJ, Franke B et al. Characterizing the heterogeneous

course of inattention and hyperactivity-impulsivity from childhood to young adulthood. Eur Child Adolesc Psychiatry 2022;31:1-11.
95. Song P, Zha M, Yang Q et al. The prevalence of adult attention-deficit hyperactivity disorder: a global systematic review and meta-analysis. J Glob Health 2021;11:04009.
96. Biederman J, Petty CR, Clarke A et al. Predictors of persistent ADHD: an 11-year follow-up study. J Psychiatr Res 2011;45:150-5.
97. Biederman J, Petty CR, O'Connor KB et al. Predictors of persistence in girls with attention deficit hyperactivity disorder: results from an 11-year controlled follow-up study. Acta Psychiatr Scand 2012;125:147-56.
98. Kessler RC, Adler LA, Barkley R et al. Patterns and predictors of attention-deficit/hyperactivity disorder persistence into adulthood: results from the National Comorbidity Survey Replication. Biol Psychiatry 2005;57:1442-51.
99. Faraone SV, Biederman J. Can attention-deficit/hyperactivity disorder onset occur in adulthood? JAMA Psychiatry 2016;73:655-6.
100. Adeyemo BO, Biederman J, Zafonte R et al. Mild traumatic brain injury and ADHD: a systematic review of the literature and meta-analysis. J Atten Disord 2014;18:576-84.
101. Norman LJ, Price J, Ahn K et al. Longitudinal trajectories of childhood and adolescent attention deficit hyperactivity disorder diagnoses in three cohorts. EClinicalMedicine 2023;60:102021.
102. Staller J, Faraone SV. Attention-deficit hyperactivity disorder in girls: epidemiology and management. CNS Drugs 2006;20:107-23.
103. Bruchmüller K, Margraf J, Schneider S. Is ADHD diagnosed in accord with diagnostic criteria? Overdiagnosis and influence of client gender on diagnosis. J Consult Clin Psychol 2012;80:128-38.
104. Cortese S, Faraone SV, Bernardi S et al. Gender differences in adult attention-deficit/hyperactivity disorder: results from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). J Clin Psychiatry 2016;77:e421-8.
105. Cénat JM, Kokou-Kpolou CK, Blais-Rochette C et al. Prevalence of ADHD among Black youth compared to White, Latino and Asian youth: a meta-analysis. J Clin Child Adolesc Psychol 2024;53:373-88.
106. Xu G, Strathearn L, Liu B et al. Twenty-year trends in diagnosed attention-deficit/hyperactivity disorder among US children and adolescents, 1997-2016. JAMA Netw Open 2018;1:e181471.
107. Zablotsky B, Alford JM. Racial and ethnic differences in the prevalence of attention-deficit/hyperactivity disorder and learning disabilities among U.S. children aged 3-17 years. NCHS Data Brief 2020;358:1-8.
108. Morgan PL, Staff J, Hillemeier MM et al. Racial and ethnic disparities in ADHD diagnosis from kindergarten to eighth grade. Pediatrics 2013;132:85-93.
109. Slobodin O, Masalha R. Challenges in ADHD care for ethnic minority children: a review of the current literature. Transcult Psychiatry 2020;57:468-83.
110. Coker TR, Elliott MN, Toomey SL et al. Racial and ethnic disparities in ADHD diagnosis and treatment. Pediatrics 2016;138:e20160407.
111. Chan AYL, Ma TT, Lau WCY et al. Attention-deficit/hyperactivity disorder medication consumption in 64 countries and regions from 2015 to 2019: a longitudinal study. EClinicalMedicine 2023;58:101780.
112. Raman SR, Man KKC, Bahmanyar S et al. Trends in attention-deficit hyperactivity disorder medication use: a retrospective observational study using population-based databases. Lancet Psychiatry 2018;5:824-35.
113. Karlstad Ø, Zoëga H, Furu K et al. Use of drugs for ADHD among adults – a multinational study among 15.8 million adults in the Nordic countries. Eur J Clin Pharmacol 2016;72:1507-14.
114. Antshel KM, Faraone SV, Maglione K et al. Is adult attention deficit hyperactivity disorder a valid diagnosis in the presence of high IQ? Psychol Med 2009;39:1325-35.
115. Antshel KM, Faraone SV, Maglione K et al. Executive functioning in high-IQ adults with ADHD. Psychol Med 2010;40:1909-18.
116. Rommelse N, Antshel K, Smeets S et al. High intelligence and the risk of ADHD and other psychopathology. Br J Psychiatry 2017;211:359-64.
117. Rommelse N, van der Kruijs M, Damhuis J et al. An evidenced-based perspective on the validity of attention-deficit/hyperactivity disorder in the context of high intelligence. Neurosci Biobehav Rev 2016;71:21-47.
118. Goodman DW, Cortese S, Faraone SV. Why is ADHD so difficult to diagnose in older adults? Expert Rev Neurother 2024;24:941-4.
119. Doshi JA, Hodgkins P, Kahle J et al. Economic impact of childhood and adult attention-deficit/hyperactivity disorder in the United States. J Am Acad Child Adolesc Psychiatry 2012;51:990-1002.e2.
120. Hinnenthal JA, Perwien AR, Sterling KL. A comparison of service use and costs among adults with ADHD and adults with other chronic diseases. Psychiatr Serv 2005;56:1593-9.
121. Du Rietz E, Jangmo A, Kuja-Halkola R et al. Trajectories of healthcare utilization and costs of psychiatric and somatic multimorbidity in adults with childhood ADHD: a prospective register-based study. J Child Psychol Psychiatry 2020;61:959-68.
122. Garcia-Argibay M, Pandya E, Ahnemark E et al. Healthcare utilization and costs of psychiatric and somatic comorbidities associated with newly diagnosed adult ADHD. Acta Psychiatr Scand 2021;144:50-9.
123. Daley D, Jacobsen RH, Lange AM et al. The economic burden of adult attention deficit hyperactivity disorder: a sibling comparison cost analysis. Eur Psychiatry 2019;61:41-8.
124. Faraone SV, Larsson H. Genetics of attention deficit hyperactivity disorder. Mol Psychiatry 2019;24:562-75.
125. Chang Z, Lichtenstein P, Asherson PJ et al. Developmental twin study of attention problems: high heritabilities throughout development. JAMA Psychiatry 2013;70:311-8.
126. Larsson H, Chang Z, D'Onofrio BM et al. The heritability of clinically diagnosed attention deficit hyperactivity disorder across the lifespan. Psychol Med 2014;44:2223-9.
127. Brikell I, Kuja-Halkola R, Larsson H. Heritability of attention-deficit hyperactivity disorder in adults. Am J Med Genet B Neuropsychiatr Genet 2015;168:406-13.
128. Wechsler DL, Rijsdijk FV, Adamo N et al. Assessing aetiological overlap between child and adult attention-deficit hyperactivity disorder symptoms in an extended family design. BJPsych Open 2023;9:e169.
129. Taylor MJ, Martin J, Lu Y et al. Association of genetic risk factors for psychiatric disorders and traits of these disorders in a Swedish population twin sample. JAMA Psychiatry 2019;76:280-9.
130. Andersson A, Tuvblad C, Chen Q et al. Research review: The strength of the genetic overlap between ADHD and other psychiatric symptoms – a systematic review and meta-analysis. J Child Psychol Psychiatry 2020;61:1173-83.
131. Brikell I, Burton C, Mota NR et al. Insights into attention-deficit/hyperactivity disorder from recent genetic studies. Psychol Med 2021;51:2274-86.
132. Zhang L, Du Rietz E, Kuja-Halkola R et al. Attention-deficit/hyperactivity disorder and Alzheimer's disease and any dementia: a multi-generation cohort study in Sweden. Alzheimers Dement 2022;18:1155-63.
133. Demontis D, Walters GB, Athanasiadis G et al. Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains. Nat Genet 2023;55:198-208.
134. Rajagopal VM, Duan J, Vilar-Ribó L et al. Differences in the genetic architecture of common and rare variants in childhood, persistent and late-diagnosed attention-deficit hyperactivity disorder. Nat Genet 2022;54:1117-24.
135. Rovira P, Demontis D, Sánchez-Mora C et al. Shared genetic background between children and adults with attention deficit/hyperactivity disorder. Neuropsychopharmacology 2020;45:1617-26.
136. LaBianca S, Brikell I, Helenius D et al. Polygenic profiles define aspects of clinical heterogeneity in attention deficit hyperactivity disorder. Nat Genet 2024;56:234-44.
137. Breunig S, Lawrence JM, Foote IF et al. Examining differences in the genetic and functional architecture of attention-deficit/hyperactivity disorder diagnosed in childhood and adulthood. Biol Psychiatry Glob Open Sci 2024;4:100307.
138. Ronald A, de Bode N, Polderman TJC. Systematic review: how the attention-deficit/hyperactivity disorder polygenic risk score adds to our understanding of ADHD and associated traits. J Am Acad Child Adolesc Psychiatry 2021;60:1234-77.
139. Leppert B, Riglin L, Wootton RE et al. The effect of attention deficit/hyperactivity disorder on physical health outcomes: a 2-sample Mendelian randomization study. Am J Epidemiol 2021;190:1047-55.
140. Brikell I, Wimberley T, Albiñana C et al. Interplay of ADHD polygenic liability with birth-related, somatic, and psychosocial factors in ADHD: a nationwide study. Am J Psychiatry 2023;180:73-88.
141. Satterstrom FK, Walters RK, Singh T et al. Autism spectrum disorder and attention deficit hyperactivity disorder have a similar burden of rare protein-truncating variants. Nat Neurosci 2019;22:1961-5.
142. Thapar A, Martin J, Mick E et al. Psychiatric gene discoveries shape evidence on ADHD's biology. Mol Psychiatry 2016;21:1202-7.
143. Harich B, van der Voet M, Klein M et al. From rare copy number variants to biological processes in ADHD. Am J Psychiatry 2020;177:855-66.
144. Gudmundsson OO, Walters GB, Ingason A et al. Attention-deficit hyperactivity disorder shares copy number variant risk with schizophrenia and autism spectrum disorder. Transl Psychiatry 2019;9:258.
145. Meijer M, Franke B, Sandi C et al. Epigenome-wide DNA methylation in externalizing behaviours: a review and combined analysis. Neurosci Biobehav Rev 2023;145:104997.

146. van Dongen J, Zilhão NR, Sugden K et al. Epigenome-wide association study of attention-deficit/hyperactivity disorder symptoms in adults. Biol Psychiatry 2019;86:599-607.
147. Pujol-Gualdo N, Sánchez-Mora C, Ramos-Quiroga JA et al. Integrating genomics and transcriptomics: towards deciphering ADHD. Eur Neuropsychopharmacol 2021;44:1-13.
148. Mortimer N, Sánchez-Mora C, Rovira P et al. Transcriptome profiling in adult attention-deficit hyperactivity disorder. Eur Neuropsychopharmacol 2020; 41:160-6.
149. de Jong S, Newhouse SJ, Patel H et al. Immune signatures and disorder-specific patterns in a cross-disorder gene expression analysis. Br J Psychiatry 2016; 209:202-8.
150. Cabana-Domínguez J, Llonga N, Arribas L et al. Transcriptomic risk scores for attention deficit/hyperactivity disorder. Mol Psychiatry 2023;28:3493-502.
151. Deng MG, Zhou X, Li X et al. Identification of risk genes for attention-deficit/hyperactivity disorder during early human brain development. J Am Acad Child Adolesc Psychiatry 2024; doi: 10.1016/j.jaac.2024.10.013.
152. Liao C, Laporte AD, Spiegelman D et al. Transcriptome-wide association study of attention deficit hyperactivity disorder identifies associated genes and phenotypes. Nat Commun 2019;10:4450.
153. Qi X, Wang S, Zhang L et al. An integrative analysis of transcriptome-wide association study and mRNA expression profile identified candidate genes for attention-deficit/hyperactivity disorder. Psychiatry Res 2019;282:112639.
154. Kim JH, Kim JY, Lee J et al. Environmental risk factors, protective factors, and peripheral biomarkers for ADHD: an umbrella review. Lancet Psychiatry 2020; 7:955-70.
155. Carlsson T, Molander F, Taylor MJ et al. Early environmental risk factors for neurodevelopmental disorders – a systematic review of twin and sibling studies. Dev Psychopathol 2021;33:1448-95.
156. Jakobi B, Vlaming P, Mulder D et al. The gut-microbiome in adult attention-deficit/hyperactivity disorder – a meta-analysis. Eur Neuropsychopharmacol 2024;88:21-9.
157. Stuss DT, Murphy KJ, Binns MA et al. Staying on the job: the frontal lobes control individual performance variability. Brain 2003;126:2363-80.
158. Pievsky MA, McGrath RE. The neurocognitive profile of attention-deficit/hyperactivity disorder: a review of meta-analyses. Arch Clin Neuropsychol 2018;33:143-57.
159. Mowinckel AM, Pedersen ML, Eilertsen E et al. A meta-analysis of decision-making and attention in adults with ADHD. J Atten Disord 2015;19:355-67.
160. Alderson RM, Kasper LJ, Hudec KL et al. Attention-deficit/hyperactivity disorder (ADHD) and working memory in adults: a meta-analytic review. Neuropsychology 2013;27:287-302.
161. Bálint S, Czobor P, Komlósi S et al. Attention deficit hyperactivity disorder (ADHD): gender- and age-related differences in neurocognition. Psychol Med 2009;39:1337-45.
162. Boonstra AM, Oosterlaan J, Sergeant JA et al. Executive functioning in adult ADHD: a meta-analytic review. Psychol Med 2005;35:1097-108.
163. Schoechlin C, Engel RR. Neuropsychological performance in adult attention-deficit hyperactivity disorder: meta-analysis of empirical data. Arch Clin Neuropsychol 2005;20:727-44.
164. Skodzik T, Holling H, Pedersen A. Long-term memory performance in adult ADHD: a meta-analysis. J Atten Disord 2017;21:267-83.
165. Lei D, Du M, Wu M et al. Functional MRI reveals different response inhibition between adults and children with ADHD. Neuropsychology 2015;29:874-81.
166. Isfandnia F, El Masri S, Radua J et al. The effects of chronic administration of stimulant and non-stimulant medications on executive functions in ADHD: a systematic review and meta-analysis. Neurosci Biobehav Rev 2024;162: 105703.
167. Nigg JT, Willcutt EG, Doyle AE et al. Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? Biol Psychiatry 2005;57:1224-30.
168. Rosso G, Portaluppi C, Teobaldi E et al. Assessing adult ADHD through objective neuropsychological measures: a critical overview. J Atten Disord 2023; 27:786-94.
169. May T, Birch E, Chaves K et al. The Australian evidence-based clinical practice guideline for attention deficit hyperactivity disorder. Aust N Z J Psychiatry 2023; 57:1101-16.
170. Samea F, Soluki S, Nejati V et al. Brain alterations in children/adolescents with ADHD revisited: a neuroimaging meta-analysis of 96 structural and functional studies. Neurosci Biobehav Rev 2019;100:1-8.
171. Cortese S, Aoki YY, Itahashi T et al. Systematic review and meta-analysis: resting-state functional magnetic resonance imaging studies of attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry 2021;60:61-75.
172. Parlatini V, Itahashi T, Lee Y et al. White matter alterations in attention-deficit/hyperactivity disorder (ADHD): a systematic review of 129 diffusion imaging studies with meta-analysis. Mol Psychiatry 2023;28:4098-123.
173. Nakao T, Radua J, Rubia K et al. Gray matter volume abnormalities in ADHD: voxel-based meta-analysis exploring the effects of age and stimulant medication. Am J Psychiatry 2011;168:1154-63.
174. Frodl T, Skokauskas N. Meta-analysis of structural MRI studies in children and adults with attention deficit hyperactivity disorder indicates treatment effects. Acta Psychiatr Scand 2012;125:114-26.
175. Hoogman M, Bralten J, Hibar DP et al. Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. Lancet Psychiatry 2017;4:310-9.
176. Faraone SV, Banaschewski T, Coghill D et al. The World Federation of ADHD International Consensus Statement: 208 evidence-based conclusions about the disorder. Neurosci Biobehav Rev 2021;128:789-818.
177. Carucci S, Narducci C, Bazzoni M et al. Clinical characteristics, neuroimaging findings, and neuropsychological functioning in attention-deficit hyperactivity disorder: sex differences. J Neurosci Res 2023;101:704-17.
178. Parlatini V, Radua J, Solanes Font A et al. Poor response to methylphenidate is associated with a smaller dorsal attentive network in adult attention-deficit/hyperactivity disorder (ADHD). Transl Psychiatry 2023;13:303.
179. Parlartini V, Andrews DS, Pretzsch CM et al. Cortical alterations associated with lower response to methylphenidate in adults with ADHD. Nat Ment Health 2024;2:514-24.
180. Li T, van Rooij D, Roth Mota N et al. Characterizing neuroanatomic heterogeneity in people with and without ADHD based on subcortical brain volumes. J Child Psychol Psychiatry 2021;62:1140-9.
181. Shaw P, Malek M, Watson B et al. Trajectories of cerebral cortical development in childhood and adolescence and adult attention-deficit/hyperactivity disorder. Biol Psychiatry 2013;74:599-606.
182. Sonuga-Barke EJ, Castellanos FX. Spontaneous attentional fluctuations in impaired states and pathological conditions: a neurobiological hypothesis. Neurosci Biobehav Rev 2007;31:977-86.
183. Sutcubasi B, Metin B, Kurban MK et al. Resting-state network dysconnectivity in ADHD: a system-neuroscience-based meta-analysis. World J Biol Psychiatry 2020;21:662-72.
184. Cortese S. Pharmacologic treatment of attention deficit-hyperactivity disorder. N Engl J Med 2020;383:1050-6.
185. Faraone SV. The pharmacology of amphetamine and methylphenidate: relevance to the neurobiology of attention-deficit/hyperactivity disorder and other psychiatric comorbidities. Neurosci Biobehav Rev 2018;87:255-70.
186. Avelar AJ, Juliano SA, Garris PA. Amphetamine augments vesicular dopamine release in the dorsal and ventral striatum through different mechanisms. J Neurochem 2013;125:373-85.
187. Easton N, Steward C, Marshall F et al. Effects of amphetamine isomers, methylphenidate and atomoxetine on synaptosomal and synaptic vesicle accumulation and release of dopamine and noradrenaline in vitro in the rat brain. Neuropharmacology 2007;52:405-14.
188. Robinson JB. Stereoselectivity and isoenzyme selectivity of monoamine oxidase inhibitors. Enantiomers of amphetamine, N-methylamphetamine and deprenyl. Biochem Pharmacol 1985;34:4105-8.
189. Markowitz JS, DeVane CL, Ramamoorthy S et al. The psychostimulant d-threo-(R,R)-methylphenidate binds as an agonist to the 5HT(1A) receptor. Pharmazie 2009;64:123-5.
190. Maletic V, Mattingly GW, Earnest J. Viloxazine extended-release capsules as an emerging treatment for attention-deficit/hyperactivity disorder in children and adolescents. Expert Rev Neurother 2024;24:443-55.
191. Association of the Scientific Medical Societies in Germany. Interdisciplinary evidence- and consensus-based (S3) guideline "Attention Deficit/Hyperactivity Disorder in Children, Young People and Adults". Frankfurt: Association of the Scientific Medical Societies in Germany, 2018.
192. Canadian ADHD Resource Alliance. Canadian ADHD Practice Guidelines, 4th ed. Toronto: Canadian ADHD Resource Alliance, 2018.
193. American Professional Society of ADHD and Related Disorders. The first US guidelines for adult ADHD to be released by the American Professional Society of ADHD and Related Disorders (APSARD.org). Brentwood: American Professional Society of ADHD and Related Disorders, 2024.
194. Cortese S, Adamo N, Del Giovane C et al. Comparative efficacy and tolerability of medications for attention-deficit hyperactivity disorder in children, adolescents, and adults: a systematic review and network meta-analysis. Lancet Psychiatry 2018;5:727-38.
195. Bellato A, Perrott NJ, Marzulli L et al. Systematic review and meta-analysis:

effects of pharmacological treatment for attention-deficit/hyperactivity disorder on quality of life. J Am Acad Child Adolesc Psychiatry 2025;64:346-61.
196. National Institute for Health and Care Excellence. NICE British National Formulary (BNF). www.nice.org.uk/bnf-uk-only.
197. Farhat LC, Flores JM, Avila-Quintero VJ et al. Treatment outcomes with licensed and unlicensed stimulant doses for adults with attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. JAMA Psychiatry 2024;81:157-66.
198. Swanson JM. Debate: are stimulant medications for attention-deficit/hyperactivity disorder effective in the long term? (Against). J Am Acad Child Adolesc Psychiatry 2019;58:936-8.
199. Brams M, Weisler R, Findling RL et al. Maintenance of efficacy of lisdexamfetamine dimesylate in adults with attention-deficit/hyperactivity disorder: randomized withdrawal design. J Clin Psychiatry 2012;73:977-83.
200. Buitelaar JK, Trott GE, Hofecker M et al. Long-term efficacy and safety outcomes with OROS-MPH in adults with ADHD. Int J Neuropsychopharmacol 2012;15:1-13.
201. Rodrigues R, Lai MC, Beswick A et al. Practitioner Review: Pharmacological treatment of attention-deficit/hyperactivity disorder symptoms in children and youth with autism spectrum disorder: a systematic review and meta-analysis. J Child Psychol Psychiatry 2021;62:680-700.
202. Adler LA, Liebowitz M, Kronenberger W et al. Atomoxetine treatment in adults with attention-deficit/hyperactivity disorder and comorbid social anxiety disorder. Depress Anxiety 2009;26:212-21.
203. Geller D, Donnelly C, Lopez F et al. Atomoxetine treatment for pediatric patients with attention-deficit/hyperactivity disorder with comorbid anxiety disorder. J Am Acad Child Adolesc Psychiatry 2007;46:1119-27.
204. Weiss M, Hechtman L. A randomized double-blind trial of paroxetine and/or dextroamphetamine and problem-focused therapy for attention-deficit/hyperactivity disorder in adults. J Clin Psychiatry 2006;67:611-9.
205. Coughlin CG, Cohen SC, Mulqueen JM et al. Meta-analysis: reduced risk of anxiety with psychostimulant treatment in children with attention-deficit/hyperactivity disorder. J Child Adolesc Psychopharmacol 2015;25:611-7.
206. Tsai FJ, Tseng WL, Yang LK et al. Psychiatric comorbid patterns in adults with attention-deficit hyperactivity disorder: treatment effect and subtypes. PLoS One 2019;14:e0211873.
207. Gabriel A. The mixed amphetamine salt extended release (Adderall XR, Mas-XR) as an adjunctive to SSRIs or SNRIs in the treatment of adult ADHD patients with comorbid partially responsive generalized anxiety: an open-label study. Atten Defic Hyperact Disord 2010;2:87-92.
208. Wilens TE, Biederman J. The stimulants. Psychiatr Clin North Am 1992;15:191-222.
209. Daviss WB, Bentivoglio P, Racusin R et al. Bupropion sustained release in adolescents with comorbid attention-deficit/hyperactivity disorder and depression. J Am Acad Child Adolesc Psychiatry 2001;40:307-14.
210. Findling RL. Open-label treatment of comorbid depression and attentional disorders with co-administration of serotonin reuptake inhibitors and psychostimulants in children, adolescents, and adults: a case series. J Child Adolesc Psychopharmacol 1996;6:165-75.
211. Shim SH, Woo YS, Kim JS et al. Comparison between atomoxetine and OROS methylphenidate as an adjunctive to SSRIs in attention-deficit/hyperactivity disorder adults with comorbid partially responsive major depressive disorder: a head-to-head, 12-week, randomized, rater-blinded clinical trial. Clin Psychopharmacol Neurosci 2022;20:143-53.
212. Biederman J, Mick E, Prince J et al. Systematic chart review of the pharmacologic treatment of comorbid attention deficit hyperactivity disorder in youth with bipolar disorder. J Child Adolesc Psychopharmacol 1999;9:247-56.
213. Scheffer RE, Kowatch RA, Carmody T et al. Randomized, placebo-controlled trial of mixed amphetamine salts for symptoms of comorbid ADHD in pediatric bipolar disorder after mood stabilization with divalproex sodium. Am J Psychiatry 2005;162:58-64.
214. Findling RL, Short EJ, McNamara NK et al. Methylphenidate in the treatment of children and adolescents with bipolar disorder and attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry 2007;46:1445-53.
215. Wilens TE, Biederman J, Spencer TJ et al. Controlled trial of high doses of pemoline for adults with attention-deficit/hyperactivity disorder. J Clin Psychopharmacol 1999;19:257-64.
216. Spencer TJ, Adler LA, Weisler RH et al. Triple-bead mixed amphetamine salts (SPD465), a novel, enhanced extended-release amphetamine formulation for the treatment of adults with ADHD: a randomized, double-blind, multicenter, placebo-controlled study. J Clin Psychiatry 2008;69:1437-48.
217. Adler LA, Goldman R, Hopkins SC et al. Dasotraline in adults with attention deficit hyperactivity disorder: a placebo-controlled, fixed-dose trial. Int Clin Psychopharmacol 2021;36:117-25.
218. Adler LA, Leon TL, Sardoff TM et al. A placebo-controlled trial of lisdexamfetamine in the treatment of comorbid sluggish cognitive tempo and adult ADHD. J Clin Psychiatry 2021;82:20m13687.
219. US Food and Drug Administration. FDA updating warnings to improve safe use of prescription stimulants used to treat ADHD and other conditions. Serious risks with misuse, abuse, addiction, and sharing these drugs. www.fda.gov.
220. United Nations on Drug Use and Crime. Treatment of stimulant use disorders: current practices and promising perspectives. Discussion paper. Vienna: United Nations on Drug Use and Crime, 2019.
221. Cassidy TA, Varughese S, Russo L et al. Nonmedical use and diversion of ADHD stimulants among U.S. adults ages 18-49: a national internet survey. J Atten Disord 2015;19:630-40.
222. Cunill R, Castells X, Tobias A et al. Pharmacological treatment of attention deficit hyperactivity disorder with co-morbid drug dependence. J Psychopharmacol 2015;29:15-23.
223. Levin FR, Mariani JJ, Specker S et al. Extended-release mixed amphetamine salts vs placebo for comorbid adult attention-deficit/hyperactivity disorder and cocaine use disorder: a randomized clinical trial. JAMA Psychiatry 2015;72:593-602.
224. Konstenius M, Jayaram-Lindström N, Guterstam J et al. Methylphenidate for attention deficit hyperactivity disorder and drug relapse in criminal offenders with substance dependence: a 24-week randomized placebo-controlled trial. Addiction 2014;109:440-9.
225. Garcia-Argibay MC, Chang Z, Brikell I et al. Evaluating ADHD medication trial representativeness: a Swedish population-based study comparing hypothetically trial eligible and ineligible individuals. Lancet Psychiatry 2025;12:131-9.
226. Li L, Zhu N, Zhang L et al. ADHD pharmacotherapy and mortality in individuals with ADHD. JAMA 2024;331:850-60.
227. Cortese S, Holtmann M, Banaschewski T et al. Practitioner review: Current best practice in the management of adverse events during treatment with ADHD medications in children and adolescents. J Child Psychol Psychiatry 2013;54:227-46.
228. Australian ADHD Professionals Association. Australian evidence-based clinical guideline for attention deficit hyperactivity disorder (ADHD). Melbourne: Australian ADHD Professionals Association, 2022.
229. Ramsay J. Adult attention-deficit/hyperactivity disorder. In Wenzel A (ed). Handbook of cognitive behavioral therapy: applications. Worcester: American Psychological Association, 2021:389-421.
230. Nimmo-Smith V, Merwood A, Hank D et al. Non-pharmacological interventions for adult ADHD: a systematic review. Psychol Med 2020;50:529-41.
231. Safren SA, Sprich SE, Perlman CA et al. Mastering your adult ADHD: a cognitive-behavioral treatment program, therapist guide. Oxford: Oxford University Press, 2017.
232. Solanto M. Cognitive-behavioral therapy for adult ADHD: targeting executive dysfunction. New York: Guilford, 2011.
233. Newark PE, Stieglitz RD. Therapy-relevant factors in adult ADHD from a cognitive behavioural perspective. Atten Defic Hyperact Disord 2010;2:59-72.
234. Liu CI, Hua MH, Lu ML et al. Effectiveness of cognitive behavioural-based interventions for adults with attention-deficit/hyperactivity disorder extends beyond core symptoms: a meta-analysis of randomized controlled trials. Psychol Psychother 2023;96:543-59.
235. Knouse LE, Teller J, Brooks MA. Meta-analysis of cognitive-behavioral treatments for adult ADHD. J Consult Clin Psychol 2017;85:737-50.
236. Halmøy A, Ring AE, Gjestad R et al. Dialectical behavioral therapy-based group treatment versus treatment as usual for adults with attention-deficit hyperactivity disorder: a multicenter randomized controlled trial. BMC Psychiatry 2022;22:738.
237. Mitchell JT, McIntyre EM, English JS et al. A pilot trial of mindfulness meditation training for ADHD in adulthood: impact on core symptoms, executive functioning, and emotion dysregulation. J Atten Disord 2017;21:1105-20.
238. Sibley MH, Link K, Torres Antunez G et al. Engagement barriers to behavior therapy for adolescent ADHD. J Clin Child Adolesc Psychol 2023;52:834-49.
239. Meinzer MC, Oddo LE, Vasko JM et al. Motivational interviewing plus behavioral activation for alcohol misuse in college students with ADHD. Psychol Addict Behav 2021;35:803-16.
240. Sibley MH, Rodriguez LM, Lopez M et al. Operationalizing in-session treatment engagement strategies and behaviors for adolescents with ADHD and their parents. J Atten Disord 2025;29:336-50.
241. Westwood SJ, Parlatini V, Rubia K et al. Computerized cognitive training in attention-deficit/hyperactivity disorder (ADHD): a meta-analysis of randomized controlled trials with blinded and objective outcomes. Mol Psychiatry

2023;28:1402-14.
242. Westwood SJ, Aggesteiner PM, Kaiser A et al. Neurofeedback for attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. JAMA Psychiatry 2025;82:118-29.
243. Young S, Adamo N, Ásgeirsdóttir BB et al. Females with ADHD: an expert consensus statement taking a lifespan approach providing guidance for the identification and treatment of attention-deficit/ hyperactivity disorder in girls and women. BMC Psychiatry 2020;20:404.
244. Biederman J, Faraone SV, Monuteaux MC et al. Gender effects on attention-deficit/hyperactivity disorder in adults, revisited. Biol Psychiatry 2004;55:692-700.
245. Willcutt EG. The prevalence of DSM-IV attention-deficit/hyperactivity disorder: a meta-analytic review. Neurotherapeutics 2012;9:490-9.
246. Madsen KB, Ravn MH, Arnfred J et al. Characteristics of undiagnosed children with parent-reported ADHD behaviour. Eur Child Adolesc Psychiatry 2018;27:149-58.
247. Skoglund C, Sundström Poromaa I, Leksell D et al. Time after time: failure to identify and support females with ADHD – a Swedish population register study. J Child Psychol Psychiatry 2024;65:832-44.
248. Agnew-Blais JC, Polanczyk GV, Danese A et al. Evaluation of the persistence, remission, and emergence of attention-deficit/hyperactivity disorder in young adulthood. JAMA Psychiatry 2016;73:713-20.
249. Gaub M, Carlson CL. Gender differences in ADHD: a meta-analysis and critical review. J Am Acad Child Adolesc Psychiatry 1997;36:1036-45.
250. Quinn PO, Madhoo M. A review of attention-deficit/hyperactivity disorder in women and girls: uncovering this hidden diagnosis. Prim Care Companion CNS Disord 2014;16:PCC.13r01596.
251. Hinshaw SP, Owens EB, Zalecki C et al. Prospective follow-up of girls with attention-deficit/hyperactivity disorder into early adulthood: continuing impairment includes elevated risk for suicide attempts and self-injury. J Consult Clin Psychol 2012;80:1041-51.
252. Wettstein R, Navarro Ovando V, Pirgon E et al. Absent or hidden? Hyperactivity in females with ADHD. J Atten Disord 2024;28:1589-97.
253. Young S, Uysal O, Kahle J et al. A systematic review and meta-analysis comparing the severity of core symptoms of attention-deficit hyperactivity disorder in females and males. Psychol Med 2024;54:1-22.
254. Waschbusch DA, King S. Should sex-specific norms be used to assess attention-deficit/hyperactivity disorder or oppositional defiant disorder? J Consult Clin Psychol 2006;74:179-85.
255. Cook J, Hull L, Crane L et al. Camouflaging in autism: a systematic review. Clin Psychol Rev 2021;89:102080.
256. Livingston LA, Shah P, Milner V et al. Quantifying compensatory strategies in adults with and without diagnosed autism. Mol Autism 2020;11:15.
257. Ginapp CM, Greenberg NR, Macdonald-Gagnon G et al. The experiences of adults with ADHD in interpersonal relationships and online communities: a qualitative study. SSM Qual Res Health 2023;3:100223.
258. Hinshaw SP, Nguyen PT, O'Grady SM et al. Annual research review: Attention-deficit/hyperactivity disorder in girls and women: underrepresentation, longitudinal processes, and key directions. J Child Psychol Psychiatry 2022;63:484-96.
259. Beauchaine TP, Hinshaw SP, Bridge JA. Nonsuicidal self-injury and suicidal behaviors in girls: the case for targeted prevention in preadolescence. Clin Psychol Sci 2019;7:643-67.
260. Mowlem FD, Rosenqvist MA, Martin J et al. Sex differences in predicting ADHD clinical diagnosis and pharmacological treatment. Eur Child Adolesc Psychiatry 2019;28:481-9.
261. Kok FM, Groen Y, Fuermaier ABM et al. The female side of pharmacotherapy for ADHD – A systematic literature review. PLoS One 2020;15:e0239257.
262. de Jong M, Wynchank D, van Andel E et al. Female-specific pharmacotherapy in ADHD: premenstrual adjustment of psychostimulant dosage. Front Psychiatry 2023;14:1306194.
263. Goodman DW, Mitchell S, Rhodewalt L et al. Clinical presentation, diagnosis and treatment of attention-deficit hyperactivity disorder (ADHD) in older adults: a review of the evidence and its implications for clinical care. Drugs Aging 2016;33:27-36.
264. Michielsen M, Semeijn E, Comijs HC et al. Prevalence of attention-deficit hyperactivity disorder in older adults in The Netherlands. Br J Psychiatry 2012;201:298-305.
265. Dobrosavljevic M, Solares C, Cortese S et al. Prevalence of attention-deficit/hyperactivity disorder in older adults: a systematic review and meta-analysis. Neurosci Biobehav Rev 2020;118:282-9.
266. Ivanchak N, Abner EL, Carr SA et al. Attention-deficit/hyperactivity disorder in childhood is associated with cognitive test profiles in the geriatric population but not with mild cognitive impairment or Alzheimer's disease. J Aging Res 2011;2011:729801.
267. Ermer J, Haffey MB, Richards C et al. Double-blind, placebo-controlled, two-period, crossover trial to examine the pharmacokinetics of lisdexamfetamine dimesylate in healthy older adults. Neuropsychiatr Dis Treat 2013;9:219-29.
268. Lensing MB, Zeiner P, Sandvik L et al. Psychopharmacological treatment of ADHD in adults aged 50+: an empirical study. J Atten Disord 2015;19:380-9.
269. Kooij JJS, Bijlenga D, Salerno L et al. Updated European consensus statement on diagnosis and treatment of adult ADHD. Eur Psychiatry 2019;56:14-34.
270. Smith MCF, Mukherjee RAS, Müller-Sedgwick U et al. UK adult ADHD services in crisis. BJPsych Bull 2024;48:1-5.
271. Adamou M, Graham K, MacKeith J et al. Advancing services for adult ADHD: the development of the ADHD Star as a framework for multidisciplinary interventions. BMC Health Serv Res 2016;16:632.
272. Asherson P, Leaver L, Adamou M et al. Mainstreaming adult ADHD into primary care in the UK: guidance, practice, and best practice recommendations. BMC Psychiatry 2022;22:640.
273. Crunelle CL, van den Brink W, Moggi F et al. International consensus statement on screening, diagnosis and treatment of substance use disorder patients with comorbid attention deficit/hyperactivity disorder. Eur Addict Res 2018;24:43-51.
274. Buitelaar JK. Optimising treatment strategies for ADHD in adolescence to minimise 'lost in transition' to adulthood. Epidemiol Psychiatr Sci 2017;26:448-52.
275. Danielson ML, Claussen AH, Arifkhanova A et al. Who provides outpatient clinical care for adults with ADHD? Analysis of healthcare claims by types of providers among private insurance and Medicaid enrollees, 2021. J Atten Disord 2024;28:1225-35.
276. Culpepper L, Fried R. Attention-deficit/hyperactivity disorder in a chronic care paradigm. Postgrad Med 2013;125:78-86.
277. Sibley MH, Faraone SV, Nigg JT et al. Sudden increases in U.S. stimulant prescribing: alarming or not? J Atten Disord 2023;27:571-4.
278. Singer J. Odd people in: the birth of community amongst people on the "autistic spectrum". A personal exploration of a new social movement based on neurological diversity. Thesis, Faculty of Humanities and Social Science, University of Technology, Sydney, 1998.
279. Sonuga-Barke E, Thapar A. The neurodiversity concept: is it helpful for clinicians and scientists? Lancet Psychiatry 2021;8:559-61.

DOI:10.1002/wps.21374

ADHD in adults: despite evidence sufficient to guide diagnosis and treatment, many questions remain

Many readers will appreciate Cortese et al's review of attention-deficit/hyperactivity disorder (ADHD) in adulthood[1]. Considerable ground is covered, providing practical recommendations for health care professionals, such as using a provisional diagnosis when symptoms are present in only a single context, and how and when to make use of neuropsychological assessment. There is a thoughtful discussion of impairment and whether it should be based on general population performance benchmarks versus individual capacity. Representation of the views of people with lived experience of ADHD is noteworthy. Here I expand upon a few points that, in my opinion, require further discussion.

Concerning assessment, I would like to elaborate on the measurement of impulsivity in ADHD. The authors correctly report that studies have shown decline in the symptoms of impulsivity-hyperactivity in samples followed longitudinally. However, impulsivity is a complex, multifactorial construct[2], that is poorly reflected in the three relevant DSM-5 symptoms. For example, the UPPS-P Impulsive Behavior Scale[3] includes five domains of impulsivity, and most factors – in particular, negative and positive urgency, premeditation (lack of), and perseverance (lack of) – are elevated in individuals with ADHD[4]. Thus, including only three items in rating scales, even when adjusted for adults, might underestimate impulsivity in adults with ADHD. Future research and clinical care may address this possibility with expanded measurement.

On the topic of prevalence, individual and cultural differences in acceptance of mental health diagnosis and treatment approaches are important to consider. The lower prevalence of ADHD amongst minoritized groups (e.g., Black and Latine people in the US) is likely to be the result of reduced access to care. However, we should also consider that cultural (as well as spiritual and philosophical) differences also drive personal constructions of mental health[5]. It is important to recognize the myriad ways in which mental health is conceptualized across individuals and cultures as we expand diagnosis and treatment globally. Understanding the mental health trajectories of individuals with satisfying lives who otherwise would have met DSM-5 or ICD-11 diagnostic criteria for ADHD could be enlightening.

On a related note, clinicians routinely wrestle with determination of impairment when diagnosing ADHD in the absence of serious consequences such as failing grades in university or employment performance warnings. Distress from failure to match performance with perceived capacity, such as inability to take on expanded responsibilities in personal and work life, are common and can lead one clinician to diagnose ADHD while another clinician will disagree. This diagnostic challenge is also relevant for older people when it might be easy to dismiss ADHD as a cause of distress in the absence of employment. However, I recommend that we consider the diagnosis of depression as a comparator. In that case, we respect and assess personally experienced distress as a cause for recommending intervention – be it pharmacologic or not. This is also where collateral informant reports have value, and where skilled clinical interviewing can separate mild, situationally specific symptoms from chronic, truly impairing experiences.

Concerning neuropsychological testing, the authors clarify that it may be helpful for identifying an individual's cognitive strengths and weaknesses and perhaps for ruling out other causes of apparent ADHD symptoms such as traumatic brain injury or emerging dementia. Neuropsychological testing is not recommended as useful in ADHD diagnosis. Interestingly, neuropsychological test batteries generally include, in addition to assessments of general intellectual ability and achievement, tests of working memory, attention and executive function, all domains frequently impaired in adults with ADHD. When test performance is poor, resulting clinical recommendations may be similar to those provided by a therapist using cognitive-behavior therapy for ADHD to address the same deficits identified in clinical interview. An example is performing work in a low-distraction environment to improve sustained attention[6].

On the topic of ADHD course throughout life, more investigation is needed. A limited number of studies has shown that relatively few individuals with ADHD have a stable remitting course, but longitudinal research with successive measurements has only reached into the fifth decade of life[7]. In the Multimodal Treatment Study of Children with ADHD, rigorous prospective examination from childhood reached to a mean age of 25[8]. Research into older adulthood is critical to understand how ADHD intersects with the aging process cognitively and physically. In addition, taking life course fluctuations into account in studies of the neurobiology and genetics of ADHD is critical. Relying on single point-in-time assessments of ADHD risks missing important biologically determined differences between individuals with different lifespan courses.

The authors refer to the extensive literature reporting the adverse health outcomes associated with ADHD. Successful health self-management requires skills, planning, delay of gratification (resisting societally reinforced unhealthy food and drink consumption), and resources (e.g., access to healthy food). Therefore, it is not surprising that individuals with ADHD have poorer health outcomes, and simple provision of education is unlikely to drive improved outcomes. One key question for research is whether ADHD-related health risks are specific to this condition or reflect a process of physical health deterioration following mental health difficulties more broadly. Cortese et al touch on this latter possibility by noting the prevalence of co-occurring psychiatric conditions that may mediate the relation between ADHD and somatic conditions (e.g., alcohol use disorder leading to liver disease).

Finally, on the topic of treatment of ADHD in adulthood, the authors provide a helpful review of the status of the literature. My hope is that readers will not stop at the first sentence identifying pharmacological treatment as the cornerstone of management of

adult ADHD. Certainly, efficacy of this treatment has been demonstrated, and a balanced discussion of treatment options follows where readers will notice that, in addition to the demonstrated efficacy of medication, there are additional factors to consider when forecasting long-term treatment needs.

Now that we understand that ADHD is often a lifelong disorder, with variable expression, severity and impact, health care practitioners must integrate this understanding into their discussions with patients proactively. Although helpful, medication does not cure ADHD. Troubling non-symptom aspects of ADHD are often not satisfactorily improved with medication and are likely to require periodic and/or sustained support with psychotherapy or other interventions throughout life, in accordance with individual need. In this respect, it is surprising that research on ADHD coaching remains in its infancy.

In a population-based cohort study in the Netherlands, less than 20% of adults followed over 25 years were able to maintain healthy lifestyles across five critical behaviors (physical activity, body weight, smoking, sleep, and alcohol consumption)[9]. If maintaining a healthy lifestyle is hard for the general population, it will be especially difficult for individuals with ADHD to follow recommended behavioral changes. Successful long-term care may need to include supports such as case managers embedded within routine clinical care settings to help patients maintain motivation and access available resources.

We have amassed sufficient literature and clinical experience to know that ADHD in adulthood exists. We should now turn our attention to the many questions, such as understanding lifespan trajectories, to inform our next steps in research and clinical care.

Brooke S.G. Molina
Departments of Psychiatry, Psychology, Pediatrics, Clinical and Translational Science, University of Pittsburgh, Pittsburgh, PA, USA

The author is partially supported by the US National Institute on Drug (grant no. DA049721) and the US National Institute on Alcohol Abuse and Alcoholism (grant no. AA027494).

1. Cortese S, Bellgrove MA, Brikell I et al. World Psychiatry 2025;24:347-71.
2. Dick DM, Smith G, Olausson P et al. Addict Biol 2010;15:217-26.
3. Lynam DR, Smith GT, Whiteside SP et al. The UPPS-P: assessing five personality pathways to impulsive behavior. West Lafayette: Purdue University, 2006.
4. Pedersen SL, Walther CA, Harty SC et al. Addiction 2016;111:1582-9.
5. Gopalkrishnan N, Babacan H. Australas Psychiatry 2015;23(Suppl. 6):6-8.
6. Solanto MV. Cognitive-behavioral therapy for adult ADHD: targeting executive dysfunction. New York: Guilford, 2011.
7. Grevet EH, Bandeira CE, Vitola ES et al. Eur Arch Psychiatry Clin Neurosci 2024;274:59-70.
8. Sibley MH, Arnold LE, Swanson JM et al. Am J Psychiatry 2022;179:142-51.
9. Schermer EE, Engelfriet PM, Blokstra A et al. Front Public Health 2022;10: 966155.

DOI:10.1002/wps.21344

Accurate assessment of adult ADHD: a key to better outcomes?

Cortese et al[1] provide a thoughtful overview of the scientific and clinical aspects of attention-deficit/hyperactivity disorder (ADHD) in adults. This is a field that has developed rapidly over the past few years, moving from the margins of most people's perception and awareness into the limelight. While this increased recognition of and attention to the needs of adults with ADHD is welcome, it has brought with it some additional demands and challenges. This is highlighted by the tension between the debate, played out publicly in the media, about whether ADHD is being "overdiagnosed", and the message that adults with ADHD are having great difficulty accessing adequate services[2].

With reference to adult ADHD, it is more helpful to talk about missed diagnosis and misdiagnosis[3]. It is almost certainly the case that in many countries we are currently seeing a combination of the two. While there are still many adults with ADHD who do not receive a diagnosis or treatment, there are also people who do not meet the criteria for ADHD but have been given a diagnosis and are receiving treatment, most commonly with stimulant medication.

The most common route to this is a quick and poorly conducted assessment. By continuing to miss a diagnosis of ADHD, we are also missing the opportunity to facilitate access to the available effective pharmacological and non-pharmacological treatment. But by diagnosing and treating people who do not meet criteria, we are changing the meaning of ADHD and moving beyond the evidence base. I understand the pressures on clinicians, particularly in private practice, when asked to conduct an "ADHD assessment" by someone who for whatever reason and by whatever route has already identified strongly or perhaps "self-diagnosed", but we must remain objective and assess properly.

Indeed, one of the most important issues facing clinicians working in this area is to ensure that assessments are accurate and comprehensive. This does not seem to always be the case. There have been many recent reports of services popping up that are offering rapid access to "specialist ADHD assessments", often over telehealth and with a hefty price tag. Feedback from people who have been through such processes is that the assessment was short, often a single appointment of around 30 min, relied heavily on questionnaires rather than clinical interview, and did not include a developmental history, or a screen for other physical or mental health conditions, or the collection of collateral information from sources other than the person being assessed. A positive diagnosis is almost always made. While medication treatment is usually recommended, this is not offered as part of the service. Instead, a referral is made to a primary care physician who is asked to start, titrate and monitor medication and outcomes.

If this is not good practice, how should we be working? I strongly advise that, as recommended in evidence-based guidelines, a structured clinical interview should always form the core of a clinical assessment[4]. As emphasized by Cortese et al, our work on screening questionnaires in children and young people highlighted problems with specificity, and are therefore associated with high rates

of false positives[5]. While we have not yet completed our review of screeners in adults, it is likely that we will find a similar result. Why are current approaches to screening so disappointing? One major factor is that they focus on symptoms and do not assess impairment. Perhaps in the future we will be able to harness multistage screening and/or artificial intelligence to help with both screening and assessment, but this is not yet the case.

Another aspect of assessment that I would take a hard line on is the use of cognitive testing, including continuous performance tasks (CPTs) such as QbTest, to aid diagnosis. At the cognitive level, ADHD is highly heterogeneous. While people with this condition have differences across a broad range of cognitive functions, there is no definitive ADHD cognitive profile[6]. No cognitive deficit is shared by everyone with ADHD, and no cognitive deficit is unique to ADHD. Just because one performs poorly on one or more task(s), including QbTest, that does not mean that he/she has ADHD. On the other hand, when someone does not show problems with executive functioning or on a CPT, this does not rule out ADHD. Cognitive testing can help you understand a person's strengths and weaknesses, but it does not aid with diagnosis[7].

This also helps to answer another question posed by Cortese et al: whether executive functioning should be considered a core feature of ADHD. The answer depends on how one is defining this cognitive domain. Performance on neuropsychological tests of executive functioning as part of a formal assessment is clearly not decisive. Less than half of those with ADHD perform poorly. On the other hand, scoring highly on the Behavior Rating Inventory of Executive Function (BRIEF) is very common. However, it is important to note that, while the BRIEF scores correlate very highly with ADHD symptom measures, their correlation with recognized tests of executive functioning is usually non-significant. I firmly believe that executive functioning difficulties should be listed among the associated features of ADHD rather than regarded as a core feature.

One area where we clearly need more evidence is for females with ADHD. As highlighted by Cortese et al, the male to female ratio changes across development, with a preponderance of males in childhood but equality in adults. What is not clear is how and why this happens. Are more male children with ADHD remitting before adulthood? Are some females who are subsyndromal during childhood developing full ADHD as they reach adulthood? Is it a combination of these, or is there another explanation?

A close examination of data from the most recent epidemiological Australian survey of child and adolescent mental health is intriguing. For children aged 6 to 12 years, the prevalence of ADHD is 10.9% in males and 5.4% in females, a ratio of around 2:1. For adolescents aged 13 to 18 years, the prevalence in males is similar (9.8%), but the prevalence in females drops by around 50% to 2.7%, giving a male to female ratio of 3.6:1[8]. This was a well-conducted population-based study. The assessment process, a structured research diagnostic interview, was similar across the age range. I believe that the results of this study pose serious questions about the validity and reliability of our clinical assessments for adolescent females. It seems very unlikely that there is an increase in the gap between males and females during adolescence and then an equalling out in adulthood. Is this problem with assessment limited to adolescents, or does it continue into adulthood? I am not suggesting here that we should have different criteria for males and females, although perhaps this argument could be made. What I argue is that we need to be sure that we are accurately identifying symptoms in females and applying the same standards across sexes and throughout development.

There are, of course, many other areas relevant to adult ADHD in which we could improve our performance. However, if we can at least get the assessment process right and make sure that we are diagnosing the right people, this will be a very significant first step.

David Coghill
Departments of Paediatrics and Psychiatry, University of Melbourne, Melbourne, VIC, Australia

1. Cortese S, Bellgrove MA, Brikell I et al. World Psychiatry 2025;24:347-71.
2. Community Affairs References Committee. Assessment and support services for people with ADHD. Canberra: Commonwealth of Australia, 2023.
3. Gyngell C, Payne JM, Coghill D. Lancet Psychiatry 2023;10:658-60.
4. May T, Birch E, Chaves K et al. Aust N Z J Psychiatry 2023;57:1101-16.
5. Mulraney M, Arrondo G, Musullulu H et al. J Am Acad Child Adolesc Psychiatry 2022;61:982-96.
6. Coghill DR, Seth S, Matthews K. Psychol Med 2014;44:1989-2001.
7. Coghill D. J Am Acad Child Adolesc Psychiatry 2021;60:1461-3.
8. Lawrence D, Hafekost J, Johnson SE et al. Aust N Z J Psychiatry 2016;50:876-86.

DOI:10.1002/wps.21345

The emotional side of adult ADHD

In their scholarly paper on the knowns and unknowns of adult attention-deficit/hyperactivity disorder (ADHD), Cortese et al[1] address the multifaceted aspects of this prevalent, early-onset, and often life-spanning disorder, including the role of emotional dysregulation and psychiatric comorbidities.

Large scale meta-analyses, investigating a six- to seven-digit number of patients, clearly demonstrate that comorbidities include major depressive disorder (MDD)[2] and bipolar disorder (BD)[3], with an odds ratio of 4.5 and 8.7, respectively[2]. These epidemiological data are corroborated by family-based studies on relatives of patients with ADHD, but also of patients with either MDD or BD. However, there is comparably little research on how mood disorders are linked to ADHD – and even lesser data on how this comorbidity should be treated.

There is significant shared heritability between ADHD, MDD and BD[4]. However – even though there are multiple common genetic risk variants, especially between ADHD and MDD, and Mendelian randomization studies further argue that the genetic liability for ADHD is causally related to MDD – effect sizes are too small to account for the considerable comorbidity. So, what else might play

a role here?

First, almost every environmental risk factor for MDD is overrepresented in ADHD. Childhood adversity has been present in many cases. This includes disturbed parental bonding (which may be due to the child's disruptive behaviors as well as to possible ADHD in a parent), repeated social rejection, being subjected to bullying and educational failure, as well as physical and emotional trauma. Also later in life, people with ADHD are more prone to traumatic life events. Furthermore, lower socioeconomic status, unstable relationships, and a plethora of other negative life events all add up to increase the risk for later depression. Comorbidities such as obesity and substance use further contribute to MDD risk. Adult ADHD-MDD comorbidity is thus a prime example for the bio-psycho-social model of the pathogenesis of mental disorders, with genetic as well as environmental risk factors adding up to finally increase the risk above threshold.

There are only few studies on the phenotype of depression in the context of adult ADHD. ADHD-MDD seems to be associated with an earlier onset of depression, a higher disease burden (e.g., more hospitalizations and a higher number of episodes), increased suicidality, higher functional impairment, lower quality of life, and a higher risk for treatment-resistant depression (TRD). Notably, ADHD polygenic risk scores are associated with TRD. According to the BRIDGE-II-MIX study[5], ADHD-MDD goes along with a higher number of (hypo)manic symptoms in MDD, a higher prevalence of mixed and atypical depression, a positive family history for (hypo)mania, and a history of manic switch upon antidepressant treatment. Thus, the ADHD-MDD phenotype may look like bipolar depression. Therefore, when it comes to clinical assessment, this requires diagnostic rigor and knowledge about the connections between these disorders.

An up to 30-year follow-up study of MDD patients[6] found the conversion risk of MDD to bipolar disorder to be 26%. However, when three or more subthreshold hypomanic symptoms were present at baseline, this proportion increased to more than 45%. As indicated above, ADHD-MDD with (hypo)manic symptoms is precisely the phenotype going along with a higher risk to conversion into BD. Bringing these data together, a clinical trajectory from ADHD to ADHD-MDD with bipolar features, then converting into ADHD-BD, seems conceivable. This course of disease could actually be described in longitudinal studies on offspring of patients with BD, especially those who were lithium non-responders[7]. This adds to the concept of a more episodic, lithium-responsive subtype of BD, as opposed to a more chronic, lithium-non-responsive subtype, where chronicity might be due to underlying ADHD.

As pointed out by Cortese et al, emotional dysregulation – usually defined as the inability to adequately control emotions, resulting in frequent, prolonged and abnormally intense emotional states – goes along with childhood as well as adulthood ADHD. Whether, and how, emotional dysregulation is linked to full-blown mood episodes is, however, unknown. The currently ongoing DynAMoND study[8], based on dense ecological momentary assessment sampling of affect and arousal in ADHD and BD, will hopefully shed light on this issue.

The therapeutic implications of ADHD-MDD and ADHD-BD remain largely unknown, and somewhat rely on which disorder has been diagnosed first. Most clinical guidance suggests that, if MDD is diagnosed together with ADHD, it should be treated first and according to pertinent guidelines. Upon remission or at least response, the impairment of comorbid ADHD should be evaluated and, if still relevant, ADHD should be treated using first-line medication, i.e. stimulants.

However, this sequence is not empirically grounded. It might well be that simultaneous treatment with antidepressant medication and stimulants increases remission rates, given that ADHD is a risk factor for TRD. Further, the summary of product characteristics of stimulants usually conveys a warning, or contraindication, regarding stimulant use when MDD is, or has been, present. This warning is mostly theoretical in nature, and clinical experience suggests that stimulant use is safe in ADHD-MDD. However, randomized clinical trials are urgently needed to address this issue, given that about 1-2% of the overall adult population at least once in their lifetime suffer from ADHD-MDD. Recommendations to preferentially use antidepressants with a dopaminergic or noradrenergic mechanism of action in ADHD-MDD rely on valid pharmaco-theoretic reasoning, yet they have no empirical underpinning. Again, such advice should be tested in controlled trials.

When it comes to ADHD-BD, there is the frequent concern that stimulant use in BD might trigger manic episodes. While this might indeed be the case in the absence of mood-stabilizing agents, it has been shown[9] that a combination of mood stabilizers (e.g., lithium) and stimulants is safe with respect to switch risk. Thus, if ADHD causes impairment outside of mood episodes in ADHD-BD, it should be treated accordingly, which might also improve adherence to mood-stabilizing therapy.

The role of psychotherapy and neurostimulation has not yet systematically assessed in either ADHD-MDD or ADHD-BD. Certainly, psychoeducation is of high relevance in both conditions.

In conclusion, the connections between ADHD, emotional dysregulation, and mood disorders are manifold, and many important questions are still unanswered. More studies on these highly prevalent comorbidities are clearly needed to improve patients' life. This requires clinicians and researchers to leave their "diagnostic silos" to fully appreciate the complexities of mental disorders.

Andreas Reif
Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Medical Center Frankfurt, Frankfurt, Germany; Fraunhofer Institute for Translational Medicine and Pharmacology, Frankfurt, Germany

1. Cortese S, Bellgrove MA, Brikell I et al. World Psychiatry 2025;24:347-71.
2. Hartmann CA, Larsson H, Vos M et al. Neurosci Biobehav Rev 2023;151:105209.
3. Schiweck C, Arteaga-Henriquez G, Aichholzer M et al. Neurosci Biobehav Rev 2021;124:100-23.
4. Hindley G, Frei O, Shadrin AA et al. Am J Psychiatry 2022;179:833-43.
5. Vannucchi G, Medda P, Pallucchini A et al. J Affect Disord 2019;246:346-54.
6. Fiedorowicz JG, Endicott J, Leon AC et al. Am J Psychiatry 2011;168:40-8.
7. Duffy A, Goodday S, Keown-Stoneman C et al. Am J Psychiatry 2019;176:720-9.
8. Bayas M, Kockler TD, Ramos-Quiroga JA et al. Int J Bipolar Disord 2024;12:44.
9. Viktorin A, Ryden E, Thase ME et al. Am J Psychiatry 2017;174:341-8.

DOI:10.1002/wps.21346

What are the long-term outcomes of ADHD treatment?

As outlined by Cortese et al[1], there has been considerable progress in the understanding of the epidemiology, genetic risk, clinical features and management of adult attention-deficit/hyperactivity disorder (ADHD). However, as this condition is increasingly being recognized and treated in millions of children, adolescents and adults worldwide, the remaining knowledge gaps also become more apparent, and the need to address these gaps more urgent.

In my meetings with patients with ADHD and their relatives, patient organizations and health care professionals, certain questions and concerns keep coming up: How will my future look like? What are the consequences of having ADHD and, in particular, what is the long-term outcome of ADHD treatment?

Although many different interventions have been developed for ADHD, the most effective and by far the most studied treatment is that with stimulant drugs, such as amphetamines and methylphenidate. Like many other important discoveries, the beneficial effects of stimulants can be traced back to accidental observations, including that Benzedrine (amphetamine sulfate) had a calming effect in children with behavioral disorders[2]. Since the systematic testing of stimulants in ADHD started in the 1960s[3], it has been demonstrated that they provide immediate symptomatic relief of ADHD symptoms such as inattention, impulsivity and hyperactivity[3].

Recent meta-analyses confirm that stimulants and atomoxetine have significant effects in reducing ADHD symptoms after 12 weeks of treatment in children and adults[4]. However, as pointed out by Cortese et al, "at about 52 weeks, no medication has supporting evidence of being more efficacious than placebo". A conservative interpretation may be that initial treatment effects disappear after one year of treatment. However, this conclusion is inconsistent with clinical observations, naturalistic treatment studies and registry-based research. Several open-label extension studies have documented that the efficacy observed during the initial placebo-controlled phase was either maintained or improved during follow-up periods of several years[5].

As already noticed in the first clinical trials[3], treatment response and side effects of stimulant therapy in ADHD are extremely variable, and difficult to predict based on clinical observations. Most adolescents and adults discontinue their ADHD medications during the first year of treatment[6], but some adults with ADHD report sustained treatment effects for decades. Although it is unclear whether better treatment adherence/persistence contributes to more favorable long-term outcomes[7], naturalistic studies suggest beneficial effects in multiple life domains.

Effect sizes based on aggregated data from thousands of individuals have limited value in clinical practice, where treatment decisions regarding prescription of ADHD medications to individual patients are mainly based on trial and error. An alternative could be to introduce a data-driven personalized approach. So far, there is limited evidence that stratification of ADHD patients into subgroups based on clinical features improves treatment efficacy or long-term outcomes. However, recent biomarker studies, in particular molecular genetic research, suggest that ADHD is a highly heterogenous condition, with multiple risk factors implicating distinct pathophysiological mechanisms and different clinical trajectories. Future treatment studies should systematically explore how the wealth of genomic data collected from ADHD patients can be used for patient stratification and "genome guided" personalized interventions[8].

Cortese et al argue that longer-lasting randomized controlled trials (RCTs) are needed to establish long-term effects, but also acknowledge the practical, financial and ethical challenges associated with such trials. They mention the possibility of supplementing conventional RCTs with discontinuation-controlled trials. These latter trials in patients who already are on stable medication may be easier to conduct and probably better mirror clinical practice than conventional RCTs, as only a minority of adults treated with ADHD medications would be eligible for those RCTs[9].

As all published RCTs have limited duration and cannot predict long-term (years or decades) treatment effects, such data will need to come from other study designs. "Real-world" registry studies have the advantages of large sample sizes (millions of individuals), less selected and more relevant study populations, potentially many years of observations, and the prospect of exploring multiple outcomes and interactions. Moreover, prescription registry data can detect rare outcomes and complications, while RCTs are underpowered to catch them. Scandinavian prescription registry studies show that, during periods of stimulant medication, ADHD patients are less likely to be involved in criminal acts or accidents, or to experience severe psychiatric comorbidities. However, registry studies also show that stimulant treatment is associated with an increased risk of cardiovascular diseases. More longitudinal data are needed to evaluate the long-term effects of stimulant use on other outcomes, such as quality of life and work participation. Although registry studies have limitations, including the inherent problem of proving causality, they will probably be – with triangulation of evidence from different data sources – the major source of new insights into the long-term trajectory of ADHD and its treatment.

In addition to the urgent need to establish the long-term efficacy and safety of ADHD treatments, this field is facing new and overarching challenges, including questions about societal consequences of treating an increasing proportion of the population with psychoactive substances. ADHD symptoms and impairments are dimensionally distributed in the general population, without obvious borders between typical and "pathological" or "neurodivergent" behaviors. This is recognized in the DSM-5, that specifies different levels of severity within the categorical diagnosis of ADHD.

As suggested by Cortese et al, a provisional diagnosis of "unspecified ADHD" could be applied to people who do not fulfil the diagnostic criteria of impairment across multiple settings. This implies that more people would be diagnosed. Critics argue that increasing the number of people who receive a diagnosis of ADHD

(or any other condition) involves a "medicalization" and inevitably more (pharmacological) treatment. This is probably not correct. In all areas of medicine, diagnoses are used not only to select people for treatment, but also to avoid unnecessary treatment, for instance if the condition is considered mild or transient, or if there are no proven treatments available. Thus, the application of more differentiated diagnoses of mild or "unspecified" ADHD could potentially lead to fewer people being treated.

In summary, we need more data and new tools to explore many aspects of adult ADHD etiology, management and long-term outcomes. This research agenda should be adapted to a new clinical reality marked by increasing rates of several psychiatric disorders. Society is rapidly changing; people at all ages are increasingly being exposed to massive amounts of potentially addictive and "neurotoxic" electronic devices and social media. Although causality has not been formally proven, there seems to be a correlation between social media use and ADHD, anxiety and mental distress. It is imperative to explore how such new and emerging risk factors could add to and interact with established environmental and genetic risks.

Jan Haavik
Department of Biomedicine, University of Bergen, Bergen, Norway; Bergen Center for Brain Plasticity, Division of Psychiatry, Haukeland University Hospital, Bergen, Norway

1. Cortese S, Bellgrove MA, Brikell I et al. World Psychiatry 2025;24:347-71.
2. Strohl MP. Yale J Biol Med 2011;84:27-33.
3. Conners CK, Eisenberg L. Am J Psychiatry 1963;120:458-64.
4. Ostinelli EG, Schulze M, Zangani C et al. Lancet Psychiatry 2025;12:32-43.
5. Fredriksen M, Halmøy A, Faraone SV et al. Eur Neuropsychopharmacol 2013;6:508-27.
6. Brikell I, Yao H, Li L et al. Lancet Psychiatry 2024;11:16-26.
7. Ferrin M, Häge A, Swanson J. Eur Child Adolesc Psychiatry 2025;34:867-82.
8. Haavik J. Front Psychiatry 2022;13:925442.
9. Garcia-Argibay M, Chang Z, Brikell I et al. Lancet Psychiatry 2025;12:131-9.

DOI:10.1002/wps.21347

ADHD, substance use disorders and stimulant treatment: understanding the relationships

Adults with attention-deficit/hyperactivity disorder (ADHD) are at increased risk for substance use disorders, with rates 2-3 times higher than the general population. Several questions about this association immediately come to mind. What might be the basis for this increased risk? Are specific drugs of abuse associated with ADHD? Is there an association with past or current stimulant use? And how concerned should we be about stimulant misuse and/or abuse, which is increasingly prevalent[1]?

Substance abuse often begins in adolescence, and its onset is earlier in individuals with ADHD. Risk is particularly elevated when there is comorbidity, most specifically with conduct disorder. But increased risk for substance use disorders is not solely accounted for by comorbidity. ADHD, conduct disorders and substance use disorders all share high levels of impulsivity and sensation seeking. They also share genetic variants. Moreover, all three disorders are characterized by a hypodopaminergic state and associated low reward responsiveness.

Drugs of abuse produce an increase in dopaminergic neurotransmission. Psychostimulants have been a mainstay of ADHD treatment, and their beneficial effects are also attributed to enhanced dopaminergic activity. Several studies in animal models have found that early exposure to stimulants may produce sensitization to later exposure[2]. Hence, the question of whether stimulant treatment increases risk of substance use disorders has been raised.

A relatively large literature has examined this question. The most recent meta-analysis found no increased or decreased risk[3]. Longitudinal data from the Multimodal Treatment Study of ADHD also did not find increased risk; stimulant use went down dramatically during adolescence, as substance use was increasing[4]. Most compelling are data from Swedish registries, which indicate a marked decrease in substance abuse in association with stimulant treatment[5]. Longer duration of treatment was associated with lower rates of substance abuse. Thus, while it is possible that for selected individuals stimulant treatment could contribute to substance abuse, available data indicate that this is rare if it occurs. More likely is that a variety of biopsychosocial risk factors, such as impulsivity, reward sensitivity, and associated comorbid conditions (such as conduct, mood and personality disorders) contribute to risk of substance abuse, particularly in the context of psychosocial stressors.

Consistent with the above, several studies have shown beneficial effects of ADHD treatment on rates of substance abuse. Early stimulant treatment is associated with lower cannabis use in adolescents. Methylphenidate treatment is associated with decreased smoking risk and greater abstinence from nicotine. Atomoxetine treatment produced a greater reduction of heavy drinking in recently abstinent adults with ADHD. Moreover, high-dose treatment with long-acting racemic amphetamine resulted in lower drug use in individuals with cocaine use disorder[6]. While these findings are encouraging, positive effects of treatment are best measured by reduction in use rather than abstinence, which remains elusive.

More concerning are the high rates of misuse, diversion and abuse of prescription stimulants. Before delving into this subject, some definitions are in order[7]. Misuse is intentional therapeutic use of a substance in an inappropriate way. Abuse is the intentional non-therapeutic use of a drug to achieve a desirable psychological or physiological effect. Of note, misuse is much more common than abuse. Non-medical use is the use of a drug without a prescription or in a way other than prescribed, which includes both misuse and abuse. Finally, diversion is giving or selling the drug to another person.

Stimulant non-medical use is a particular problem among late

adolescents and young adults – the 18-25 year old group is most vulnerable – and is particularly problematic in communal social settings. This phenomenon has been amply reported among US college students, but the problem has also been documented in other countries. Rates of misuse are highest with immediate release stimulant formulations, which have the fastest onset of effect and are the easiest to obtain. In the US, amphetamines are the most frequently misused stimulants[1]. Whether this is because they are more abusable or simply more available remains unknown.

The most common motivation for stimulant non-medical use is to improve academic performance[1,7]. Other motivations include self-medication of suspected ADHD, which is indeed more prevalent in stimulant misusers than non-misusing controls, weight reduction, and increasing energy or staying awake. More concerning is the desire for euphoria or to heighten the effects of alcohol. A minority of stimulant non-medical use is specifically to get "high". This is sometimes achieved via insufflation or injecting the drug. Non-oral use is especially concerning, because it is associated with the highest rates of untoward medical consequences.

A variety of environmental and psychological factors are known to contribute to stimulant non-medical use, including lack of awareness by prescribers, the perception that stimulant misuse is common, and that it is harmless and morally acceptable. Methods to combat stimulant misuse build on knowledge regarding the types of medications most often misused and abused, and the psychosocial factors which either breed or enable this behavior.

Recommendations include using non-stimulants or long-acting stimulants[8], restricting the prescription of immediate release stimulants, limiting the number of pills in each dispensation and monitoring use, obtaining toxicology testing when indicated, counseling patients and families about potential medical and legal dangers of misuse and diversion, and conducting educational and preventive intervention programs for prescribers and students – ideally addressing the psychosocial and perceptual risk factors described above.

Taking all of the above into consideration, how should we understand the complex relationships among ADHD, substance use and stimulant medication? And how should abuse of stimulant medication be managed? In many countries, some or all of the prescription stimulants are classified as drugs of abuse, and their use is either forbidden or severely restricted. Seen from one vantage point, this is certainly understandable. Abuse of prescription stimulants is a public health problem and efforts to curtail it are warranted. But, on the other hand, so is ADHD. This is the most prevalent child neuropsychiatric disorder worldwide. It is highly impairing for individuals, families and society. It is associated with numerous behaviors and clinical features that carry high morbidity and mortality. It increases risk for other psychiatric disorders later in life – including substance abuse.

Most importantly, stimulant treatment does not in itself increase risk for substance abuse, and has been shown to be protective at the population level. In addition, stimulant treatment can partially mitigate the severity of substance abuse and can aid in treating selected individuals with the condition. Moreover, the large majority of people do not abuse their stimulant medication, and people who abuse stimulants also abuse other drugs. Indeed, early stimulant misuse likely indicates emerging substance use disorder[9].

In summary, the relationship between ADHD and substance use disorder is complex. The two conditions frequently co-occur, complicating management. Stimulant treatment does not in itself cause substance use disorders, and can be used to advantage provided certain precautions are taken. Stimulant misuse is more likely to be associated with substance use disorders[1]. Mitigation strategies include prioritizing non-stimulant medications, and using long-acting formulations if stimulant treatment is needed. Monitoring for misuse, offering abuse prevention programs to high-risk populations, and combining psychosocial treatment with medication are also effective methods for decreasing risk.

Jeffrey H. Newcorn
Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA

1. Han B, Jones CM, Volkow ND et al. JAMA Psychiatry 2025;19:e250054.
2. Ivanov I, Bjork JM, Blair J et al. Neurosci Biobehav Rev 2022;135:104575.
3. Humphreys KL, Eng T, Lee SS. JAMA Psychiatry 2013;70:740-9.
4. Molina BSG, Kennedy TM, Howard AL et al. JAMA Psychiatry 2023;80:933-41.
5. Chang Z, Lichtenstein P, Halldner L et al. J Child Psychol Psychiatry 2014;55:878-85.
6. Levin FR, Mariani JJ, Specker S et al. JAMA Psychiatry 2015;72:593-602.
7. Faraone SV, Rostain AL, Montano CB et al. J Am Acad Child Adolesc Psychiatry 2020;59:100-12.
8. Crunelle CL, van den Brink W, Moggi F et al. Eur Addict Res 2018;24:43-51.
9. McCabe SE, Schulenberg JE, Wilens TE et al. Psychiatr Serv 2024;75:622-9.

DOI:10.1002/wps.21348

The efficacy of cognitive-behavioral therapy for adults with ADHD

Cognitive-behavioral therapy (CBT) for attention-deficit/hyperactivity disorder (ADHD) (henceforth referred to as CBT-ADHD) differs from traditional CBT, as the latter is most often practiced to treat disorders that primarily involve emotions such as anxiety and depression. Instead, the primary focus of CBT-ADHD is to target executive dysfunction, which has been shown to be a major predictor of functional impairment in school and in the workplace for children and adults with ADHD[1].

CBT-ADHD aims to improve the executive functions of time management, organization, and planning. As such, it features specific strategies – both cognitive and behavioral – to facilitate time awareness, prioritizing, scheduling, tracking, and overcoming distraction and procrastination[2,3]. These intervention programs also typically include components of traditional CBT with respect to identifying and restructuring negative automatic thoughts that generate anxiety and depression, which are prevalent among individuals with ADHD[4] and contribute further to distress and impairment.

The UK and Australian ADHD guidelines recommend to utilize CBT as a first-line intervention, in conjunction with medication, for adults with ADHD. Although Cortese et al[5] state that "CBT is not designed to treat the core symptoms of ADHD", the DSM-5 set of ADHD symptoms has been utilized as a primary outcome measure in virtually all the studies of the efficacy of CBT-ADHD, with significant positive findings. In fact, multiple core symptoms of ADHD overlap with executive dysfunctions and are addressed in the program. These include failure to complete tasks; difficulty with organization; avoidance of tasks requiring sustained mental effort; losing things; and forgetting things.

The earliest trials of CBT-ADHD in adults were published in 2010, and included respectively 86[2] and 88[3] participants, who were randomly assigned to receive CBT-ADHD in either individual[2] or group[3] modality, or to receive an active control condition ("psychological placebo") intended to control for the non-specific effects of treatment, principally social support. The control condition was either relaxation with educational support[2] or a support group[3]. Targets and strategies, as well as the number of sessions (twelve) were otherwise quite similar in the two studies. Results in both studies were based on blinded, investigator-rated, well-validated structured assessments of core ADHD symptoms, which, it should be noted, is the "gold standard" for outcomes of clinical trials, and has been determined to be more reliable and valid than self-report measures for this purpose. Results revealed moderate effect sizes of 0.52 and 0.58, respectively, favoring CBT-ADHD. The adult protocol was subsequently revised for the needs of college students with ADHD, with comparable positive results[6].

A total of 17 randomized controlled trials completed by 2023, including the two studies above, were entered into a meta-analysis[7], categorized with respect to whether the control condition was waitlist, treatment as usual (TAU), or an active control intervention. TAU typically involved some combination of medication management and individual supportive follow-up visits. The active control condition was psychoeducation, support, or relaxation (as described in one of the studies above[2]). Effect sizes for investigator-rated core ADHD symptoms for waitlist, TAU, and active control were 1.03, 0.66 and 0.32, highlighting the importance of an active control to isolate the specific benefits of CBT over and above the generic effects of social support or psychoeducation.

Importantly, from a clinical perspective, a separate meta-analysis found that CBT-ADHD added significantly to the benefits of stimulant therapy, thereby providing support for combination treatment[8].

In a meta-analysis including 20 randomized controlled trials of CBT-ADHD, of which five had active controls and 12 were uncontrolled pre-test/post-test comparisons, CBT significantly improved symptoms of anxiety and depression, as well as quality of life and emotion dysregulation[9]. These changes were predicted by the reduction in ADHD symptoms, and thus may be an indirect effect of facilitating the individual's performance and management of daily life functions.

As Cortese et al highlight, maintaining engagement is an obvious concern in treating individuals with attentional difficulties. In the CBT-ADHD programs for adults[3] and college students[6], engagement is elicited and maintained by the use of multiple strategies, including: a) initially highlighting the importance of attending all sessions, in that each participant is effectively serving as a "co-therapist" for the others; b) highlighting that the number of home exercises has been shown to predict benefit from the program; c) conducting a round-table review of each participant's experience with the home exercise at the start of each session; d) utilizing the Socratic method and facilitating discussion when new strategies are presented; e) reassuring participants that they should never skip a session because they have not completed the home exercise ("because any and all experience shared is helpful as 'grist for the mill'"); f) reviewing the upcoming home exercise at the end of each session, with anticipatory trouble-shooting.

Feasibility was monitored in these studies via the completion rates of treatment, defined as having attended at least 9 of the 12 group sessions, which were reported as 87% and 83% for the adult and college programs, respectively. Acceptability of treatment was assessed via participant-completed ratings of the "helpfulness" of each of the eleven strategies presented in the program on a 4-point scale: 0 ("not at all helpful"); 1 ("slightly helpful"); 2 ("moderately helpful") and 3 ("very helpful"). Results revealed that, in the adult study, six of the eleven strategies were rated by more than half of the participants as either "moderately helpful" or "very helpful". For college students, the corresponding figure was eight of the eleven strategies. There was notable overlap between the strategies rated most helpful by the adults and college students.

There is a clear need for a definitive large-scale study in which CBT and medication, along with their respective placebo control conditions, are compared separately and together. Furthermore, there has been little investigation of the long-term maintenance of gains of CBT-ADHD, or of the potential utility of booster sessions. There would also be value in tailoring and testing CBT-ADHD for treatment of individuals with specific co-occurring symptoms, including anxiety, depression, and substance use, and in assessing the relative benefits of individual and group modalities for these subgroups.

In conclusion, research to date provides strong support for the efficacy of CBT-ADHD for clinical use both as a standalone treatment for ADHD, and as a beneficial adjunct to medication.

Mary V. Solanto
Zucker School of Medicine at Hofstra-Northwell, Lake Success, NY, USA

1. Barkley RA, Murphy KR. Arch Clin Neuropsychol 2010;25:157-73.
2. Safren SA, Sprich S, Mimiaga MJ et al. JAMA 2010;304:875-80.
3. Solanto MV, Marks DJ, Wasserstein J et al. Am J Psychiatry 2010;167:958-68.
4. Kessler RC, Adler LA, Barkley RA et al. Am J Psychiatry 2006;163:716-23.
5. Cortese S, Bellgrove MA, Brikell I et al. World Psychiatry 2025;24:347-71.
6. Solanto MV, Scheres A. J Atten Disord 2021;25:2068-82.
7. Liu CI, Hua MH, Lu ML et al. Psychol Psychother 2023;96:543-59.
8. Li Y, Zhang L. J Atten Disord 2024;28:279-92.
9. Lopez-Pinar C, Martinez-Sanchis S, Carbonell-Vaya E et al. J Atten Disord 2020; 24:456-78.

DOI:10.1002/wps.21349

Cut from the same cloth: neurobiological continuity between childhood and adult ADHD

The growing recognition of adult attention-deficit/hyperactivity disorder (ADHD) has been accompanied by discussions over the strength of its evidence base, that in turn feed into debates over the "reality" – the ontological status – of ADHD itself. Against this context, it is helpful to consider in detail the literature on the neurobiological foundations of adult ADHD, expanding upon the points already raised by Cortese et al[1] in their excellent paper. It is particularly salient to ascertain the extent to which the neural features tied to childhood ADHD extend into adulthood among those who retain the diagnosis, and the genetic variants underpinning childhood ADHD overlap with those observed in adult ADHD.

To what extent are the neural differences seen in adult ADHD carried forward from childhood? This question is most directly answered by combining clinical and neuroimaging observations acquired in tandem from childhood into adulthood. There are only a handful of such studies, reflecting the challenges of longitudinal work conducted over decades, compounded by instability in the measurement of neural features arising from rapid advances in neuroimaging technologies.

One longitudinal *in vivo* structural magnetic resonance imaging (MRI) study, conducted over 20 years on the same magnetic resonance scanner, found that young adults whose ADHD symptoms had persisted from childhood showed neural differences that also persisted in a relatively "fixed" manner from childhood[2]. These neuroanatomic features localized to the cognitive control and default mode networks, with changes to the thickness of the cortex in the posterior cingulate, right inferior parietal, and dorsolateral prefrontal regions. By contrast, those whose childhood ADHD symptoms largely resolved by adulthood showed an accompanying significant convergence toward typical cortical dimensions, rectifying early anomalies.

A similar theme emerged in a longitudinal study of cerebellar anatomy, albeit over a much shorter adolescent time window. While some midline cerebellar regions (superior vermis) showed fixed differences, atypical features of the cerebellar hemispheres persisted only among those who had persisting symptoms. Finally, a study spanning late childhood into early adolescence showed that persisting atypical microstructure of thalamic, striatal, and long association white matter tracts was found among those youth whose ADHD symptoms persisted[3]. In short, longitudinal neuroimaging studies point to subtle anatomic and white matter microstructural differences in children with ADHD that are often carried forward into adulthood if core symptoms persist.

Given the sparsity of longitudinal imaging data, other studies have focused on adults who have been followed clinically since childhood, thus enhancing diagnostic certainty, but who have had neuroimaging for the first time in adulthood. Such studies allow not only diagnostic comparisons but also contrasts of adults whose childhood ADHD has persisted against those whose childhood ADHD has remitted[4]. They find that those with adult ADHD (i.e., with the persistent form of ADHD) show atypical anatomic features, ranging from a thinner cortex and decreased thalamic grey matter density, to atypical microstructure of the white matter tracts within attention control (e.g., the inferior fronto-occipital fasciculus) and reward processing (e.g., the uncinate fasciculus) networks. Atypical features in adults with ADHD symptoms persisting from childhood are also found for the brain's intrinsic functional architecture, mapped by both resting state functional MRI and resting state magnetoencephalography, and for brain activity during ADHD-related cognitive processes, such as response inhibition. It seems likely that these neural features, which resemble those reported in studies of childhood ADHD, are carried forward from childhood in tandem with symptom persistence, while they are not present in adults whose childhood ADHD has remitted.

Counter to this evidence for neural differences in adult ADHD, several meta-analytic studies report either minimal or no diagnostic differences[5]. The meta-analytic null findings may stem from the reliance on mostly small, underpowered cross-sectional studies. Faced with the limitations of meta-analyses, some initiatives such as the ENIGMA consortium have taken a mega-analytic approach. This involves the use of individual-level imaging data acquired from multiple cohorts, which provides impressive sample sizes that can be analyzed using methods that account for "site of acquisition" effects. Additionally, many mega-analytic studies process the "raw" imaging data on uniform pipelines and can thus employ consistent quality control standards. The use of individual-level data also allows individual-level confounds to be controlled, including medication history and co-occurring conditions.

What do these mega-analytic studies find? Considering neuroanatomy, the ENIGMA consortium reported that ADHD diagnostic differences are most marked in childhood, and present only at trend level, if at all, in adults, though the limited number of adults does not allow definitive conclusions[6]. Mega-analytic studies of brain's functional architecture in children with ADHD find significant but small differences in the connectivity between the default mode network and task-positive networks, and within the brain's cortico-striatal information processing loops[7]. It will be fascinating to see if similar differences are present in forthcoming well-powered mega-analyses of adult brain function.

Looking to the future, the neurobiological understanding of adult ADHD will be transformed by rapid technological advances, such as imaging at ultra-high field strengths (currently of 7 Tesla). Among its many advantages, high field strength imaging allows the precise quantification of key neurotransmitters, both inhibitory (such as GABA) and excitatory (such as glutamate). It is noteworthy that early *in vivo* imaging studies suggest an altered balance between GABA and glutamate levels in ADHD, and this finding is complemented by genetic studies that also point to these neurotransmitters[8].

Turning to genomics, there is a compelling case for an overlap between the genetic features underpinning childhood and adult

ADHD. The common genetic variants that explain part of the high heritability of adult ADHD are very similar to those found in childhood ADHD, with a genetic correlation around 0.8[9]. Indeed, polygenic scores, that reflect genome-wide measures of common variants tied to ADHD, are higher among those with ADHD that persists into adulthood compared to childhood-limited forms. Furthermore, longitudinal twin studies show that genetic factors account for most of the adolescent change in hyperactive-impulsive symptoms and around half of the change in inattention, with much more modest contributions from environmental factors. In short, the high heritability of adult ADHD is partly explained by common genetic variation that is shared with childhood ADHD. The next step will be to quantify the role of rarer forms of genetic variation, such as copy number variants and deleterious point mutations.

Due partly to collaborative efforts, brain-based and genomic models of adult ADHD are being rigorously tested. Several neural features robustly tied to childhood ADHD extend into adulthood when symptoms persist, and the genetic variants that underpin childhood ADHD overlap considerably with those seen in adult ADHD. This neurobiological continuity provides an important evidence base to inform both scientific thinking and the public understanding of adult ADHD.

Philip Shaw
King's Maudsley Partnership for Children and Young People and Department of Child and Adolescent Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

1. Cortese S, Bellgrove MA, Brikell I et al. World Psychiatry 2025;24:347-71.
2. Sudre G, Mangalmurti A, Shaw P. Neurosci Biobehav Rev 2018;94:198-209.
3. Fuelscher I, Hyde C, Thomson P. Biol Psychiatry Cogn Neurosci Neuroimaging 2023;8:1103-12.
4. Shaw P, Sudre G. Biol Psychiatry 2021;89:152-61.
5. Cortese S, Aoki YY, Itahashi T et al. J Am Acad Child Adolesc Psychiatry 2021;60:61-75.
6. Hoogman M, van Rooij D, Klein M et al. Hum Brain Mapp 2022;43:37-55.
7. Norman LJ, Shaw P. Neuropsychopharmacology 2024;50:332-4.
8. Faraone SV, Larsson H. Mol Psychiatry 2019;24:562-75.
9. Rovira P, Demontis D, Sánchez-Mora C et al. Neuropsychopharmacology 2020;45:1617-26.

DOI:10.1002/wps.21350

New developments and potential future research directions in adult ADHD

The review produced by S. Cortese and nineteen other worldwide renowned scientists[1] provides a comprehensive update on the current position of our knowledge and understanding of attention-deficit/hyperactivity disorder (ADHD) in adults. The first conclusion is that, although a lot of progress has been made by research and in clinical practice, the science of adult ADHD lags behind compared to childhood ADHD and to other psychiatric disorders in adulthood. Key unmet needs according to individuals with lived experience may lead the way to better understanding and treatment of ADHD in adulthood. Here I highlight some new developments and potential future research directions.

Regarding the decision about whether impairment is severe enough in adults with mild symptoms, I would like to challenge the notion of "negligible impairment" in adults who suffer from low self-esteem, inner distress and self-blame. A recent paper compared the endorsement of DSM-5 criteria for ADHD between genders, using the Diagnostic Interview for ADHD in Adults (DIVA-5) in 2,257 adults. Of the five potential domains of impairment, particularly self-esteem issues were highly common (in 89% of women and 81% of men)[2]. The high recognition of self-esteem problems in adults with ADHD may be interpreted as a plea to reconsider this impairment as important, having high impact in both genders.

Sleep loss is present in around 80% of both children and adults with ADHD, even before the start of any stimulant treatment. The most common comorbid sleep disorder is the circadian rhythm sleep-wake disorder, delayed sleep phase type, leading to chronic late and short sleep[3]. Melatonin onset in the evening has been shown to be 1.5 hrs delayed in adults with ADHD, which may point to a dysregulation of the circadian clock[4]. This understudied area needs more attention in ADHD.

Recently, an alarmingly high number of health conditions have been found to be more common in adults with ADHD, including obesity, diabetes type 1 and 2, cardiovascular diseases, dementia and Parkinson's disease, migraine, asthma, allergies, irritable bowel syndrome and ulcerative colitis, arthritis, many autoimmune disorders, epilepsy, early menopause, and chronic obstructive pulmonary disease. It may be time to consider ADHD as a systemic disease, including both mental and physical manifestations, and start rethinking from there. We know that ADHD is a heritable condition, but what the 76 genes currently identified exactly do is still subject of investigation[5]. All these comorbidities may help to find the key(s) to more general factors involved in both mental and physical diseases.

The high rate of autoimmune disorders and inflammatory diseases with a strong genetic load may point in this direction. ADHD has been found to be associated with weak connective tissue, manifesting in a variety of hypermobility syndromes, Ehlers-Danlos syndrome, musculoskeletal pain syndromes; inflammation of the gut resulting in food intolerance; as well as dysautonomia or orthostatic intolerance, resulting in dysregulation of blood pressure, dizziness, and palpitations[6]. Also, failure of the immune system to deal with infections such as COVID-19, resulting in long COVID, has been detected more often in children with ADHD. Other immune-related disorders have also been associated with ADHD, such as selective immunoglobulin A deficiency, and familial mediterranean fever[7].

In summary, if genetic susceptibility for ADHD is associated with weakness of connective tissue, failure of the immune system,

and low-grade inflammation related to obesity, chronic sleep loss and an unhealthy lifestyle, we may get a better understanding of the complex etiology of ADHD and its broad range of mental and physical manifestations.

Women have a share of 50% in the total prevalence of ADHD, but they have been remarkably understudied. Cortese et al correctly state that we need high-quality research including large numbers of female participants across all the different phases of life. I would like to add that this research should focus on the specific female presentation of ADHD, the late recognition and underdiagnosis in girls and women, the interaction between neurotransmitters such as dopamine and female hormones associated with increased premenstrual, postpartum and peri-menopausal mood and ADHD symptoms, the best treatment of ADHD in women, as well as the understudied comorbidity with hormonal, gynaecological and cardiovascular disorders[8,9].

Regarding ADHD in the elderly, further research is clearly needed, especially on differentiation from cognitive decline, which may not be so easy as it seems. A growing number of patients ask for help for their parent who is already living in a nursing home, when it may be really too late to disentangle ADHD symptoms from cognitive decline and dementia. They may occur together, and studies have not sufficiently looked into differences and overlap using biological and neuropsychological measures. There are no randomized controlled trials using stimulants in older people, and clinical guidance on treatment of somatic comorbidities, such as hypertension or cardiovascular diseases, when using stimulants is missing. It is therefore not surprising that elderly people living with ADHD get treatment only in 0.09% of cases. The number of these people is rapidly increasing, and the relevant questions to science are becoming more urgent.

People living with ADHD pointed out that there are important health care gaps due to "not being seen, not being recognized, not being diagnosed nor treated". When they are treated, the available treatment often does not include family support and cognitive behavior therapy. At the same time, there is an enduring stigma from society towards people with a diagnosis of ADHD. This is even more true for women, people of color, other minorities, and people with low income, multiplying inequities. The elephant in the room here may be education, that is needed for health care professionals, society, schools and workplaces, in order to reduce stigma and increase recognition and support for people living with ADHD.

J.J. Sandra Kooij
Psycho-Medical Center, Amsterdam UMC/VUMc and PsyQ, The Hague, The Netherlands

1. Cortese S, Bellgrove MA, Brikell I et al. World Psychiatry 2025;24:347-71.
2. Platania NM, Starreveld DEJ, Wynchank D et al. Front Glob Womens Health 2025;6:1549028.
3. Van Veen MM, Kooij JJ, Boonstra AM et al. Biol Psychiatry 2010;67:1091-6.
4. Bijlenga D, Vollebregt MA, Kooij JJS et al. Atten Defic Hyperact Disord 2019;11:5-19.
5. Demontis D, Walters RK, Martin J et al. Nat Genet 2019;51:63-75.
6. Csecs JL, Iodice V, Rae CL et al. Front Psychiatry 2022;12:786916.
7. Merzon E, Farag R, Ashkenazi S et al. J Clin Med 2024;13:6075.
8. Dorani F, Bijlenga D, Beekman AT et al. J Psychiatr Res 2021;133:10-5.
9. Ter Beek LS, Böhmer MN, Wittekoek ME et al. Arch Womens Ment Health 2023;26:851-5.

DOI:10.1002/wps.21351

Charting the evolution of artificial intelligence mental health chatbots from rule-based systems to large language models: a systematic review

Yining Hua[1,2], Steve Siddals[2], Zilin Ma[3], Isaac Galatzer-Levy[4,5], Winna Xia[2], Christine Hau[2], Hongbin Na[6], Matthew Flathers[2], Jake Linardon[7], Cyrus Ayubcha[7], John Torous[2]

[1]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA; [2]Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA; [3]Intelligent Interactive Systems Group, Harvard School of Engineering and Applied Sciences, Allston, MA, USA; [4]Department of Psychiatry, New York University Grossman School of Medicine, New York, NY, USA; [5]Google Research, New York, NY, USA; [6]Australian Artificial Intelligence Institute, University of Technology Sydney, Sydney, NSW, Australia; [7]SEED Lifespan Strategic Research Centre, School of Psychology, Faculty of Health, Deakin University, Geelong, VIC, Australia

The rapid evolution of artificial intelligence (AI) chatbots in mental health care presents a fragmented landscape with variable clinical evidence and evaluation rigor. This systematic review of 160 studies (2020-2024) classifies chatbot architectures – rule-based, machine learning-based, and large language model (LLM)-based – and proposes a three-tier evaluation framework: foundational bench testing (technical validation), pilot feasibility testing (user engagement), and clinical efficacy testing (symptom reduction). While rule-based systems dominated until 2023, LLM-based chatbots surged to 45% of new studies in 2024. However, only 16% of LLM studies underwent clinical efficacy testing, with most (77%) still in early validation. Overall, only 47% of studies focused on clinical efficacy testing, exposing a critical gap in robust validation of therapeutic benefit. Discrepancies emerged between marketed claims ("AI-powered") and actual AI architectures, with many interventions relying on simple rule-based scripts. LLM-based chatbots are increasingly studied for emotional support and psychoeducation, yet they pose unique ethical concerns, including incorrect responses, privacy risks, and unverified therapeutic effects. Despite their generative capabilities, LLMs remain largely untested in high-stakes mental health contexts. This paper emphasizes the need for standardized evaluation and benchmarking aligned with medical AI certification to ensure safe, transparent and ethical deployment. The proposed framework enables clearer distinctions between technical novelty and clinical efficacy, offering clinicians, researchers and regulators ordered steps to guide future standards and benchmarks. To ensure that AI chatbots enhance mental health care, future research must prioritize rigorous clinical efficacy trials, transparent architecture reporting, and evaluations that reflect real-world impact rather than the well-known potential.

Key words: Artificial intelligence, chatbots, rule-based systems, machine learning, large language models, foundational bench testing, pilot feasibility testing, clinical efficacy testing, mental health care

(World Psychiatry 2025;24:383–394)

Mental disorders remain a major contributor to the global burden of disease. An estimated 970 million individuals live with mental health or substance use disorders worldwide, with depression and anxiety among the leading causes of disability-adjusted life years lost[1,2].

Despite the increasing recognition of mental health as a critical component of public health, care systems remain underfunded and overwhelmed[3], with fewer than five mental health professionals available per 100,000 people globally[4]. This unmet need is more pronounced in low- and middle-income countries, where more than 75% of individuals with mental health conditions fail to receive treatment[5]. The need for scalable, accessible solutions has driven interest in digital interventions, particularly conversational agents or chatbots, as potential tools to support mental health care by way of screening, psychoeducation, and therapy augmentation[6].

While the rapid growth of artificial intelligence (AI)-driven chatbots offers new possibilities for mental health applications, their potential remains unclear. Many clinicians, policy makers and researchers are uncertain about whether these tools are appropriate for clinical deployment. Much of this uncertainty arises from the heterogeneous nature of chatbot technologies, which range from basic rule-based systems to advanced large language models (LLMs). While rule-based chatbots rely on pre-programmed scripts or decision trees, LLMs leverage deep neural networks trained on vast datasets to produce more versatile, human-like conversational capabilities.

The conflation of these technologies has led to several challenges. While most health care chatbots remain rule-based, companies often continue to market them as "AI" or even "LLM-driven". This creates misconceptions about their sophistication and reliability, as these systems differ substantially in their capabilities and limitations. Additionally, exaggerated claims about LLMs, such as their ability to pass professional exams or display empathy surpassing that of human clinicians, have blurred the distinction between experimental results and real-world applicability[7,8]. For instance, passing a written test or simulating empathetic dialogue in controlled conditions does not necessarily equate to making accurate clinical diagnoses or supporting patients in complex and dynamic clinical scenarios.

The current state of the AI chatbot literature reflects these challenges. Meta-analyses and reviews often conflate simple feasibility testing results with more complex clinically focused research. In this paper, we applied a staged approach to translational research and clinical development that mirrors the progression from basic science/preclinical research to early phase human testing (phase 1 or 2 clinical trials) to clinical efficacy testing (phase 3 clinical trials). This scheme organizes chatbot studies into three tiers: foundational bench testing for technical feasibility, pilot feasibility testing for assessment of usability and acceptability in humans, and clinical efficacy testing. These categories reflect increasing levels of clinical applicability, providing a structured approach to understanding the field's progress. This approach contextualizes the

current discussion around AI certification with a staged structure[9], and provides a useful scaffold for understanding prior studies and developing future work.

The evolution of AI chatbots reflects decades of advancements in computational approaches to human language, which we categorize into three paradigms: rule-based systems, non-LLM machine learning models, and LLM-based systems. This tripartite framework is designed to clarify distinctions between these paradigms, particularly as the term "AI" has historically encompassed a wide range of technologies, from early deterministic systems to modern probabilistic models.

Rule-based systems were the earliest form of conversational AI. ELIZA, developed in the 1960s, exemplified these systems by simulating a Rogerian psychotherapist through simple pattern-matching and substitution rules[10]. Currently popular mental health chatbots, such as *Woebot*, have been primarily rule-based systems, underscoring the enduring effectiveness of this approach[11]. Most chatbots continue to be employed in narrowly defined tasks, such as structured screening tools or symptom checkers[12], where deterministic outputs remain sufficient. However, their inability to adapt to novel inputs or provide personalized responses has constrained their utility where dynamic and context-sensitive interactions are critical[13]. Despite limitations, rule-based systems laid the groundwork for subsequent innovations by demonstrating the feasibility of automated dialogue.

The transition from rule-based systems to machine learning introduced greater adaptability and probabilistic reasoning into chatbot designs. Machine learning refers to algorithms that identify patterns in data, enabling models to generalize beyond explicit rules and make predictions based on statistical likelihoods[14,15]. Non-LLM machine learning-based chatbots marked a pivotal shift by moving beyond scripted interactions. These systems incorporated natural language processing techniques, such as sentiment analysis and intent recognition, to infer user emotions and tailor responses to context[16]. For example, conversational agents such as *Wysa* employ a combination of machine learning algorithms and rule-based scripts to deliver cognitive-behavioral therapy (CBT) interventions, demonstrating efficacy in reducing symptoms of depression and anxiety in pilot studies[17,18].

While non-LLM machine learning models expanded the scope of chatbot applications, they faced inherent challenges. Their performance often depends on domain-specific training data, which limits generalizability across diverse conversational scenarios. Additionally, these systems struggle with generating coherent and human-like language, as they were typically designed to classify or process inputs rather than produce contextually appropriate outputs.

LLMs represent a paradigm shift in AI chatbots, driven by their ability to generate human-like language with fluency and contextual awareness. Built on Transformer architectures, LLMs such as OpenAI's *GPT* series and Meta's *Llama* models leverage self-attention mechanisms to understand relationships between words and concepts across extensive text passages[19,20]. This architecture enables LLMs to maintain coherence within complex, multi-turn dialogues in ways that previous approaches could not achieve[21]. Unlike earlier machine learning-based systems that primarily focused on classification or limited response selection, LLMs are fundamentally generative in nature – they create novel text rather than selecting from predefined responses.

LLMs' advanced linguistic capabilities have sparked explosive interest in their mental health applications[22], offering the potential to enhance psychoeducation, triage, and supportive interactions. However, the transformative potential of LLMs is accompanied by significant challenges. Their reliance on vast, uncurated datasets introduces risks such as bias, misinformation, and the generation of fabricated or harmful content[23]. In psychiatry, where the consequences of errors can be severe, these limitations raise ethical concerns about reliability and safety.

METHODS

We classified chatbots into three systems:

- *Rule-based systems*. These rely on deterministic scripts (e.g., rule-based conversation systems, simple decision trees), with no data-driven learning. They are ideal for structured, low-risk tasks (e.g., symptom checklists) where predictability ensures safety. However, their rigidity limits their utility in dynamic therapeutic contexts.
- *Machine learning-based systems*. These include traditional machine learning (e.g., support vector machine, SVM) and non-generative deep learning (e.g., recurrent neural networks, RNN; long short-term memory, LSTM; and bidirectional encoder representations from transformers, BERT). While RNN/LSTM and traditional machine learning differ technically (e.g., sequential vs. static data processing), both lack natural language fluency. Grouping these under "machine learning-based systems" reflects their shared limitation in mental health: adaptability without generative capacity.
- *LLM-based systems*. These leverage generative models trained on vast text corpora to produce human-like dialogue. This category includes multimodal models that can process images, audio or other modalities in addition to text, as long as they maintain the core LLM architecture for language generation.

We used a tiered framework that categorizes studies by their evaluation rigor, akin to the translational pipeline from technical validation to real-world clinical impact:

- *T1. Foundational bench testing*. This focuses on technical validation in controlled settings (e.g., scripted scenarios, expert assessments) to ensure that chatbots meet baseline functional and safety standards. For mental health, this stage is critical to verify adherence to clinical guidelines (e.g., suicide risk protocols) before human interaction.
- *T2. Pilot feasibility testing*. This assesses usability and acceptability with human participants (e.g., patients, clinicians) over short-term interactions. While it provides insights into engagement, it often overlooks sustained therapeutic outcomes – a

gap particularly problematic in mental health, where longitudinal efficacy is paramount.

- *T3. Clinical efficacy testing*. This measures clinically meaningful outcomes (e.g., symptom reduction via validated rating scales) over extended periods. It is essential for mental health chatbots, as transient usability gains (T2) do not necessarily equate to therapeutic benefit.

This framework ensures that mental health interventions undergo rigorous validation before clinical deployment. A chatbot that performs well in scripted tests (T1) may still fail in real-world empathy or crisis management, while short-term usability (T2) does not guarantee long-term adherence or relapse prevention. By stratifying evidence into three tiers, the classification enables clinicians and regulators to distinguish technically functional tools from those with proven clinical impact[23].

We systematically reviewed mental health chatbot studies published from January 1, 2020 to January 1, 2025 across PubMed, APA PsycNet, Scopus and Web of Science, following PRISMA 2020 guidelines[24]. Search strings were adapted from prior scoping reviews and optimized for lexical coverage (see supplementary information). Additional records were identified through manual searches of Google Scholar and major AI conference proceedings.

To ensure relevance, only studies evaluating chatbots within a mental health care context were included. Papers focused on psycholinguistics, psychosocial demographics, or predictive models without conversational interfaces were excluded. Only full, peer-reviewed papers written in English were considered. Reviews, meta-analyses and retracted papers were excluded. Eligible studies were required to include an actual evaluation of chatbot performance, excluding protocol descriptions.

The screening and annotation process was a collaborative effort involving the entire research team. Each study was randomly assigned to at least two reviewers, who independently evaluated its eligibility based on predefined inclusion criteria. Any disagreements between reviewers were resolved through group discussions to ensure consistency.

Figure 1 presents the PRISMA flow diagram illustrating the study selection process. A total of 1,727 records were identified through database and manual searches, including 620 from PubMed, 18 from APA PsycNet, 419 from Scopus, 480 from Web of Science, 131 from Google Scholar, and 59 from major AI conferences. After removing 790 duplicates, 937 unique records were screened. Of these, 734 studies were excluded based on title and abstract screening, due to irrelevance or failure to meet the inclusion criteria. Following this, 203 reports were sought for full-text retrieval, but four could not be retrieved. The remaining 199 full-text reports were assessed for eligibility, leading to the exclusion of eight studies that lacked chatbot evaluation, 21 that were duplicates published under different titles, and ten whose models did not meet our definition of LLMs. Ultimately, 160 studies met the inclusion criteria and were included in the systematic review[25-184].

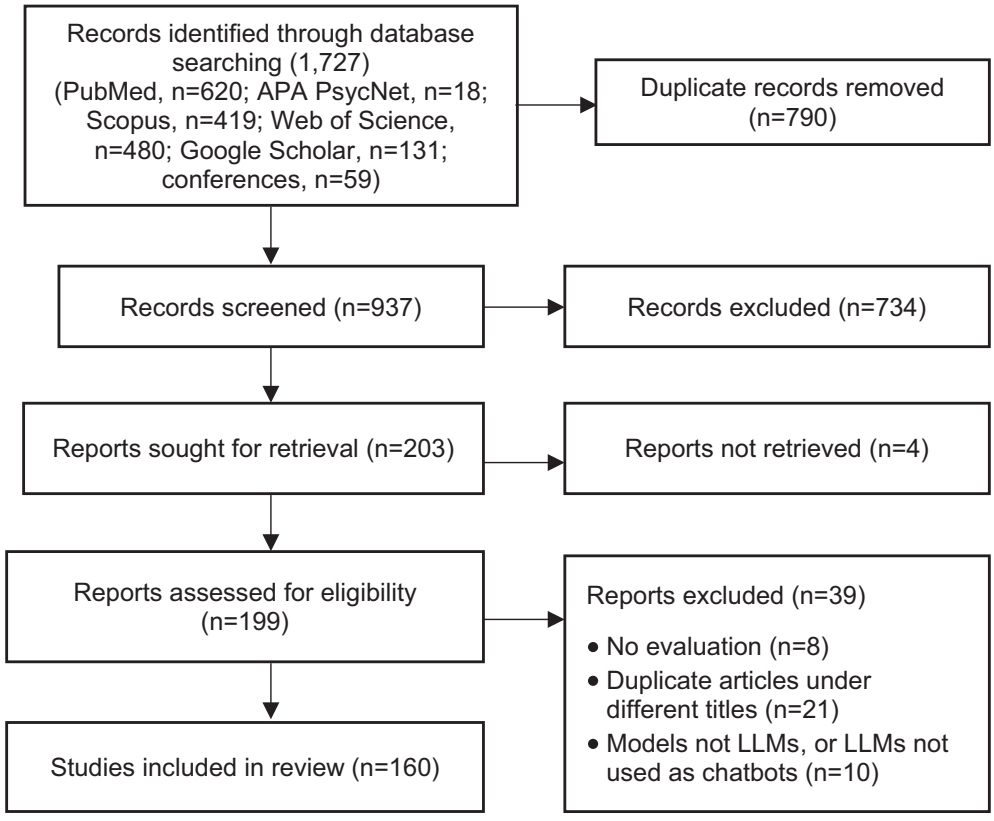

Figure 1 PRISMA flow diagram of study selection process. LLMs – large language models

To ensure consistency and depth in data extraction, senior team members (YH and JT) implemented a structured annotation protocol (see supplementary information), formalized through team training. Key elements included chatbot architecture, evaluation methodology, target conditions, functional purpose, and outcome measures. Each study was also annotated with type-specific information based on its evaluation tier (T1, T2 or T3), capturing evaluator characteristics, usage duration, and relevant clinical instruments. Missing or non-applicable data were systematically flagged to support transparent synthesis.

To analyze the 160 annotated studies, we implemented a structured multi-step methodology to transform raw annotations into standardized themes. In cases where studies were classified into multiple categories, each classification was weighted proportionally (e.g., a study targeting both depression and anxiety would be counted twice, each with weighting 0.5). Reported subtotals and percentages were rounded to the nearest whole number, which may result in apparent summation discrepancies. Computational tools supported initial categorization, with all results refined and validated by domain experts to ensure fidelity and interpretability (see also supplementary information).

RESULTS

Evolution of chatbot architectures

Research interest in mental health chatbots increased substantially over the review period, with the annual number of studies quadrupling from 14 in 2020 to 56 in 2024. Coinciding with this growth, the underlying chatbot architectures underwent a significant transformation (see Figure 2).

Initial research in 2020 (n=14)[25-38] focused exclusively on rule-based systems (100%). The landscape diversified from 2021 (n=28)[39-66], with the emergence of machine learning-based (21%) and the first LLM-based studies (11%), although rule-based systems remained dominant (68%). Machine learning-based approaches peaked in 2022 (40% of 25 studies[67-91]), before stabilizing as a smaller component (14% in 2024).

LLM-based architectures, after comprising 16% of studies in 2022 and 19% in 2023 (n=37)[92-128], surged to represent 45% of studies in 2024 (n=56)[129-184]. This rapid rise made LLMs the most frequently studied architecture in 2024, surpassing rule-based systems (41%), despite the absolute number of rule-based studies remaining relatively consistent in 2023-2024. This trend indicates a decisive shift towards investigating advanced generative models within the field.

Distribution by evaluation methodology and architecture

Research effort across the evaluation stages was primarily focused on clinical efficacy testing (n=75), followed by pilot feasibility testing (n=72), and foundational bench testing (n=13) (see Figure 3).

Chatbot architecture distribution varied markedly across these evaluation stages. Foundational bench testing was dominated by LLM-based systems, which accounted for over two-thirds (77%) of studies at this stage, with smaller contributions from machine learning-based (15%) and rule-based (8%) systems. Rule-based architectures predominated in later stages, accounting for over half of both pilot feasibility studies (58%) and clinical efficacy trials (65%). The proportion of LLM-based studies decreased substantially in these stages, accounting for only 24% of pilot feasibility studies and 16% of clinical efficacy studies. Machine learning-based systems remained a minority across all stages, ranging from 15% in foundational bench testing to 19% in clinical efficacy testing.

This stark contrast between stages indicates that, while LLMs are the primary focus of early technical validation, rule-based systems remain the principal architecture undergoing human testing and clinical evaluation.

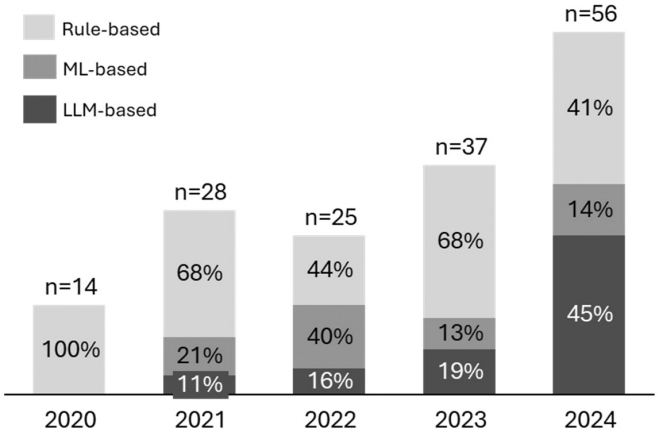

Figure 2 Evolution of chatbot architectures studied from 2020 to 2024. Percentages indicate the proportion of studies utilizing each architecture type within a given year. The number above each bar indicates the total number of studies for that year. ML – machine learning, LLM – large language model.

Target conditions, functional purpose, and outcome measures of chatbot studies

Analysis of the studies' target conditions, functional purpose, and outcome measures revealed distinct patterns in the application and evaluation of different chatbot architectures (see Figure 4).

Examining target conditions, research most frequently addressed general mental well-being (n=51), depression (n=50), and anxiety (n=41). Rule-based systems were the predominant architecture for studies targeting depression (58%) and anxiety (62%), compared to LLM (20-25%) and machine learning-based (13-23%) systems. LLM-based systems showed higher relative representation in studies targeting general mental well-being (28%), compared with rule-based (49%) and machine learning-based (22%) approaches.

Methodologically, studies of general mental well-being were largely in pilot feasibility testing (66%), with fewer in clinical efficacy testing (27%) or foundational bench testing (7%). By contrast, both depression and anxiety interventions had mostly advanced

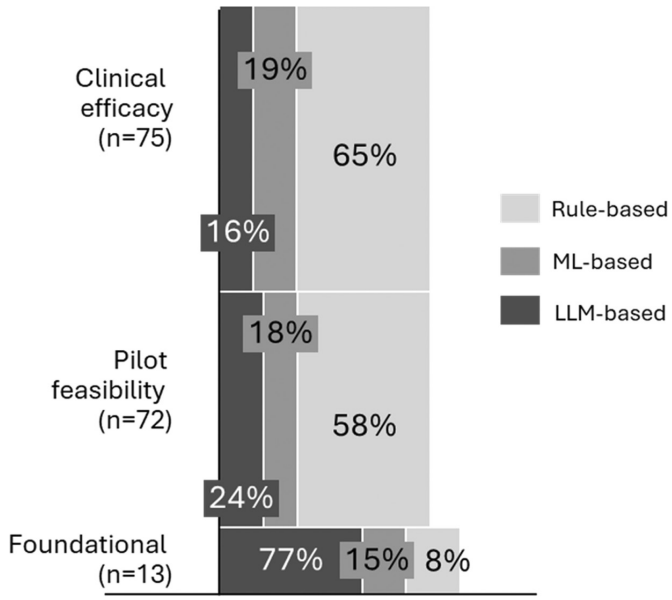

Figure 3 Distribution of studies by evaluation methodology and chatbot architecture. Percentages indicate the proportion of chatbot architectures within each evaluation methodology category. Subtotals and percentages are rounded to the nearest whole number, which may result in apparent summation discrepancies. ML – machine learning, LLM – large language model.

to clinical efficacy trials (57% and 58%, respectively), with smaller proportions in pilot feasibility (38% for depression; 33% for anxiety) and foundational testing (6-8%).

Studies were grouped into five functional purposes, including emotional support (n=46), therapeutic interventions (n=42), education and skills training (n=27), assessment and monitoring (n=21), and general and other functions (n=24). In every category, rule-based systems dominated, ranging from 48% in emotional support to 83% in general and other functions. LLM-based approaches were most represented in emotional support (30%) and assessment and monitoring (29%), while machine learning-based systems peaked at 27% for therapeutic interventions and dipped as low as 4% for general and other functions.

Evaluation methodology varied notably by functional purpose. Pilot feasibility testing was the most common method for assessment and monitoring (78%), emotional support (53%), and education and skills training (49%). By contrast, clinical efficacy testing led in therapeutic interventions (65%), and general and other functions (71%), with education and skills (43%) and emotional support (34%) also seeing substantial clinical work. Foundational bench testing remained minimal, accounting for 7-13% of studies in four categories and 0% for general and other studies.

The choice of chatbot architecture and evaluation stage was associated with the outcome measures prioritized. Studies measuring clinical outcomes (n=99) predominantly employed rule-based (59%) or machine learning-based (18%) systems, with fewer using LLM-based approaches (23%). These clinical outcome studies were most often evaluated via clinical efficacy testing (65%), followed by pilot feasibility (31%) and foundational bench testing (4%).

User experience evaluations (n=46) similarly favored rule-based (64%) and machine learning-based (19%) architectures over LLM-based approaches (17%), and were overwhelmingly conducted as pilot feasibility studies (78%), with smaller proportions in clinical efficacy (17%) and foundational testing (5%). In contrast, technical performance studies (n=15) were dominated by LLM-based systems (54%), with rule-based and machine learning-based approaches at 31% and 15% respectively, and were primarily assessed at the foundational bench (45%) and pilot feasibility (37%) stages, with only 18% reaching clinical efficacy testing.

Characteristics of evaluation stages

The nature of evaluations conducted on mental health chatbots evolved significantly across the research pipeline, particularly concerning the types of participants involved (see Figure 5).

Foundational bench testing (n=13) primarily involved evaluations conducted with clinicians (62%) or represented technical assessments where participant type was "not applicable" (38%). Transitioning to the pilot feasibility testing (n=72), evaluation efforts focused predominantly on general users (78%) to assess usability and acceptability, alongside a substantial inclusion of patients (19%) for initial target population testing. The clinical efficacy testing (n=75) presented a more varied participant profile, characterized by the engagement of clinicians as evaluators (19%) and continued recruitment of general users (25%). Notably, explicitly identified patient participants were less frequently involved (5%) in this final stage compared to pilot studies, while a large proportion of these efficacy trials reported participant type as "not applicable" (37%) or "not specified" (13%), potentially reflecting diverse study designs or reporting practices.

During the pilot feasibility phase (n=72), study durations ranged from under one hour to two years (see Figure 6). Under one hour applied to eight studies, between one hour and one day to eighteen studies, and between one day and one week to sixteen studies. Ten studies did not report a duration. Rule-based architectures appeared in half to two-thirds of studies across every duration band. LLM-based systems featured in 20% to 33% of studies, with a 31% share in the one-day to one-week group. Machine learning-based approaches ranged from 0% in the not-applicable category up to 25% in the one-to-four-week interval. These results show that, although rule-based chatbots dominate pilot feasibility testing, LLM and machine learning systems have also been trialed across the full spectrum of study lengths.

Terminology discrepancies: the meaning of "AI"

Beyond the core findings related to chatbot architectures and evaluation stages, the terminology used to describe these technologies in study titles also warranted examination (see Figure 7).

Analysis revealed ambiguity in the use of the term "AI". Of the 160 included studies, a small proportion (n=21, 13%) explicitly

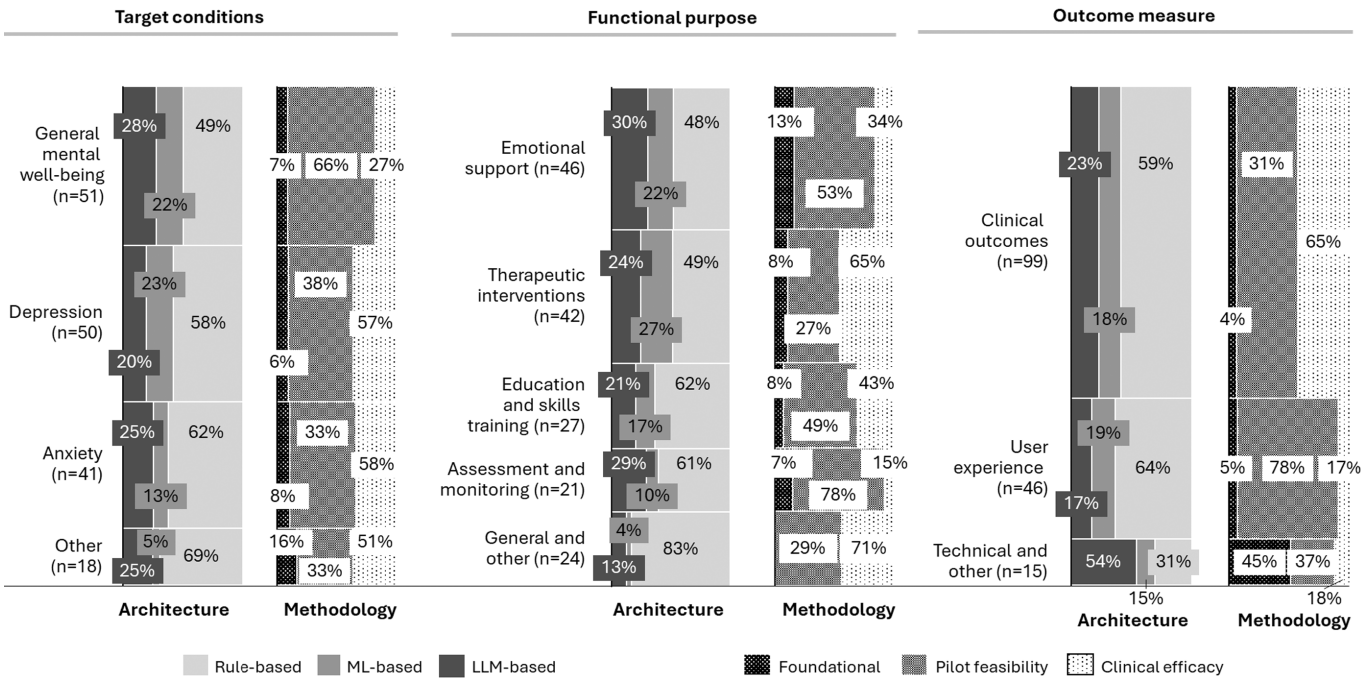

Figure 4 Distribution of chatbot studies by target condition, functional purpose, and outcome measure. Within each subcategory, the left bar indicates the percentage distribution of chatbot architectures used, and the right bar shows the percentage distribution of evaluation methodologies employed. Subtotals and percentages are rounded to the nearest whole number, which may result in apparent summation discrepancies. ML – machine learning, LLM – large language model.

used the term "AI" in their titles; the vast majority (n=139, 87%) did not. Among the 21 studies that did use the "AI" label, the majority (57%) employed advanced LLM-based systems. A further 19% utilized machine learning-based approaches. Notably, however, the "AI" label was also applied in nearly one-quarter (24%) of these cases to studies utilizing rule-based architectures. As rule-based systems operate on predefined scripts without the adaptive learning capabilities characteristic of contemporary machine learning and LLM systems, their inclusion under the general "AI" descriptor contributes to terminological ambiguity and potential misrepresentation of chatbot sophistication within the field.

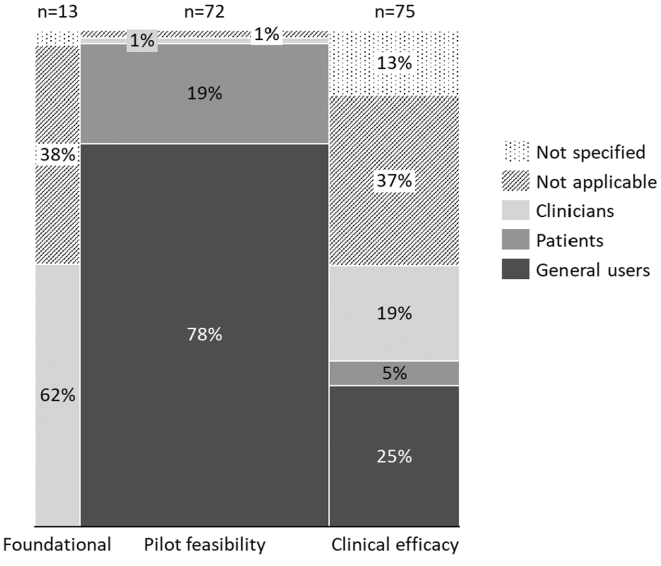

Figure 5 Evaluation participant types across study methodologies. Percentages indicate the distribution of participant types within each methodology. Subtotals and percentages are rounded to the nearest whole number, which may result in apparent summation discrepancies.

DISCUSSION

The rapid adoption of generative LLMs (e.g., *GPT-4*) reflects broader AI trends, but introduces unique risks in mental health contexts[185]. The complexity of mental health conditions, the subjective nature of diagnosis, and the need for contextual understanding further complicate AI integration[186-188]. While early rule-based systems such as *Woebot* prioritized safety through scripted dialogues, LLMs such as *Replika* now risk generating unvalidated advice due to their reliance on uncurated datasets[189]. This tension between innovation and safety, reflected in a December 2024 complaint by the American Psychological Association to the US Federal Trade Commission accusing a generative AI chatbot of harming children[190], underscores the need for structured validation frameworks and research to fill the gaps identified in our results.

A persistent challenge in the field is the misalignment between the marketed rhetoric of "AI-driven" systems and their underlying technical realities. Platforms such as *Woebot* and *Replika* market themselves as "AI-driven", yet the term "AI" remains ambiguously defined and is often employed without clear specification of the

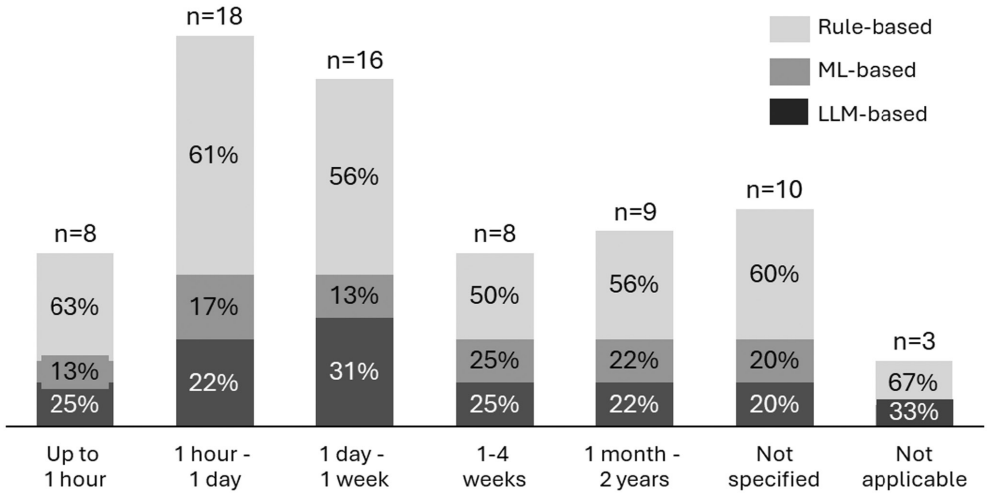

Figure 6 Distribution of study durations for the various chatbot architectures in the 72 pilot feasibility studies. Percentages indicate the proportion of chatbot architectures within each study duration category. Subtotals and percentages are rounded to the nearest whole number, which may result in apparent summation discrepancies. ML – machine learning, LLM – large language model.

underlying model architecture[191].

Early iterations of "AI" chatbots predominantly operated through scripted, rule-based interactions with only rudimentary machine learning enhancements. These rule-based tools, emblematic of the good old-fashioned artificial intelligence (GOFAI) paradigm, have been critiqued for their limited adaptability and depth, standing in sharp contrast to modern data-driven, connectionist approaches[192]. As both technological capabilities and stakeholder expectations evolve, clinicians and patients now increasingly expect AI to represent more sophisticated, dynamic and autonomous connectionist systems, such as LLMs, capable of generating contextually rich, free-form dialogue[193].

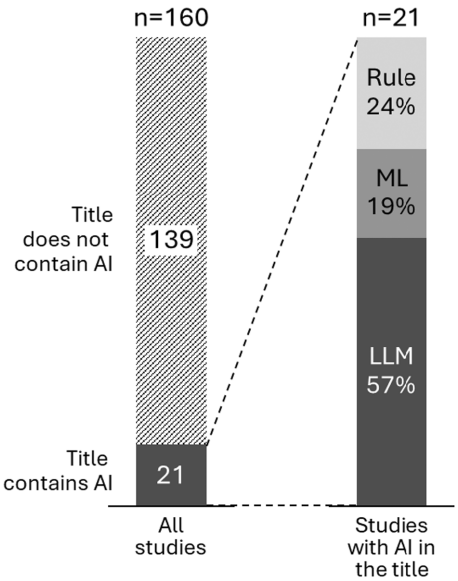

Figure 7 Usage of "AI" in study titles versus actual chatbot architectures. The left bar shows the proportion of all studies containing "AI" in the title. The right bar shows the percentage distribution of underlying chatbot architectures for the subset of studies with "AI" in the title. AI – artificial intelligence, Rule – rule-based, ML – machine learning, LLM – large language model.

This evolving expectation creates confusion, as legacy systems continue to be marketed under the same broad AI umbrella, making it difficult to distinguish between LLM-driven innovations and older rule-based or hybrid approaches. To address this, our classification system offers a structured way to categorize mental health chatbots based on architecture and function, distinguishing rule-based systems that operate through deterministic scripts, machine learning-based systems that enhance adaptability through data-driven models, and LLM-based systems that generate free-form, contextually rich dialogue. This architectural classification is critical, as it allows clinicians, researchers and regulators to appropriately evaluate chatbot capabilities, ensuring that expectations align with actual functionalities rather than misleading claims.

With the increasing number of chatbot studies and the sharp rise in LLM-based chatbot research, the need for standardized evaluation frameworks is more urgent than ever. While chatbots were once predominantly rule-based, requiring only basic assessments of functionality and engagement, the integration of machine learning- and LLM-based systems has generated complexity in evaluation. Unlike rule-based chatbots, which can be tested through predefined workflows, LLM chatbots introduce elements of unpredictability, requiring assessments beyond technical performance. Without a standardized nomenclature for evaluation, studies report outcomes inconsistently, making it difficult to compare findings across research. The introduction of a standardized three-tier evaluation continuum – foundational bench testing, pilot feasibility testing, and clinical efficacy testing – addresses this need by providing a clear progression of evidentiary rigor. It also fits well with recent calls for graded regulation of LLM systems, with certification linked to the role of the chatbot and rigor of its real-world efficacy[9].

Architectural analysis reveals that the focus and stage of evaluation vary significantly depending on chatbot type. Rule-based systems, lending themselves to structured interactions such as

symptom monitoring or delivering psychoeducational content, continue to be the most common architecture evaluated in clinical efficacy trials (65% of T3 studies). This likely reflects their longer history and suitability for interventions where predictability and safety are paramount. Machine learning-based chatbots, offering more adaptability than rule-based systems but lacking the generative fluency of LLMs, are represented modestly across all evaluation tiers (15-19%). In stark contrast, LLM-based systems heavily dominate foundational bench testing (77% of T1 studies), indicating that current research primarily investigates their technical capabilities, such as conversational quality or adherence to specific prompts, often in simulated scenarios. Despite their potential for nuanced, high-context interactions relevant to applications such as assessment or emotional support, LLMs are infrequently evaluated in T3 trials (16%). This suggests that, while LLMs are being actively explored for their technical promise, they have yet to undergo widespread, rigorous testing for clinical benefit in high-stakes mental health contexts, leaving a critical gap in evidence.

A major challenge highlighted by the T1-T3 framework is the heterogeneity in the types of evidence generated across the evaluation pipeline, often tied to the predominant chatbot architecture at each stage. Foundational bench testing (T1), where LLM-based studies are most prevalent, typically yields evidence related to technical performance – such as conversational coherence, linguistic accuracy, or safety in controlled tests. As studies progress to T2 pilot feasibility testing, the focus shifts towards usability, engagement, and user acceptance, evaluated across diverse groups including general users and patients. Rule-based studies are more common here (58%) than LLM-based studies (24%). Evidence of clinically meaningful impact, such as symptom reduction measured by validated scales over time, is primarily generated in T3 clinical efficacy testing, and rule-based systems are the main architecture assessed at this highest tier (65%).

The current concentration of LLM research in T1 and T2 stages means that these advanced models are often validated based on technical feasibility or short-term user experience metrics, rather than demonstrated therapeutic effectiveness. This disparity underscores a crucial limitation: strong performance in T1 stage or positive user feedback in T2 stage does not necessarily translate to T3 clinical efficacy. Furthermore, while T2 studies often specify diverse participant groups, T3 evaluations show less consistency in reporting participant characteristics, sometimes hindering the assessment of real-world applicability and long-term impact.

Ethical, safety and regulatory concerns are becoming increasingly critical as chatbots move closer to clinical deployment. LLM-based systems introduce considerable risks, including potentially greater data privacy violations than rule-based or machine learning-driven systems, algorithmic bias, and the potential for "hallucinations" (i.e., false or misleading responses) that could lead to harmful advice. All this may be exacerbated by the richer and more in-depth conversational capabilities of LLMs, which could encourage users to disclose more sensitive information[193]. Unlike rule-based chatbots, which are constrained to predefined responses, LLMs rely on large, uncurated datasets, making them susceptible to misinformation. These risks are not hypothetical: real-world examples, such as *Replika*'s backlash for generating inappropriate responses[194], illustrate the consequences of insufficient safeguards in generative AI systems[195]. While rule-based chatbots mitigate some of these risks through structured outputs, they lack the adaptive empathy needed for sustained mental health support. Machine learning-based systems occupy an intermediary position, balancing adaptability with limited generative capacity but often struggling with transparency and interpretability.

Addressing these risks requires regulatory bodies to establish clear certification pathways for LLM-driven chatbots, ensuring that innovations in generative AI are balanced with accountability and user safety. However, given the widespread availability and increasing use of base models (e.g., *GPT*[8], *Gemini*[196], *Claude*[197], *Llama*[20], *DeepSeek*[198]) for mental health applications, this alone may not be sufficient. An urgent research avenue is to independently establish the safety and clinical utility of these models, ideally through rapid, automated and repeatable evaluations as their capabilities continue to evolve.

Moving forward, mental health chatbot research must prioritize rigorous clinical efficacy trials for LLM-based systems, ensuring that chatbots progress beyond foundational and feasibility testing to real-world clinical validation. The development of standardized clinical endpoints, transparency in chatbot architectures, and regulatory alignment with AI-driven mental health tools will be essential in bridging the gap between feasibility and efficacy. As chatbots continue to evolve, robust validation methodologies will be necessary to ensure that they serve as effective, ethical, and clinically reliable tools for global mental health care.

CONCLUSIONS

Mental health chatbots have rapidly evolved from deterministic rule-based systems to sophisticated LLMs, signalling a transformative shift in digital psychiatry. Despite this promising advancement, our systematic analysis highlights a fragmented landscape with limited rigorous clinical validation, particularly concerning generative AI technologies. The proposed three-tier classification system clarifies evaluation rigor and reveals that most LLM-based interventions remain in early development phases.

Future research should prioritize rigorous clinical efficacy trials, transparent reporting of chatbot architecture, and ethical evaluations, to ensure that these technologies reliably enhance mental health care. Clinicians and policy makers must distinguish between marketing claims and technical realities, advocating for evidence-based standards analogous to established medical AI certification processes.

ACKNOWLEDGEMENTS

J. Torous is supported by the Argosy Foundation and Shifting Gears Foundation. Y. Hua and S. Siddals are joint first authors of this paper. Supplementary information on this study is available at https://osf.io/tcwv8/wiki/home.

REFERENCES

1. James SL, Abate D, Abate KH et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195

countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018;392:1789-858.
2. World Health Organization. Depression and other common mental disorders. www.who.int.
3. World Health Organization. Mental health. www.who.int.
4. World Health Organization. Global health workforce statistics database. www.who.int.
5. Patel V, Saxena S, Lund C et al. The Lancet Commission on global mental health and sustainable development. Lancet 2018;392:1553-98.
6. Abd-Alrazaq AA, Rababeh A, Alajlani M et al. Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. J Med Internet Res 2020;22:e16021.
7. Brown TB, Mann B, Ryder N et al. Language models are few-shot learners. In: Proceedings of the 34th International Conference on Neural Information Processing Systems, 2020:1877-901.
8. Ayers JW, Poliak A, Dredze M et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023;183:589-96.
9. Rajpurkar P, Topol EJ. A clinical certification pathway for generalist medical AI systems. Lancet 2025;405:20.
10. Weizenbaum J. ELIZA – a computer program for the study of natural language communication between man and machine. Communications of the ACM 1966;9:36-45.
11. Darcy A. Why generative AI is not yet ready for mental healthcare. Woebot Health, March 1, 2023.
12. Parmar P, Ryu J, Pandya S et al. Health-focused conversational agents in person-centered care: a review of apps. NPJ Digit Med 2022;5:1-9.
13. Yeh PL, Kuo WC, Tseng BL et al. Does the AI-driven chatbot work? Effectiveness of the Woebot app in reducing anxiety and depression in group counseling courses and student acceptance of technological aids. Curr Psychol 2025;44:8133-45.
14. Bishop CM. Pattern recognition and machine learning. New York: Springer, 2006.
15. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436-44.
16. Mikolov T, Sutskever I, Chen K et al. Distributed representations of words and phrases and their compositionality. Adv Neural Inf Process Syst 2013;26:3111-9.
17. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health 2017;4:e19.
18. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR mHealth uHealth 2018;6:e12106.
19. Vaswani A, Shazeer N, Parmar N et al. Attention is all you need. Adv Neural Inf Process Syst 2017;30:5998-6008.
20. Touvron H, Lavril T, Izacard G et al. LLaMA: open and efficient foundation language models. arXiv 2023;230213971.
21. Hatch SG, Goodman ZT, Vowels L et al. When ELIZA meets therapists: a Turing test for the heart and mind. PLoS Ment Health 2025;2:e0000145.
22. Na H, Hua Y, Wang Z et al. A survey of large language models in psychotherapy: current landscape and future directions. arXiv 2025;2502.11095.
23. Bender EM, McMillan-Major A, Shmitchell S et al. On the dangers of stochastic parrots: can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 2021:610-23.
24. Page MJ, McKenzie JE, Bossuyt PM et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:71.
25. Gabrielli S, Rizzi S, Carbone S et al. A chatbot-based coaching intervention for adolescents to promote life skills: pilot study. JMIR Hum Factors 2020;7:e16762.
26. Denecke K, Vaaheesan S, Arulnathan A. A mental health chatbot for regulating emotions (SERMO) – concept and usability test. IEEE Trans Emerg Topics Comput 2020;9:1170-82.
27. Hauser-Ulrich S, Künzli H, Meier-Peterhans D et al. A smartphone-based health care chatbot to promote self-management of chronic pain (SELMA): pilot randomized controlled trial. JMIR mHealth uHealth 2020;8:e15806.
28. Linden B, Tam-Seto L, Stuart H. Adherence of the Here4U App - Military Version to criteria for the development of rigorous mental health apps. JMIR Form Res 2020;4:e18890.
29. Gaffney H, Mansell W, Tai S. Agents of change: understanding the therapeutic processes associated with the helpfulness of therapy for mental health problems with relational agent MYLO. Digit Health 2020;6:2055207620911580.
30. Dosovitsky G, Pineda BS, Jacobson NC et al. Artificial intelligence chatbot for depression: descriptive study of usage. JMIR Form Res 2020;4:e17065.
31. Kawasaki M, Yamashita N, Lee YC et al. Assessing users' mental status from their journaling behavior through chatbots. Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, 2020:32.
32. Lee YC, Yamashita N, Huang Y. Designing a chatbot as a mediator for promoting deep self-disclosure to a real mental health professional. Proceedings of the ACM Conference on Human-Computer Interaction, 2020;4:31.
33. de Gennaro M, Krumhuber EG, Lucas G. Effectiveness of an empathic chatbot in combating adverse effects of social exclusion on mood. Front Psychol 2020;10:3061.
34. De Nieva JO, Joaquin JA, Tan CB et al. Investigating students' use of a mental health chatbot to alleviate academic stress. Proceedings of the 6th International ACM In-Cooperation HCI and UX Conference, 2020:1-10.
35. Daley K, Hungerbuehler I, Cavanagh K et al. Preliminary evaluation of the engagement and effectiveness of a mental health chatbot. Front Digit Health 2020;2:576361.
36. Narain J, Quach T, Davey M et al. Promoting wellbeing with Sunny, a chatbot that facilitates positive messages within social groups. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020:1-8.
37. Ryu H, Kim S, Kim D et al. Simple and steady interactions win the healthy mentality: designing a chatbot service for the elderly. Proceedings of the ACM Conference on Human-Computer Interaction, 2020;4:152.
38. Bennion MR, Hardy GE, Moore RK et al. Usability, acceptability, and effectiveness of web-based conversational agents to facilitate problem solving in older adults: controlled study. J Med Internet Res 2020;22:e16794.
39. Park S, Thieme A, Han J et al. "I wrote as if I were telling a story to someone I knew": designing chatbot interactions for expressive writing in mental health. Proceeding of the ACM Designing Interactive Systems Conference, 2021:926-41.
40. Williams R, Hopkins S, Frampton C et al. 21-Day Stress Detox: open trial of a universal well-being chatbot for young adults. Soc Sci 2021;10:416.
41. Chung K, Cho HY, Park JY. A chatbot for perinatal women's and partners' obstetric and mental health care: development and usability evaluation study. JMIR Med Inform 2021;9:e18607.
42. Loveys K, Sagar M, Pickering I et al. A digital human for delivering a remote loneliness and stress intervention to at-risk younger and older adults during the COVID-19 pandemic: randomized pilot trial. JMIR Ment Health 2021;8:e31586.
43. Prochaska JJ, Vogel EA, Chieng A et al. A therapeutic relational agent for reducing problematic substance use (Woebot): development and usability study. J Med Internet Res 2021;23:e24850.
44. Oliveira ALS, Matos LN, Junior MC et al. An initial assessment of a chatbot for rumination-focused cognitive behavioral therapy (RFCBT) in college students. Proceeding of the International Conference on Computational Science and Its Applications, 2021:549-64.
45. Klos MC, Escoredo M, Joerin A et al. Artificial intelligence-based chatbot for anxiety and depression in university students: pilot randomized controlled trial. JMIR Form Res 2021;5:e20678.
46. Ellis-Brush K. Augmenting coaching practice through digital methods. Int J Evid Based Coach Mentor 2021;Spec. Issue 15:187-97.
47. Dosovitsky G, Bunge EL. Bonding with Bot: user feedback of a chatbot for social isolation. Front Digit Health 2021;3:735053.
48. Valtolina S, Hu L. Charlie: a chatbot to improve the elderly quality of life and to make them more active to fight their sense of loneliness. Proceedings of the 14th Biannual Conference of the Italian SIGCHI Chapter, 2021:19.
49. Hungerbuehler I, Daley K, Cavanagh K et al. Chatbot-based assessment of employees' mental health: design process and pilot implementation. JMIR Form Res 2021;5:e21678.
50. Lavelle J, Dunne N, Mulcahy HE et al. Chatbot-delivered cognitive defusion versus cognitive restructuring for negative self-referential thoughts: a pilot study. Psychol Rec 2021;72:247-61.
51. Kaywan P, Ahmed K, Miao Y et al. DEPRA: an early depression detection analysis chatbot. Proceedings of the International Conference on Health Information Science, 2021:193-204.
52. Beilharz F, Sukunesan S, Rossell SL et al. Development of a positive body image chatbot (KIT) with young people and parents/carers: qualitative focus group study. J Med Internet Res 2021;23:e27807.
53. Narynov S, Zhumanov Z, Gumar A et al. Development of chatbot psychologist applying natural language understanding techniques. Proceedings of the 21st International Conference on Control, Automation and Systems, 2021:636-41.
54. Romanovskyi O, Pidbutska N, Knysh A. Elomia chatbot: the effectiveness of artificial intelligence in the fight for mental health. Proceedings of the 5th In-

ternational Conference on Computational Linguistics and Intelligent Systems, 2021:1215-24.
55. van Cuylenburg HC, Ginige T. Emotion Guru: a smart emotion tracking application with AI conversational agent for exploring and preventing depression. Proceedings of the International Conference on UK-China Emerging Technologies, 2021:1-6.
56. Mariamo A, Temcheff CE, Léger PM et al. Emotional reactions and likelihood of response to questions designed for a mental health chatbot among adolescents: experimental study. JMIR Hum Factors 2021;8:e24343.
57. Gabrielli S, Rizzi S, Bassi G et al. Engagement and effectiveness of a healthy-coping intervention via chatbot for university students during the COVID-19 pandemic: mixed methods proof-of-concept study. JMIR mHealth uHealth 2021;9:e27965.
58. Darcy A, Daniels J, Salinger D et al. Evidence of human-level bonds established with a digital conversational agent: cross-sectional, retrospective observational study. JMIR Form Res 2021;5:e27868.
59. Bendig E, Erb B, Meißner E et al. Feasibility of a software agent providing a brief Intervention for Self-help to Uplift psychological wellbeing ("SISU"). A single-group pretest-posttest trial investigating the potential of SISU to act as therapeutic agent. Internet Interv 2021;24:100377.
60. Sia DE, Yu MJ, Daliva JL et al. Investigating the acceptability and perceived effectiveness of a chatbot in helping students assess their well-being. Proceedings of the Asian CHI Symposium, 2021:34-40.
61. Rakib AB, Rumky EA, Ashraf AJ et al. Mental healthcare chatbot using sequence-to-sequence learning and BiLSTM. Brain Informatics Lecture Notes in Computer Science, 2021:378-87.
62. Jang S, Kim JJ, Kim SJ et al. Mobile app-based chatbot to deliver cognitive behavioral therapy and psychoeducation for adults with attention deficit: a development and feasibility/usability study. Int J Med Inform 2021;150:104440.
63. Dosovitsky G, Kim E, Bunge EL. Psychometric properties of a chatbot version of the PHQ-9 for adults and older adults. Front Digit Health 2021;3:645805.
64. Lee J, Liang B, Fong H. Restatement and question generation for counsellor chatbot. Proceedings of the 1st Workshop on NLP for Positive Impact, 2021:1-7.
65. Salhi I, Guemmat KE, Qbadou M et al. Towards developing a pocket therapist: an intelligent adaptive psychological support chatbot against mental health disorders in a pandemic situation. Indones J Electr Eng Comput Sci 2021;23:1200.
66. Kraus M, Seldschopf P, Minker W. Towards the development of a trustworthy chatbot for mental health applications. Proceedings of the 27th International Conference on MultiMedia Modeling, 2021:354-66.
67. Goonesekera Y, Donkin L. A cognitive behavioral therapy chatbot (Otis) for health anxiety management: mixed methods pilot study. JMIR Form Res 2022;6:e37877.
68. Leo AJ, Schuelke MJ, Hunt DM et al. A digital mental health intervention in an orthopedic setting for patients with symptoms of depression and/or anxiety: feasibility prospective cohort study. JMIR Form Res 2022;6:e34889.
69. Beredo JL, Ong EC. A hybrid response generation model for an empathetic conversational agent. International Conference on Asian Language Processing, 2022:300-5.
70. Rathnayaka P, Mills N, Burnett D et al. A mental health chatbot with cognitive skills for personalised behavioural activation and remote health monitoring. Sensors 2022;22:3653.
71. Vertsberger D, Naor N, Winsberg M. Adolescents' well-being while using a mobile artificial intelligence-powered acceptance commitment therapy tool: evidence from a longitudinal study. JMIR AI 2022;1:e38171.
72. Liu I, Xiao Y, Liu F et al. Assessing the effectiveness of using chatbots for positive psychological intervention: a randomized control study. Proceedings of the 10th International Symposium of Chinese CHI, 2022:227-34.
73. Medeiros L, Bosse T, Gerritsen C. Can a chatbot comfort humans? Studying the impact of a supportive chatbot on users' self perceived stress. IEEE Trans Hum Mach Syst 2022;52:343-53.
74. Nicol G, Wang R, Graham S et al. Chatbot-delivered cognitive behavioral therapy in adolescents with depression and anxiety during the COVID-19 pandemic: feasibility and acceptability study. JMIR Form Res 2022;6:e40242.
75. Nie J, Shao H, Zhao M et al. Conversational AI therapist for daily function screening in home environments. Proceedings of the 1st ACM International Workshop on Intelligent Acoustic Systems and Applications, 2022:31-6.
76. Collins C, Arbour S, Beals N et al. Covid Connect: chat-driven anonymous story-sharing for peer support. Proceedings of the ACM Designing Interactive Systems Conference, 2022:301-18.
77. Harshini Raji VP, Maheswari PU. COVID-19 lockdown in India: an experimental study on promoting mental wellness using a chatbot during the coronavirus. Int J Ment Health Prom 2022;24:189-205.
78. Maeng W, Lee J. Designing and evaluating a chatbot for survivors of image-based sexual abuse. Proceedings of the CHI Conference on Human Factors in Computing, 2022:344.
79. Shah J, DePietro B, D'Adamo L et al. Development and usability testing of a chatbot to promote mental health services use among individuals with eating disorders following screening. Int J Eat Disord 2022;55:1229-44.
80. Trappey AJC, Lin APC, Hsu KYK et al. Development of an empathy-centric counseling chatbot system capable of sentimental dialogue analysis. Processes 2022;10:930.
81. Leo AJ, Schuelke MJ, Hunt DM et al. Digital mental health intervention plus usual care compared with usual care only and usual care plus in-person psychological counseling for orthopedic patients with symptoms of depression or anxiety: cohort study. JMIR Form Res 2022;6:e36203.
82. Fitzsimmons-Craft EE, Chan WW, Smith et al. Effectiveness of a chatbot for eating disorders prevention: a randomized clinical trial. Int J Eat Disord 2022;55:343-53.
83. Liu I, Chen W, Ge Q et al. Enhancing psychological resilience with chatbot-based cognitive behavior therapy: a randomized control pilot study. Proceedings of the 10th International Symposium of Chinese CHI, 2022:216-21.
84. Beatty C, Malik T, Meheli S et al. Evaluating the therapeutic alliance with a free-text CBT conversational agent (Wysa): a mixed-methods study. Front Digit Health 2022;4.
85. He Y, Yang L, Zhu X et al. Mental health chatbot for young adults with depression symptoms during the COVID-19 pandemic: a single-blind, three-arm, randomized controlled trial. J Med Internet Res 2022;24:e40719.
86. Santa-Cruz J, Moran L, Tovar M et al. Mobilizing digital technology to implement a population-based psychological support response during the COVID-19 pandemic in Lima, Peru. Glob Ment Health 2022;9:355-65.
87. Boyd K, Potts C, Bond R et al. Usability testing and trust analysis of a mental health and wellbeing chatbot. Proceedings of the 33rd European Conference on Cognitive Ergonomics, 2022:18.
88. Burger F, Neerincx MA, Brinkman WP. Using a conversational agent for thought recording as a cognitive therapy task: feasibility, content, and feedback. Front Digit Health 2022;4.
89. Liu H, Peng H, Song X et al. Using AI chatbots to provide self-help depression interventions for university students: a randomized trial of effectiveness. Internet Interv 2022;27:100495.
90. Schick A, Feine J, Morana S et al. Validity of chatbot use for mental health assessment: experimental study. JMIR mHealth uHealth 2022:e28082.
91. Todorov M, Avramova-Todorova G, Dimitrova et al. Virtual assisted technologies as a helping tool for therapists in assessment of anxiety. Outcomes of a pilot trial with chatbot assistance. Proceedings of the International Symposium on Bioinformatics and Biomedicine, 2022:60-6.
92. Sanabria G, Greene KY, Tran JT et al. "A great way to start the conversation": evidence for the use of an adolescent mental health chatbot navigator for youth at risk of HIV and other STIs. J Technol Behav Sci 2023;8:382-91.
93. Chung LL, Kang J. "I'm hurt too": the effect of a chatbot's reciprocal self-disclosures on users' painful experiences. Arch Design Res 2023;36:67-85.
94. Sabour S, Zhang W, Xiao X et al. A chatbot for mental health support: exploring the impact of Emohaa on reducing mental distress in China. Front Digit Health 2023;5.
95. Escobar-Viera CG, Porta G, Coulter RW et al. A chatbot-delivered intervention for optimizing social media use and reducing perceived isolation among rural-living LGBTQ+ youth: development, acceptability, usability, satisfaction, and utility. Internet Interv 2023;34:100668.
96. Brinsley J, Singh B, Maher CA. A digital lifestyle program for psychological distress, wellbeing and return-to-work: a proof-of-concept study. Arch Phys Med Rehabil 2023;104:1903-12.
97. Potts C, Lindstrm F, Bond R et al. A multilingual digital mental health and well-being chatbot (ChatPal): pre-post multicenter intervention study. J Med Internet Res 2023;25:e43051.
98. Alonso-Mencía J, Castro-Rodríguez M, Herrero-Pinilla B et al. ADELA: a conversational virtual assistant to prevent delirium in hospitalized older persons. J Supercomput 2023;79:17670-90.
99. Gabor-Siatkowska K, Sowanski M, Rzatkiewicz R et al. AI to train AI: using ChatGPT to improve the accuracy of a therapeutic dialogue system. Electronics 2023;12:4694.
100. Wrightson-Hester AR, Anderson G, Dunstan J et al. An artificial therapist (Manage Your Life Online) to support the mental health of youth: co-design and case series. JMIR Hum Factors 2023;10:e46849.
101. Omarov B, Zhumanov Z, Gumar A et al. Artificial intelligence enabled mobile chatbot psychologist using AIML and cognitive behavioral therapy. Int J Adv

Comput Sci Appl 2023;14.
102. Viduani A, Cosenza V, Fisher HL et al. Assessing mood with the identifying depression early in adolescence chatbot (IDEABot): development and implementation study. JMIR Hum Factors 2023;10:e44388.
103. Ollier J, Suryapalli P, Fleisch E et al. Can digital health researchers make a difference during the pandemic? Results of the single-arm, chatbot-led Elena+: care for COVID-19 interventional study. Front Public Health 2023;11:1185702.
104. Yorita A, Egerton S, Chan C et al. Chatbot persona selection methodology for emotional support. Proceedings of the 62nd Annual Conference of the Society of Instrument and Control Engineers, 2023:333-8.
105. Lee CH, Liaw GH, Yang WC et al. Chatbot-assisted therapy for patients with methamphetamine use disorder. Front Psychiatry 2023;14:1159399.
106. Selaskowski B, Reiland M, Schulze M et al. Chatbot-supported psychoeducation in adult attention-deficit hyperactivity disorder: randomised controlled trial. BJPsych Open 2023;9:e192.
107. Sezgin E, Chekeni F, Lee J et al. Clinical accuracy of large language models and Google search responses to postpartum depression questions: cross-sectional study. J Med Internet Res 2023;25:e49240.
108. Kang E, Kang YA. Counseling chatbot design: the effect of anthropomorphic chatbot characteristics on user self-disclosure and companionship. Int J Hum-Comput Interact 2023;40:2781-95.
109. Lee M, Contreras Alejandro J, IJsselsteijn W. Cultivating gratitude with a chatbot. Int J Hum-Comput Interact 2023;40:4957-72.
110. Islam A, Chaudhry BM. Design validation of a relational agent by COVID-19 patients: mixed methods study. JMIR Hum Factors 2023;10:e42740.
111. Dosovitsky G, Bunge E. Development of a chatbot for depression: adolescent perceptions and recommendations. Child Adolesc Ment Health 2023;28:124-7.
112. Park G, Chung J, Lee S. Effect of AI chatbot emotional disclosure on user satisfaction and reuse intention for mental health counseling: a serial mediation model. Curr Psychol 2023;42:28663-73.
113. Mishra K, Priya P, Burja M et al. e-THERAPIST: I suggest you to cultivate a mindset of positivity and nurture uplifting thoughts. Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2023:13952-67.
114. Iglesias M, Sinha C, Vempati R et al. Evaluating a digital mental health intervention (Wysa) for workers' compensation claimants: pilot feasibility study. J Occup Environ Med 2023;65:e93-9.
115. Kang A, Hetrick S, Cargo T et al. Exploring young adults views about Aroha, a chatbot for stress associated with the COVID-19 pandemic: interview study among students. JMIR Form Res 2023;7:e44556.
116. Suharwardy S, Ramachandran M, Leonard SA et al. Feasibility and impact of a mental health chatbot on postpartum mental health: a randomized controlled trial. AJOG Glob Rep 2023;3:100165.
117. Bird JJ, Lotfi A. Generative transformer chatbots for mental health support: a study on depression and anxiety. Proceedings of the 16th International Conference on PErvasive Technologies Related to Assistive Environments, 2023:475-9.
118. Sharma A, Lin IW, Miner AS et al. Human-AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. Nat Mach Intell 2023;5:46-57.
119. Han HJ, Mendu S, Jaworski BK et al. Preliminary evaluation of a conversational agent to support self-management of individuals living with posttraumatic stress disorder: interview study with clinical experts. JMIR Form Res 2023;7:e45894.
120. Rodriguez-Martinez A, Amezcua-Aguilar T, Cortés-Moreno J et al. Qualitative analysis of conversational chatbots to alleviate loneliness in older adults as a strategy for emotional health. Healthcare 2023;12:62.
121. Heston TF. Safety of large language models in addressing depression. Cureus 2023;15:e50729.
122. Mancone S, Diotaiuti P, Valente G et al. The use of voice assistant for psychological assessment elicits empathy and engagement while maintaining good psychometric properties. Behav Sci 2023;13:550.
123. Sinha C, Meheli S, Kadaba M. Understanding digital mental health needs and usage with an artificial intelligence-led mental health app (Wysa) during the COVID-19 pandemic: retrospective analysis. JMIR Form Res 2023;7:e41913.
124. Ma Z, Mei Y, Su Z. Understanding the benefits and challenges of LLM-based agents for mental well-being. AMIA Annu Symp Proc 2023;2023:1105-14.
125. Inkster B, Kadaba M, Subramanian V. Understanding the impact of an AI-enabled conversational agent mobile app on users' mental health and well-being with a self-reported maternal event: a mixed method real-world data mHealth study. Front Glob Womens Health 2023;4:1084302.
126. Matheson EL, Smith HG, Amaral ACS et al. Using chatbot technology to improve Brazilian adolescents' body image and mental health at scale: randomized controlled trial. JMIR mHealth uHealth 2023;11:e39934.
127. Anmella G, Sanabra M, Primé-Tous M et al. Vickybot, a chatbot for anxiety-depressive symptoms and work-related burnout in primary care and health care professionals: development, feasibility, and potential effectiveness studies. J Med Internet Res 2023;25:e43293.
128. Ge Q, Liu L, Zhang H et al. Designing Philobot: a chatbot for mental health support with CBT techniques. Proceedings of the Chinese Intelligent Automation Conference, 2023:361-71.
129. Siddals S, Torous J, Coxon A. "It happened to be the perfect thing": experiences of generative AI chatbots for mental health. Npj Ment Health Res 2024;3:48.
130. da Costa FL, Matzenbacher LS, Maia IS et al. "She has become my best friend": a qualitative study on the perspective of elderly with type 2 diabetes regarding the use of an interactive virtual assistant device for diabetes care and mental health promotion. Acta Diabetol 2024; doi: 10.1007/s00592-024-02377-z.
131. Mancinelli E, Magnolini S, Gabrielli S et al. A chatbot (Juno) prototype to deploy a behavioral activation intervention to pregnant women: qualitative evaluation using a multiple case study. JMIR Form Res 2024;8:e58653.
132. Yasukawa S, Tanaka T, Yamane K et al. A chatbot to improve adherence to internet-based cognitive-behavioural therapy among workers with subthreshold depression: a randomised controlled trial. BMJ Ment Health 2024;27:e300881.
133. Ulrich S, Lienhard N, Künzli H et al. A chatbot-delivered stress management coaching for students (MISHA app): pilot randomized controlled trial. JMIR mHealth uHealth 2024;12:e54945.
134. Boian R, Bucur AM, Todea D et al. A conversational agent framework for mental health screening: design, implementation, and usability. Behav Inform Technol 2024; doi: 10.1080/0144929X.2024.2332934.
135. Allan DD. A mobile messaging-based conversational agent-led stress mindset intervention for New Zealand small-to-medium-sized enterprise owner-managers: effectiveness and acceptability study. Behav Inf Technol 2024;43:2785-98.
136. Shidara K, Tanaka H, Adachi H et al. Adapting the number of questions based on detected psychological distress for cognitive behavioral therapy with an embodied conversational agent: comparative study. JMIR Form Res 2024;8:e50056.
137. So J-H, Chang J, Kim E et al. Aligning large language models for enhancing psychiatric interviews through symptom delineation and summarization: pilot study. JMIR Form Res 2024;8:e58418.
138. Foran HM, Kubb C, Mueller J et al. An automated conversational agent self-help program: randomized controlled trial. J Med Internet Res 2024;26:e53829.
139. Marmol-Romero AM, Garcia-Vega M, Garcia-Cumbreras M et al. An empathic GPT-based chatbot to talk about mental disorders with Spanish teenagers. Int J Hum-Comput Interact 2024; doi: 10.1080/10447318.2024.2344355.
140. Ruggiano N, Brown EL, Clarke PJ et al. An evidence-based IT program with chatbot to support caregiving and clinical care for people with dementia: the CareHeroes development and usability pilot. JMIR Aging 2024;7:e57308.
141. Han HJ, Mendu S, Jaworski BK et al. Assessing acceptance and feasibility of a conversational agent to support individuals living with post-traumatic stress disorder. Digit Health 2024;10:20552076241286133.
142. Maurya RK, Montesinos S, Bogomaz M et al. Assessing the use of ChatGPT as a psychoeducational tool for mental health practice. Couns Psychother Res 2024; doi: 10.1002/capr.12759.
143. Heissler R, Jonaa J, Carre N et al. Can AI digital personas for well-being provide social support? A mixed-method analysis of user reviews. Hum Behav Emerg Technol 2024;2024:6738001.
144. Hodson N, Williamson S. Can large language models replace therapists? Evaluating performance at simple cognitive behavioral therapy tasks. JMIR AI 2024;3:e52500.
145. Li Y, Chung TY, Lu W et al. Chatbot-based mindfulness-based stress reduction program for university students with depressive symptoms: intervention development and pilot evaluation. J Am Psychiatr Nurses Assoc 2024; doi: 10.1177/10783903241302092.
146. Striegl J, Fekih F, Weber G et al. Chatbot-based mood and activity journaling for resource-oriented CBT support of students. Proceedings of the International Conference on Human-Computer Interaction, 2024:177-88.
147. Dergaa I, Fekih-Romdhane F, Hallit S et al. ChatGPT is not ready yet for use in providing mental health assessment and interventions. Front Psychiatry 2024;14:1277756.
148. Gargari OK, Habibi G, Nilchian N et al. Comparative analysis of large language models in psychiatry and mental health: a focus on GPT, AYA, and Nemotron-3-8B. Asian J Psychiatr 2024;99:104148.
149. Liu T, Zhao H, Liu Y et al. ComPeer: a generative conversational agent for pro-

active peer support. Proceedings of the 37th ACM Symposium on User Interface Software and Technology, 2024:1-22.
150. Aghakhani S, Rousseau A, Mizrahi S et al. Conversational agent utilization patterns of individuals with autism spectrum disorder. J Technol Behav Sci 2024; doi: 10.1007/s41347-024-00451-5.
151. Ferrández A, Lavigne-Cerván R, Peral J et al. CuentosIE: can a chatbot about "tales with a message" help to teach emotional intelligence? PeerJ Comput Sci 2024;10:e1866.
152. Chiauzzi E, Williams A, Mariano TY et al. Demographic and clinical characteristics associated with anxiety and depressive symptom outcomes in users of a digital mental health intervention incorporating a relational agent. BMC Psychiatry 2024;24:79.
153. Raut R, Kolambe S, Borkar P et al. Depression therapy chat-bot using natural language processing. Int J Intell Syst Appl Eng 2024;12:181-7.
154. Nasiri Y, Fulda N. Designing a language-model-based chatbot that considers user's personality profile and emotions to support caregivers of people with dementia. Proceedings of the 1st International OpenKG Workshop: Large Knowledge-Enhanced Models, 2024:59-70.
155. Abilkaiyrkyzy A, Laamarti F, Hamdi M et al. Dialogue system for early mental illness detection: toward a digital twin solution. IEEE Access 2024;12:2007-24.
156. Kleinau EF, Lamba T, Jaskiewicz W et al. Effectiveness of a chatbot in improving the mental wellbeing of health workers in Malawi during the COVID-19 pandemic: a randomized, controlled trial. PLoS One 2024;19:e0303370.
157. MacNeill AL, Doucet S, Luke A. Effectiveness of a mental health chatbot for people with chronic diseases: randomized controlled trial. JMIR Form Res 2024;8:e50025.
158. Karkosz S, Szymanski R, Sanna K et al. Effectiveness of a web-based and mobile therapy chatbot on anxiety and depressive symptoms in subclinical young adults: randomized controlled trial. JMIR Form Res 2024;8:e47960.
159. Chew HSJ, Chew NW, Loong SSE et al. Effectiveness of an artificial intelligence-assisted app for improving eating behaviors: mixed methods evaluation. J Med Internet Res 2024;26:e46036.
160. Vereschagin M, Wang AY, Richardson CG et al. Effectiveness of the Minder mobile mental health and substance use intervention for university students: randomized controlled trial. J Med Internet Res 2024;26:e54287.
161. Schillings C, Meißner E, Erb B et al. Effects of a chatbot-based intervention on stress and health-related parameters in a stressed sample: randomized controlled trial. JMIR Ment Health 2024;11:e50454.
162. Fitzsimmons-Craft EE, Rackoff GN, Shah J et al. Effects of chatbot components to facilitate mental health services use in individuals with eating disorders following online screening: an optimization randomized controlled trial. Int J Eat Disord 2024;57:2204-16.
163. Yeo YH, Peng Y, Mehra M et al. Evaluating for evidence of sociodemographic bias in conversational AI for mental health support. Cyberpsychol Behav Soc Netw 2024; doi: 10.1089/cyber.2024.0199.
164. Sinha C, Dinesh D, Phang YS et al. Examining a brief web and longitudinal app-based intervention [Wysa] for mental health support in Singapore during the COVID-19 pandemic: mixed-methods retrospective observational study. Front Digit Health 2024;6:1443598.
165. Bowman R, Cooney O, Newbold JW et al. Exploring how politeness impacts the user experience of chatbots for mental health support. Int J Hum Comput Stud 2024;184:103181.
166. Kim Y, Kang Y, Kim B et al. Exploring the role of engagement and adherence in chatbot-based cognitive training for older adults: memory function and mental health outcomes. Behav Inform Technol 2024; doi: 10.1080/0144929x.2024.2362406.
167. Sharma A, Rushton K, Lin IW et al. Facilitating self-guided mental health interventions through human-language model interaction: a case study of cognitive restructuring. Proceedings of the CHI Conference on Human Factors in Computing Systems, 2024:1-29.
168. Berrezueta-Guzman S, Kandil M, Martin-Ruiz ML et al. Future of ADHD care: evaluating the efficacy of ChatGPT in therapy enhancement. Healthcare 2024;12:683.
169. Hu Z, Hou H, Ni S. Grow with your AI buddy: designing an LLMs-based conversational agent for the measurement and cultivation of children's mental resilience. Proceedings of the 23rd Annual ACM Interaction Design and Children Conference, 2024:811-7.
170. Park G, Chung J, Lee S. Human vs. machine-like representation in chatbot mental health counseling: the serial mediation of psychological distance and trust on compliance intention. Curr Psychol 2024;43:4352-63.
171. Lee J, Lee J-G, Lee D. Influence of rapport and social presence with an AI psychotherapy chatbot on users' self-disclosure. Int J Hum-Comput Interact 2024;40:1620-31.
172. Liu I, Liu F, Xiao Y et al. Investigating the key success factors of chatbot-based positive psychology intervention with retrieval- and generative pre-trained transformer (GPT)-based chatbots. Int J Hum-Comput Interact 2024;41:341-52.
173. Dinesh DN, Rao MN, Sinha C. Language adaptations of mental health interventions: user interaction comparisons with an AI-enabled conversational agent (Wysa) in English and Spanish. Digit Health 2024; doi: 10.1177/20552076241255616.
174. Qi Y. Pilot quasi-experimental research on the effectiveness of the Woebot AI chatbot for reducing mild depression symptoms among athletes. Int J Hum-Comput Interact 2024;41:452-9.
175. Mori D, Matsumoto K, Kang X et al. SMACS: Stress Management AI Chat System. Proceedings of the 16th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, 2024:167-74.
176. Vossen W, Szymanski M, Verbert K. The effect of personalizing a psychotherapy conversational agent on therapeutic bond and usage intentions. Proceedings of the 29th International Conference on Intelligent User Interfaces, 2024:761-71.
177. Zsoldos I, Tran E, Fournier H et al. The value of a virtual assistant to improve engagement in computerized cognitive training at home: exploratory study. JMIR Rehabil Assist Technol 2024;11:e48129.
178. Andrade-Arenas L, Yactayo-Arias C, Pucuhuayla-Revatta F. Therapy and emotional support through a chatbot. Int J Online Biomed Eng 2024;20:114-30.
179. Thunstrm AO, Carlsen HK, Ali L et al. Usability comparison among healthy participants of an anthropomorphic digital human and a text-based chatbot as a responder to questions on mental health: randomized controlled trial. JMIR Hum Factors 2024;11:e54581.
180. Farrand P, Raue PJ, Ward E et al. Use and engagement with low-intensity cognitive behavioral therapy techniques used within an app to support worry management: content analysis of log data. JMIR mHealth uHealth 2024;12:e47321.
181. Chaudhry BM, Debi HR. User perceptions and experiences of an AI-driven conversational agent for mental health support. Mhealth 2024;10:22-2.
182. Chou YH, Lin C, Lee SH et al. User-friendly chatbot to mitigate the psychological stress of older adults during the COVID-19 pandemic: development and usability study. JMIR Form Res 2024;8:e49462.
183. Deepaisarn S, Imkome E, Wongpatikaseree K et al. Validation of a Thai artificial chatmate designed for cheering up the elderly during the COVID-19 pandemic. F1000Research 2024;11:1411.
184. Kostenius C, Lindstrom F, Potts C et al. Young peoples' reflections about using a chatbot to promote their mental wellbeing in northern periphery areas – a qualitative study. Int J Circumpolar Health 2024;83:2369349.
185. Li H, Zhang R, Lee YC et al. Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. NPJ Digit Med 2023;6:236.
186. Torous J, Myrick K. Aguilera A. The need for a new generation of digital mental health tools to support more accessible, effective and equitable care. World Psychiatry 2023;22:1-2.
187. Torous J, Blease C. Generative artificial intelligence in mental health care: potential benefits and current challenges. World Psychiatry 2024;23:1-2.
188. Galderisi S, Appelbaum PS, Gill N et al. Ethical challenges in contemporary psychiatry: an overview and an appraisal of possible strategies and research needs. World Psychiatry 2024;23:364-86.
189. Lawrence HR, Schneider RA, Rubin SB et al. The opportunities and risks of large language models in mental health. JMIR Ment Health 2024;11:e59479.
190. Evans AC. Letter to the Federal Trade Commission regarding concerns about the perils and unintended consequences to the public resulting from the underregulated development and deceptive deployment of generative AI or enabled technologies. American Psychological Association, December 20, 2024.
191. Wang P. On defining artificial intelligence. J Artif Gen Intell 2019;10:1-37.
192. Boden MA. Gofai. In: Frankish K, Ramsey W (eds). The Cambridge handbook of artificial intelligence. Cambridge: Cambridge University Press, 2014:89-107.
193. Nong P, Ji M. Expectations of healthcare AI and the role of trust: understanding patient views on how AI will impact cost, access, and patient-provider relationships. J Am Med Inform Assoc 2025;32:795-9.
194. Roose K. Can A.I. be blamed for a teen's suicide? New York Times, October 23, 2024.
195. Blease C, Torous J. ChatGPT and mental healthcare: balancing benefits with risks of harms. BMJ Ment Health 2023;26:e300884.
196. Team G, Anil R, Borgeaud S et al. Gemini: a family of highly capable multimodal models. arXiv 2025;231211805.
197. Anthropic. Introducing the next generation of Claude. www.anthropic.com.
198. Deep Seek-AI, Liu A, Feng B et al. DeepSeek-V3 technical report. arXiv 2025;241219437.

DOI:10.1002/wps.21352

The epidemiology of ICD-11 bodily distress disorder and DSM-5 somatic symptom disorder in new large-scale population surveys within the World Mental Health Survey Initiative

Oye Gureje[1,2], Maria Lourdes Rosanna De Guzman[3], Salma M. Khaled[4], Michael Y. Ni[5-7], Peter W. Woodruff[8], Geoffrey M. Reed[9], Corine S.M. Wong[5], Candi M.C. Leung[5], Marilyn Crisostomo[3], Shelby Borowski[10], Irving Hwang[10], Maria V. Petukhova[10], Nancy Sampson[10], Ronald C. Kessler[10]

[1]Department of Psychiatry, University of Ibadan, Ibadan, Nigeria; [2]Department of Psychiatry, Stellenbosch University, Stellenbosch, South Africa; [3]Department of Psychiatry and Behavioral Medicine, University of the Philippines College of Medicine and Philippine General Hospital, Manila, The Philippines; [4]Department of Population Medicine, College of Medicine, Qatar University, Doha, Qatar; [5]School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong SAR, China; [6]State Key Laboratory of Brain and Cognitive Sciences, University of Hong Kong, Hong Kong SAR, China; [7]Urban Systems Institute, University of Hong Kong, Hong Kong SAR, China; [8]School of Medicine and Population Health, University of Sheffield, Sheffield, UK; [9]Department of Psychiatry, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA; [10]Department of Health Care Policy, Harvard Medical School, Boston, MA, USA

Distressing somatic symptoms are common and disabling, but a lack of reliable classification of the underlying disorders has limited our understanding of the extent of their population burden. The new categories of bodily distress disorder (BDD) in the ICD-11 and somatic symptom disorder (SSD) in the DSM-5 were designed to address the fundamental weaknesses of previous conceptualizations, but have important differences in their criteria specifications. Three new large-scale population surveys within the World Mental Health (WMH) Survey Initiative, conducted in socially and culturally diverse settings, provide the opportunity to address questions regarding population prevalence, mental and physical health correlates, and associations with role impairment of BDD and SSD. WMH surveys were carried out in representative household samples of adults in Hong Kong, the Philippines, and Qatar (combined N=18,105 respondents). Multivariable regression analysis examined associations of BDD and SSD with socio-demographic variables, comorbid DSM-5 mental disorders, and chronic physical conditions. Role impairment was assessed by examining the mean number of health-related days out of role (DOR) in the 30 days before the interview, adjusting for socio-demographic variables and comorbidities. The point prevalence across the three settings was 2.0% for BDD, 3.5% for SSD, and 4.1% for either diagnosis. The point prevalence of BDD and especially of SSD was highest in Hong Kong, suggesting a role of cultural and social factors. Females were twice as likely as males to meet the criteria for either disorder. Prevalence increased with age. BDD and SSD were significantly associated with generalized anxiety, panic, post-traumatic stress, major depressive, and bipolar spectrum disorders, and associations were consistently stronger for BDD than SSD. More modest comorbidities were found with common chronic physical conditions (arthritis, asthma, diabetes mellitus, hypertension, and stomach or intestinal ulcer). BDD and SSD were both significantly associated with increased mean DOR after adjusting for comorbid mental disorders and chronic physical conditions, but the adjusted mean DOR was significantly higher in the BDD-only than in the SSD-only subsample (4.7 vs. 3.1, p<0.001). These findings attest to the high public health importance of BDD and SSD. Even though both are not highly prevalent in the community, their co-occurrence with common physical and mental disorders, and the fact that they are associated significantly with role impairment, provide strong reason for clinical attention.

Key words: Bodily distress disorder (BDD), somatic symptom disorder (SSD), World Mental Health Surveys, prevalence, psychiatric comorbidities, chronic physical conditions, days out of role (DOR), culture

(World Psychiatry 2025;24:395–403)

Even though distressing and burdensome somatic symptoms are common[1,2], their classification has been challenging[3,4]. For example, the category of "somatoform disorders" in both the ICD-10 and DSM-IV was criticized for being either too restrictive or too broad, and for lacking reliability[3,5,6]. One major consequence of these perceived shortcomings has been that somatoform disorders were excluded from most large-scale community epidemiological surveys of mental disorders[7]. Researchers conducting studies in health care settings have tended to use other conceptualizations of somatic distress[1,8]. As a result, we lack robust information about the community prevalence of these disorders, even though there is general agreement that persons experiencing burdensome somatic concerns commonly require health service attention, and that quantifying the extent of the population burden of these disorders would be important for policy planning purposes[2].

In response to the problems with the previous classification of these disorders, their defining characteristics underwent extensive revision in the ICD-11 and DSM-5[4,9-11]. Compared to ICD-10 and DSM-IV somatoform disorders, the bodily distress disorder (BDD) construct in the ICD-11 and the somatic symptom disorder (SSD) construct in the DSM-5 are substantially simplified, principally by subsuming several previous categories.

Broadly similar in their conceptualization, the two disorders are however not identical[12]. Both exclude the previous requirement that symptoms are "medically unexplained", given the demonstrated unreliability of this criterion[5], and require the presence of specific psychological and cognitive-behavioral features accompanying the distressing symptoms, such as excessive preoccupation, and anxiety about health or symptoms[13]. In persons with an "established medical condition that may be causing or contributing to the symptoms", the ICD-11 Clinical Descriptions and Diagnostic Requirements (CDDR)[11] request, for a diagnosis of BDD, "a degree of attention related to the symptoms (that) is clearly excessive in relation to the nature and severity of the medical condition".

Both systems require that symptoms are persistent, but they differ in the specified duration: for BDD, it is "several months (e.g. three months or more)", while it is "typically more than 6 months"

for SSD. For BDD, excessive attention to the symptoms can be demonstrated by "repeated contacts with health care providers" and its persistence "despite appropriate clinical examination and investigations or appropriate reassurance by health care providers". There is no requirement for clinical help-seeking for SSD.

The diagnosis of BDD also requires the presence of distress and "significant impairment in personal, family, social, educational, occupational or other important areas of functioning". For SSD, the specification is for distress or "significant disruption of daily life". Finally, while the diagnosis of BDD requires that "symptoms or the associated distress and preoccupation" are not better accounted for by another mental disorder, such as an anxiety or mood disorder, the diagnosis of SSD does not require such an exclusion.

A scoping review of studies that have examined the reliability, validity, and clinical utility of the SSD construct provides evidence for its considerable improvement over DSM-IV somatoform disorders[14], strengthened in particular by the inclusion of criteria specification of psychological symptoms[5]. We are aware of only one study reporting the performance of the BDD construct relative to ICD-10 somatoform disorders[15], which was conducted in the context of the development of the ICD-11, and indicated that the use of BDD criteria led to an improvement in clinicians' diagnostic accuracy, and in the clinical utility of the construct. The importance of exploring the epidemiological profiles of these new diagnostic constructs has been repeatedly highlighted[16].

This paper presents data from the World Mental Health (WMH) Survey Initiative on the prevalence and correlates of BDD and SSD. Although WMH surveys have been conducted in close to 30 countries at different times over the past three decades[17], BDD and SSD were only included in the three most recent surveys. These new surveys: a) provide data on the point prevalence of BDD and SSD in three culturally and socially diverse settings; b) allow examination of the associations of these disorders with other common mental disorders and chronic physical health conditions; and c) allow an exploration of the association of BDD and SSD with role impairment both in the presence and absence of comorbid mental and chronic physical disorders.

METHODS

Samples

The WMH Survey Initiative is a coordinated series of community epidemiologic surveys carried out in countries around the world using a consistent methodology in order to make cross-national comparisons of the prevalence and correlates of mental disorders[18-20]. This report uses data from the three most recent WMH surveys, each based on a general population household survey of respondents aged 18 or older.

Two of these surveys were carried out in jurisdictions classified as high-income by the World Bank: a 2022-2024 regional household survey of residents in the Hong Kong Special Administrative Region (SAR) of the People's Republic of China (N=3,053), and a 2019-2022 national phone survey of citizens and Arab expatriates in Qatar (N=5,195). The third survey was carried out in 2021-2022 with a national household sample of the Philippines, a country classified by the World Bank as middle-income (N=9,857).

In Hong Kong, households for the WMH survey were selected randomly from the FAMILY Cohort sample[21] (which is representative of all 18 districts in Hong Kong and has sample sizes proportionate to the population of each district) as well as from households in a supplemental sample of three new towns in the SAR[22].

The Philippines survey was based on a national area probability sample of households selected specifically for the WMH survey. The sample was recruited independently in each of the 17 regions of the country, with the number of respondents in each region selected to be proportional to population size[23].

The Qatar survey was based on a stratified random sample of telephone numbers selected from a national list. This exception to the general WMH area household sampling scheme was dictated by the fact that the survey was initiated shortly before the onset of the COVID-19 pandemic, and then shut down as soon as the lockdown order made it impossible to carry out in-person interviews. The survey was then re-designed for telephone administration, re-initiated in the summer of 2020, and completed in January 2022[24].

In Hong Kong, participants were recruited using a stratified random sampling method. In the Philippines, all eligible adult respondents in each sampled household were interviewed, with a weight used to adjust for differential response rates by household member age and sex. In Qatar, a post-stratification calibration weight was used to adjust the overall distribution of the sample to match the census distribution of the population on socio-demographics. The weighted (by sample size) average response rate across the three surveys was 51.6% using the American Association for Public Opinion Research RR1w definition[25].

At all survey sites, the local ethics or institutional review committees reviewed and approved the protocol to ensure protection of human subjects, in line with appropriate international and local guidelines.

Measures

The Composite International Diagnostic Interview, version 5.0 (CIDI 5.0)[26] was administered by lay interviewers who had undergone extensive standardized training and carried out a series of monitored practice interviews to confirm proficiency in administration. The interview was translated into local languages using a standardized translation, back-translation, and harmonization protocol[27]. Standardized remote quality control monitoring was performed using a field quality software linked to computerized interview schedules[28]. Supervisors also made follow-up assessments with probability subsamples of respondents to repeat certain key questions as checks of interviewer accuracy.

The socio-demographic variables considered in this study include sex, age (18-24, 25-39, 40-54, 55+ years), education (categorized into four levels based on the country-specific education

system)[29], marital status (categorized into three levels: married or cohabitating, previously married, and never married), family income (coded into quartiles of high, high-middle, low-middle, and low, using a country-specific coding schema[30]), and employment status (employed, self-employed, retired, disabled, student, homemaker, and other).

The assessment of SSD and BDD was limited to point prevalence at the time of interview. SSD and BDD were operationalized based on the definitions and criteria in DSM-5 and ICD-11 (see supplementary information). Even though the ICD-11 CDDR[11] specify that a diagnosis of BDD requires a determination that symptoms and their associated distress are not "better accounted for by another mental disorder", this criterion was not operationalized, because of our interest in exploring the associations of SSD and BDD with common mental disorders[31], and determining whether this criterion makes any meaningful difference to the conceptualization of these diagnostic constructs. Otherwise, the diagnostic algorithms were designed to capture the specific requirements of each of the constructs, with particular attention paid to their similarities and differences.

The presence of comorbid mental disorders was assessed by the CIDI 5.0. In this report, we focus on 12-month DSM-5 generalized anxiety disorder (GAD), panic disorder, post-traumatic stress disorder (PTSD), major depressive disorder, bipolar spectrum disorders (including bipolar I disorder, bipolar II disorder, and subthreshold bipolar disorder), and alcohol use disorder. DSM-5 organic exclusion rules were not applied in making these diagnoses, and diagnostic hierarchy rules were not applied other than between major depressive disorder and bipolar spectrum disorders.

Respondents were asked whether they ever seriously thought about suicide and, if so, whether they ever made a suicide attempt. Respondents who reported these lifetime experiences were then asked whether each of them occurred at any time in the past 12 months. Responses were coded yes/no without regard to frequency or intensity.

Chronic physical conditions were assessed by a standard checklist. Checklists of this type have been shown to yield more complete and accurate reports of disorder prevalence than estimates derived from responses to open-ended questions[32,33], and to have moderate to good concordance with medical records[34,35]. In this report, we explored the associations of BDD and SSD with common chronic physical conditions – arthritis, asthma, diabetes mellitus, hypertension, and stomach or intestinal ulcer – along with a summary measure of any other less common conditions.

Role impairment was assessed by a single item from the WMH Survey Initiative version of the WHO Disability Assessment Schedule (WMH WHODAS-II)[36] about days out of role (DOR) due to health problems. The question asked respondents how many days in the past 30 days they were "totally unable to work or carry out their usual daily activities" because of problems with either their physical health, mental health, or use of alcohol or drugs. Good concordance of these reports has been documented with payroll records of employed people[37,38] and prospective daily diary reports[39].

Analysis methods

As noted above, weights were applied to adjust for differences in within-household probabilities of selection and to calibrate the data to match census population distributions on socio-demographic and geographic variables. The Taylor series linearization method implemented in SAS 9.4[40] was used to adjust standard errors for the effects of these weights as well as of geographic clustering.

Cross-tabulations were used to estimate BDD and SSD point prevalence. We then applied univariable and multivariable regression models for the associations of BDD and SSD with socio-demographics, followed by parallel models controlling for socio-demographics that examined associations of comorbid 12-month mental disorders and chronic physical conditions with BDD and SSD. Finally, we explored the joint associations of BDD and SSD with role impairment adjusting for jurisdiction, socio-demographics, comorbid mental disorders, and comorbid physical conditions.

The adjustments were based on the stable balancing weight method[41]. This adjusts for differences in the distributions of covariates (in our case, jurisdiction, socio-demographics and comorbidities) across categories of a primary variable (in our case, a four-category variable for BDD-only, SSD-only, both, and neither) by weighting individual observations in a way that minimizes covariance imbalance across categories of the primary variable while minimizing variance in weights. Our assumption in doing this was that BDD and SSD would be associated significantly with role impairment, but that this association would become smaller once we adjusted for covariates. The other question was whether BDD and SSD would remain associated significantly with role impairment after this adjustment.

All regression models were applied using a logistic link function with robust standard error estimates in SAS 9.4[40]. Regression coefficients for models in which BDD and SSD were dichotomous outcomes were exponentiated to create odds ratios (ORs). Coefficients ±2 design-based standard errors were exponentiated to create design-based 95% confidence intervals (CIs). Significance of OR sets defining a single categorical variable was evaluated by Wald X^2 tests based on design-corrected coefficient variance-covariance matrices. The stable balancing weight adjustment was made using the R 'sbw' package[42]. Statistical significance was evaluated consistently using two-sided design-based 0.05-level tests.

RESULTS

Point prevalence

The point prevalence of BDD across the three settings was 2.0% (ranging from 1.2% in the Philippines to 3.5% in Hong Kong). The point prevalence of SSD was 3.5% (ranging from 2.4% in the Philippines to 7.2% in Hong Kong). The point prevalence of either diagnosis was 4.1% (ranging from 2.8% in the Philippines to 8.0% in Hong Kong) (see Table 1).

Table 1 Point prevalence of bodily distress disorder (BDD) and somatic symptom disorder (SSD)

	Total		Hong Kong		Philippines		Qatar	
	%	(SE)	%	(SE)	%	(SE)	%	(SE)
BDD	2.0	(0.1)	3.5	(0.3)	1.2	(0.2)	3.3	(0.3)
SSD	3.5	(0.2)	7.2	(0.5)	2.4	(0.2)	3.2	(0.4)
Either	4.1	(0.2)	8.0	(0.5)	2.8	(0.2)	4.8	(0.4)
Both	1.4	(0.1)	2.7	(0.2)	0.9	(0.1)	1.7	(0.3)
BDD-only	0.6	(0.1)	0.8	(0.2)	0.3	(0.1)	1.6	(0.2)
SSD-only	2.1	(0.1)	4.5	(0.5)	1.6	(0.1)	1.5	(0.3)
X^2	127.4*		78.1*		82.9*		0.1	
OR (95% CI)	99.0**	(76.0-128.9)	67.9**	(42.0-110.0)	175.5**	(112.8-273.1)	68.6**	(40.4-116.5)

OR – odds ratio, SE – standard error, *significant difference between BDD and SSD at the 0.05 level, design-based X^2 test, **significant OR between the two disorders at the 0.05 level, two-sided design-based test

Point prevalence estimates for SSD were significantly higher than those for BDD in Hong Kong (X^2=78.1, p<0.001) and the Philippines (X^2=82.9, p<0.001), but not in Qatar (X^2=0.1, p=0.73). However, the two diagnoses were highly correlated: OR=99.0 (95% CI: 76.0-128.9) in the total sample (see Table 1).

Socio-demographic correlates

Respondent age was significantly and positively associated with odds of both BDD (X^2=13.4, p=0.004) and SSD (X^2=15.2, p=0.002), as well as with meeting criteria for either (X^2=17.0, p=0.001) or both (X^2=14.0, p=0.003) diagnoses. The ORs for respondents in the youngest age category (ages 18-24) were in the range from 0.2 to 0.5 relative to respondents in the oldest age category (ages 55+) (see Table 2).

Females had significantly higher point prevalence rates than males of either (X^2=22.6, p<0.001) as well as both disorders (X^2= 13.1, p<0.001), with ORs ranging between 1.7 and 2.2. Respondent education (X^2=3.1 to 4.3, p=0.38 to 0.23), marital status (X^2= 0.9 to 2.9, p=0.63 to 0.24), and family income (X^2=1.5 to 3.7, p=0.68 to 0.29) were not associated significantly with either disorder (see Table 2).

Associations with employment status were significant (X^2=22.3 to 40.6, p=0.001 to <0.001), due to extremely high ORs for the retired (7.5 to 9.9) and less consistently elevated ORs for the disabled (1.2 to 1.5) and for respondents in the residual category of "other" employment status (1.9 to 2.2), relative to the employed (see Table 2).

Comorbidity with mental disorders

The 12-month prevalence of anxiety and mood disorders was significantly associated with odds of both BDD and SSD in multivariable models, but with ORs consistently higher for BDD than SSD. Specifically, the ORs of the association with BDD and SSD were, respectively, 3.3 and 1.7 for GAD; 3.0 and 2.6 for panic disorder; 4.1 and 2.4 for PTSD; 7.8 and 4.4 for major depressive disorder; and 5.3 and 3.0 for bipolar spectrum disorders. Consistent with this observation, the ORs of these comorbid disorders with BDD-only (ORs: 2.4 to 9.2) were for the most part significantly larger than those with SSD-only (ORs: 1.5 to 3.3) (see Table 3).

Twelve-month suicide ideation was significantly associated with both BDD and SSD in multivariable models (OR=1.9, 95% CI: 1.0-3.6, and OR=2.9, 95% CI: 1.9-4.3, respectively), but ORs did not differ significantly in predicting BDD-only versus SSD-only. Alcohol use disorder was not significantly associated with either BDD or SSD in multivariable models, despite substantively elevated ORs (OR=2.5, 95% CI: 0.7-9.0 for BDD; OR=2.1, 95% CI: 0.8-5.5 for SSD). Suicide attempt was also not significantly associated with either BDD or SDD (see Table 3).

Comorbidity with chronic physical conditions

All the common chronic physical conditions considered here had elevated ORs in multivariable models predicting both BDD (ORs ranging from 1.4 to 3.5) and SSD (ORs ranging from 1.2 to 3.3). About half the ORs were significant. There was no consistent evidence for the ORs predicting BDD-only differing significantly from those predicting SSD-only, although stomach or intestinal ulcer was more strongly associated with BDD-only than SSD-only (OR: 4.0 vs. 1.5), whereas the reverse was true for asthma (OR: 0.6 vs. 2.0) (see Table 4).

Impairment

The mean DOR was significantly higher among respondents with BDD or SSD than those without in the total sample, both before weighting (5.5±12.4 vs. 1.4±5.0, X^2=15,550.1, p<0.001) and after stable balancing weights were used to adjust for differences in covariance distributions (3.9±13.0 vs. 1.5±4.0, X^2=3,229.6, p<0.001) (see Table 5).

The same general pattern held in the subsample of respondents who had none of the comorbid mental disorders or chronic

Table 2 Pooled multivariable associations of socio-demographic variables with bodily distress disorder (BDD) and somatic symptom disorder (SSD)

	BDD		SSD		Either		Both		BDD-only		SSD-only	
	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)
Age												
18-24	0.4	(0.2-0.9)	0.4	(0.2-0.7)	0.5	(0.3-0.8)	0.2	(0.1-0.6)	1.4	(0.4-4.2)	0.6	(0.3-1.2)
25-39	0.5	(0.3-0.7)	0.6	(0.4-0.8)	0.6	(0.4-0.8)	0.4	(0.3-0.7)	0.8	(0.4-1.7)	0.7	(0.4-1.1)
40-54	0.8	(0.5-1.2)	0.8	(0.6-1.2)	0.9	(0.7-1.2)	0.6	(0.4-1.1)	1.4	(0.7-2.9)	1.0	(0.7-1.6)
55+ (ref.)	-	-	-	-	-	-	-	-	-	-	-	-
X^2		13.4*		15.2*		17.0*		14.0*		3.8		6.5
Sex												
Male (ref.)	-	-	-	-	-	-	-	-	-	-	-	-
Female	2.2	(1.6-3.0)	1.7	(1.3-2.3)	1.9	(1.4-2.4)	2.1	(1.4-3.1)	2.5	(1.4-4.7)	1.6	(1.1-2.2)
X^2		22.9*		15.7*		22.6*		13.1*		8.7*		7.1*
Education												
Low	1.3	(0.9-2.1)	1.1	(0.7-1.5)	1.1	(0.8-1.6)	1.2	(0.7-2.1)	1.5	(0.7-3.2)	0.9	(0.6-1.5)
Low-middle	0.9	(0.6-1.4)	0.8	(0.6-1.2)	0.9	(0.6-1.2)	0.9	(0.6-1.4)	0.9	(0.4-1.7)	0.8	(0.5-1.3)
High-middle	1.1	(0.6-2.0)	0.8	(0.5-1.2)	1.0	(0.7-1.4)	0.7	(0.3-1.4)	2.2	(1.0-4.6)	0.8	(0.5-1.5)
High (ref.)	-	-	-	-	-	-	-	-	-	-	-	-
X^2		3.1		3.8		3.3		4.3		7.4		1.4
Marital status												
Married/cohabitating (ref.)	-	-	-	-	-	-	-	-	-	-	-	-
Never married	1.0	(0.7-1.6)	0.9	(0.6-1.2)	0.9	(0.7-1.2)	1.0	(0.6-1.6)	1.1	(0.6-2.1)	0.8	(0.5-1.3)
Previously married	0.7	(0.4-1.2)	0.9	(0.6-1.2)	0.9	(0.6-1.2)	0.6	(0.3-1.1)	1.0	(0.4-2.7)	1.1	(0.7-1.7)
X^2	2.1			1.3		0.9		2.9		0.1		0.9
Family income												
Low	1.3	(0.9-1.9)	1.2	(0.9-1.7)	1.2	(0.9-1.6)	1.4	(0.9-2.1)	1.1	(0.5-2.4)	1.1	(0.8-1.6)
Low-middle	1.4	(0.9-2.1)	1.1	(0.8-1.6)	1.2	(0.9-1.6)	1.2	(0.7-2.2)	1.6	(0.8-3.4)	1.0	(0.7-1.6)
High-middle	1.0	(0.6-1.7)	1.0	(0.7-1.5)	1.1	(0.8-1.5)	0.8	(0.4-1.6)	1.5	(0.7-3.3)	1.2	(0.8-1.8)
High (ref.)	-	-	-	-	-	-	-	-	-	-	-	-
X^2		3.5		1.5		1.7		3.7		2.2		0.7
Employment												
Employed (ref.)	-	-	-	-	-	-	-	-	-	-	-	-
Self-employed	0.9	(0.5-1.4)	1.1	(0.8-1.5)	1.0	(0.8-1.5)	0.9	(0.5-1.6)	0.8	(0.4-1.9)	1.2	(0.7-1.9)
Retired	7.5	(2.3-24.7)	9.9	(4.0-24.7)	9.5	(4.0-22.4)	9.6	(2.4-39.1)	4.6	(1.0-20.7)	9.7	(3.2-29.0)
Disabled	1.2	(0.6-2.2)	1.5	(1.1-2.2)	1.5	(1.1-2.1)	1.2	(0.6-2.5)	1.0	(0.3-3.3)	1.8	(1.1-3.0)
Student	1.1	(0.4-3.2)	1.0	(0.3-2.9)	0.9	(0.4-2.1)	1.8	(0.4-8.1)	0.5	(0.1-1.6)	0.6	(0.1-2.9)
Homemaker	1.1	(0.8-1.7)	1.3	(0.9-1.8)	1.3	(0.9-1.7)	1.2	(0.7-2.0)	1.2	(0.5-2.7)	1.4	(0.9-2.1)
Other	1.9	(1.1-3.2)	2.1	(1.4-3.0)	2.0	(1.4-2.8)	2.2	(1.2-4.0)	1.4	(0.5-3.7)	2.0	(1.2-3.2)
X^2		22.3*		39.5*		40.6*		23.9*		9.7		35.7*

OR – odds ratio, *significant at the 0.05 level, two-sided design-based test

physical conditions considered here (1.9±5.0 vs. 1.0±3.8, X^2=857.3, p<0.001), as well as in the subsamples of those with both comorbid mental disorders and chronic physical conditions (11.7±16.2 vs. 4.0±7.7, X^2=4,374.5, p<0.001), only comorbid mental disorders (7.4±13.3 vs. 3.5±9.8, X^2=1,676.1, p<0.001), and only chronic physical conditions (4.7±9.7 vs. 1.7±5.7, X^2=1,584.1, p<0.001) (see Table 5).

Respondents in the BDD-only subsample had a significantly

Table 3 Pooled multivariable associations of bodily distress disorder (BDD) and somatic symptom disorder (SSD) with 12-month DSM-5 mental disorders and suicidal ideation/behavior

	BDD		SSD		Either		Both		BDD-only		SSD-only	
	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)
GAD	3.3*	(1.4-7.7)	1.7	(0.8-3.7)	2.3*	(1.3-4.4)	3.1*	(1.1-9.3)	7.0*	(3.0-16.3)	1.5	(0.6-3.6)
Panic disorder	3.0*	(1.6-5.8)	2.6*	(1.1-5.8)	2.8*	(1.4-5.8)	3.0*	(1.4-6.7)	2.4*	(1.2-4.8)	2.4	(0.6-8.7)
PTSD	4.1*	(2.6-6.5)	2.4*	(1.7-3.5)	2.8*	(2.0-3.9)	4.2*	(2.3-7.7)	4.5*	(2.1-9.5)	1.6	(1.0-2.5)
Major depressive disorder	7.8*	(4.1-14.7)	4.4*	(2.3-8.4)	5.5*	(3.1-9.8)	8.0*	(3.5-18.2)	9.2*	(3.7-22.9)	3.3*	(1.4-7.6)
Bipolar spectrum disorders	5.3*	(2.5-11.4)	3.0*	(1.4-6.4)	3.2*	(1.7-6.1)	7.0*	(2.6-18.6)	3.8*	(1.5-9.2)	1.5	(0.5-4.8)
Alcohol use disorder	2.5	(0.7-9.0)	2.1	(0.8-5.5)	2.0	(0.8-5.1)	2.9	(0.8-11.4)	0.3	(0.1-1.3)	1.6	(0.5-5.8)
Suicidal ideation	1.9*	(1.0-3.6)	2.9*	(1.9-4.3)	2.9*	(2.0-4.1)	2.4*	(1.1-5.3)	2.0	(0.7-5.8)	3.5*	(2.2-5.6)
Suicide attempt	1.2	(0.3-5.5)	0.8	(0.2-2.5)	0.9	(0.3-2.8)	0.9	(0.2-4.4)	0.5	(0.0-5.8)	0.8	(0.2-3.1)

OR – odds ratio, GAD – generalized anxiety disorder, PTSD – post-traumatic stress disorder, *significant at the 0.05 level, two-sided design-based test

higher mean DOR than those in the SSD-only subsample, both before weighting (6.1±9.5 vs. 3.5±8.0, X^2=1,926.3, p<0.001) and after stable balancing weights were used to adjust for differences in covariate distributions (4.7±9.2 vs. 3.1±7.5, X^2=661.5, p<0.001).

The same general pattern was found among respondents who had none of the comorbid mental disorders or chronic physical conditions considered here (4.1±7.2 vs. 1.6±5.0, X^2=81.0, p<0.001) and those who had both comorbid mental disorders and chronic physical conditions (11.5±15.6 vs. 6.3±9.6, X^2=188.2, p<0.001). The mean DOR was also significantly higher in the BDD-only than the SSD-only subsample among respondents with only comorbid mental disorders (7.0±7.4 vs. 4.9±9.0, X^2=729.4, p<0.001), but the opposite was true among respondents with only chronic physical conditions (3.6±4.6 vs. 4.2±8.4, X^2=143.0, p<0.001) (see Table 5).

DISCUSSION

This is the first report providing population prevalence estimates for BDD and SSD derived from interviews using standardized diagnostic assessments. A few general population studies provided prevalence estimates of proxy diagnoses of SSD operationalized by either a combination of self-report questionnaires or unstructured clinical assessment, but none of those studies used standardized diagnostic interviews or operationalized full criterion sets[13]. We are aware of no previous prevalence studies of BDD in any setting.

The point prevalence across the three settings was 2.0% for BDD, 3.5% for SSD, and 4.1% for either diagnosis. Although the two diagnoses were highly correlated, prevalence estimates for SSD were significantly higher than those for BDD in two of the settings. There are at least two possible reasons for this observation. First, a diagnosis of BDD requires the presence of distress *and* impairment, while a diagnosis of SSD requires distress *or* impairment. Second, a diagnosis of BDD requires that somatic symptoms persist despite reassurance by a health care provider, whereas a diagnosis of SSD does not have this requirement. It is plausible that the first difference had the effect of restricting the diagnosis of BDD relative to SSD across sites, while the second difference may have had differential impact reflecting health service practice as well as culturally influenced patterns of health seeking in specific settings[43].

The prevalence of BDD and especially of SSD was highest in Hong Kong. Cross-cultural comparisons of the population occurrence of disorders of somatic distress using similar tools and comparable methodologically rigorous design as described here are uncommon. However, there is evidence from previous communi-

Table 4 Pooled multivariable associations of bodily distress disorder (BDD) and somatic symptom disorder (SSD) with common chronic physical conditions

	BDD		SSD		Either		Both		BDD-only		SSD-only	
	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)
Arthritis	2.9*	(1.5-5.6)	3.1*	(1.9-5.1)	2.9*	(1.8-4.7)	3.9*	(1.8-8.3)	1.5	(0.3-8.9)	2.6*	(1.4-4.8)
Asthma	2.4*	(1.4-4.3)	2.6*	(1.7-4.1)	2.3*	(1.5-3.5)	3.7*	(2.0-6.9)	0.6	(0.2-1.8)	2.0*	(1.0-3.8)
Diabetes mellitus	1.4	(0.8-2.5)	1.9*	(1.3-2.8)	1.8*	(1.2-2.7)	1.6	(0.8-3.2)	1.4	(0.5-3.5)	2.1*	(1.3-3.3)
Hypertension	1.5	(0.9-2.6)	1.4*	(1.0-2.1)	1.4	(1.0-2.0)	1.6	(0.8-2.9)	1.2	(0.4-3.6)	1.3	(0.8-2.0)
Ulcer	1.4	(0.6-3.3)	1.2	(0.7-2.0)	1.5	(0.9-2.4)	0.8	(0.3-2.1)	4.0*	(1.1-13.9)	1.5	(0.9-2.4)
Any other	3.5*	(2.4-5.2)	3.3	(2.5-4.3)	3.5*	(2.8-4.5)	3.4*	(2.1-5.6)	4.4*	(2.5-7.8)	3.2*	(2.2-4.7)

OR – odds ratio, ulcer – stomach or intestinal ulcer, *significant at the 0.05 level, two-sided design-based test

Table 5 Pooled associations of bodily distress disorder (BDD) and somatic symptom disorder (SSD) with mean number of health-related days out of role (DOR) in the 30 days before survey, unadjusted or adjusted based on stable balancing weights

	BDD-only		SSD-only		Both		Either		Neither		Total	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Total sample												
Unadjusted	6.1	(9.5)	3.5	(8.0)	8.4	(16.5)	5.5	(12.4)	1.4	(5.0)	1.6	(5.9)
Adjusted	4.7	(9.2)	3.1	(7.5)	3.1	(6.3)	3.9	(13.0)	1.5	(4.0)	1.6	(4.7)
Comorbidity subsamples												
No comorbidity	4.1	(7.2)	1.6	(5.0)	1.4	(2.9)	1.9	(5.0)	1.0	(3.8)	1.0	(3.8)
Both mental and physical	11.5	(15.6)	6.3	(9.6)	14.8	(18.0)	11.7	(16.2)	4.0	(7.7)	5.4	(10.2)
Only mental	7.0	(7.4)	4.9	(9.0)	10.5	(18.2)	7.4	(13.3)	3.5	(9.8)	3.8	(10.6)
Only physical	3.6	(4.6)	4.2	(8.4)	6.2	(12.5)	4.7	(9.7)	1.7	(5.7)	1.0	(6.5)

All the differences between respondents with either diagnosis vs. neither, and between those with BDD-only vs. SSD-only, were significant ($p<0.001$, two-sided design-based test)

ty-based studies that the experience of burdensome somatic symptoms is common in China in general and Hong Kong in particular[44], which may be regarded as a culturally determined expression of distress. Moreover, social factors – such as Hong Kong's status as one of the most densely populated cities globally, with its intensely fast-paced lifestyle and long working hours – may contribute to the high rates of reported somatic symptoms[2]. The higher prevalence rate of SSD compared to BDD in Hong Kong may be due to the criterion specification of the latter requiring "repeated contacts with health care providers". As stated earlier, this specification may lower the prevalence of BDD relative to SSD in settings where help-seeking for physical symptoms that accompany mental disorders is low. There is evidence to suggest that this is the case among Chinese people[45].

Our SSD point prevalence estimates are considerably lower than the mean "frequency" of 12.9% in previous reports of population-based studies[14]. However, the diagnosis of SSD in all those earlier studies was based on cut-off scores of screening questionnaires rather than direct interviews using standardized tools and diagnostic criteria. It has been repeatedly pointed out that screening questionnaires should not be used for the assessment of prevalence rates of mental health conditions[e.g., 46]. Indeed, the scoping review reporting the above "frequency" data repeatedly acknowledged that they were overestimates[14].

A major reason for discarding previous constructs of disorders of somatic distress as described in the DSM-IV was the implausible rarity of somatization disorder at one extreme of the spectrum and the extremely high rates of undifferentiated somatoform disorder at the other extreme. For example, two German population studies[47,48], with a total sample of 7,096 respondents, identified only one participant who met criteria for somatization disorder, while the lifetime prevalence of undifferentiated somatoform disorder was in the range between 9.1% and 19.7%. The point prevalence estimates of BDD and SSD in the current surveys are more plausible, and lend no support to the concern expressed by some critics that the new disorders are likely to be overinclusive and result in overdiagnosis[49,50]. Rather, as argued by others[51], it appears that the elimination of the criterion requiring that symptoms be medically unexplained and the inclusion of specific psychological symptoms in the criterion specifications have produced improvement over earlier conceptualizations of the conditions.

Striking patterns in our data are the higher prevalence of BDD and SSD among females, older individuals and retired people. The association with female gender is similar in magnitude to that observed in previous epidemiological studies on disorders of somatic distress[52]. However, unlike what has been reported in a few studies of DSM-IV defined somatic disorders[1], we did not observe an association of either BDD or SDD with low socio-economic status.

The associations of BDD and SSD with anxiety and mood disorders are consistent with what is commonly reported for somatoform disorders[31], reflecting the common observation of symptom overlap between these conditions, especially in primary care settings[53,54]. A large proportion of persons with chronic pain, a common symptom presentation of somatic distress, will meet the diagnostic criteria for anxiety or mood disorder[55]. The associations of BDD and SSD with suicidal ideation are also consistent with previous studies of disorders of somatic distress[56]. In general, there is similarity in the pattern of psychiatric comorbidity of BDD-only and SSD-only groups.

The pattern of comorbidity with common chronic physical conditions suggests that there is a meaningful but relatively modest increase in the prevalence of disorders of somatic distress, irrespective of whether defined as BDD or SSD, among persons with these physical conditions. This pattern suggests that the decision to eliminate the previous criterion requiring that symptoms are not medically explained has not led to implausibly high levels of comorbidity with these conditions.

There is a trend for people with SSD-only to have a significant association with more physical conditions than those with BDD-only. This may reflect the difference in the criterion specification. For BDD, there is a requirement for symptoms to persist "despite appropriate clinical examination and investigations or appropri-

ate reassurance by health care providers". SSD does not have such requirement. In the presence of physical health conditions, the requirement may limit the number of persons who receive the diagnosis of BDD. However, the difference is actually very modest. Moreover, SSD-only was more strongly associated with asthma than BDD-only, while BDD-only was more strongly associated with stomach and intestinal ulcer than SSD-only. So, our findings do not support the argument that the definition of SSD leads to mislabelling of persons with chronic physical conditions as having a mental disorder[50,57].

The presence of somatic distress (defined as either BDD or SSD) is associated with a significant decrement in role functioning even among persons with no co-occurring physical or mental disorder. The pattern of associations with role impairment would suggest that BDD-only is more impairing than SSD-only, although the picture is ambiguous across subgroups defined by the presence or absence of comorbid mental disorders and physical conditions. Nevertheless, it is clear that either disorder involves some role functioning difficulties for those experiencing it, and the common comorbidity with either physical or mental disorders increases the level of role impairment.

Some limitations of this study need to be noted. First, the weighted average response rate across the three surveys (51.6%) was relatively low, and it is possible that persons with very burdensome somatic symptoms were more likely to decline interviews. Second, surveys in Qatar were conducted by telephone, due to the COVID-19 pandemic, unlike the other two settings where survey interviews were conducted face-to-face. Third, in a cross-sectional design, no causality can be inferred regarding direction of association of the correlates. Finally, the measures of the correlates capture different time frames: 30 days for DOR, 12 months for mental disorders, and lifetime for chronic physical conditions, with the potential for lack of precision in the temporal associations.

In conclusion, this study provided a unique opportunity for the exploration of the profile and correlates of disorders of somatic distress in the community, using direct interviews of respondents conducted with a standardized questionnaire. In addition, we are able to present information relating to the differences and similarities between the relatively new diagnostic constructs of these disorders. Our findings suggest that disorders of somatic distress, defined as either SSD or BDD, occur in a considerable proportion of the population, and are associated with significant role impairment also after adjusting for comorbid mental disorders and chronic physical conditions. These findings suggest the need for focused public health attention to these distressing and burdensome conditions.

ACKNOWLEDGEMENTS

The WMH Survey Initiative is supported by the US National Institute of Mental Health (grant no. R01 MH070884), the MacArthur Foundation, the Pfizer Foundation, the US Public Health Service (grants nos. R13-MH066849, R01-MH069864 and R01 DA016558), the Fogarty International Center (grant no. R03-TW006481), the Pan American Health Organization, Eli Lilly & Co., Ortho-McNeil Pharmaceutical Inc., GlaxoSmithKline, and Bristol-Myers Squibb. The authors are grateful to the staff of the WMH Data Collection and Data Analysis Coordination Centres for their assistance. The World Mental Health Hong Kong Study was funded by the WYNG Foundation, the State Key Laboratory of Brain and Cognitive Sciences, and the University of Hong Kong. The establishment of the original FAMILY Cohort was funded by the Hong Kong Jockey Club Charities Trust. The Philippine National Survey for Mental Health and Well-being was funded by the Department of Health under the Advancing Health through Evidence-Assisted Decisions with Health Policy and Systems Research Program of the Philippine Council for Health Research and Development, with support from the University of the Philippines. The World Mental Health Qatar Study was conducted by the Social and Economic Survey Research Institute at Qatar University. This study was funded by Hamad Medical Corporation through Cambridgeshire and Peterborough National Health Service Foundation Trust, UK. None of the funders had any role in the design, analysis, interpretation of results, or preparation of this paper. Supplementary information on this study is available at https://osf.io/qve7h.

REFERENCES

1. Creed FH, Davies I, Jackson J et al. The epidemiology of multiple somatic symptoms. J Psychosom Res 2012;72:311-7.
2. Lowe B, Toussaint A, Rosmalen JGM et al. Persistent physical symptoms: definition, genesis, and management. Lancet 2024;403:2649-62.
3. Creed F, Gureje O. Emerging themes in the revision of the classification of somatoform disorders. Int Rev Psychiatry 2012;24:556-67.
4. Gureje O. Classification of somatic syndromes in ICD-11. Curr Opin Psychiatry 2015;28:345-9.
5. Voigt K, Nagel A, Meyer B et al. Towards positive diagnostic criteria: a systematic review of somatoform disorder diagnoses and suggestions for future classification. J Psychosom Res 2010;68:403-14.
6. Mayou R, Kirmayer LJ, Simon G et al. Somatoform disorders: time for a new approach in DSM-V. Am J Psychiatry 2005;162:847-55.
7. Creed F. Can DSM-V facilitate productive research into the somatoform disorders? J Psychosom Res 2006;60:331-4.
8. Gureje O, Simon GE, Ustun TB et al. Somatization in cross-cultural perspective: a World Health Organization study in primary care. Am J Psychiatry 1997;154:989-95.
9. Dimsdale JE, Creed F, Escobar J et al. Somatic symptom disorder: an important change in DSM. J Psychosom Res 2013;75:223-8.
10. Gureje O, Reed GM. Bodily distress disorder in ICD-11: problems and prospects. World Psychiatry 2016;15:291-2.
11. World Health Organization. Clinical descriptions and diagnostic requirements for ICD-11 mental, behavioural and neurodevelopmental disorders. Geneva: World Health Organization, 2024.
12. First MB, Gaebel W, Maj M et al. An organization- and category-level comparison of diagnostic requirements for mental disorders in ICD-11 and DSM-5. World Psychiatry 2021;20:34-51.
13. Lowe B, Mundt C, Herzog W et al. Validity of current somatoform disorder diagnoses: perspectives for classification in DSM-V and ICD-11. Psychopathology 2008;41:4-9.
14. Lowe B, Levenson J, Depping M et al. Somatic symptom disorder: a scoping review on the empirical evidence of a new diagnosis. Psychol Med 2022;52:632-48.
15. Keeley J, Reed GM, Rebello T et al. Case-controlled field study of the ICD-11 clinical descriptions and diagnostic requirements for bodily distress disorders. J Affect Disord 2023;333:271-7.
16. Creed F. Progress in understanding functional somatic symptoms and syndromes in light of the ICD-11 and DSM-5. World Psychiatry 2023;22:474-5.
17. Kessler RC, Ustun TB. The WHO World Mental Health Surveys: global perspectives on the epidemiology of mental disorders. New York: Cambridge University Press, 2008.
18. McGrath JJ, Al-Hamzawi A, Alonso J et al. Age of onset and cumulative risk of mental disorders: a cross-national analysis of population surveys from 29 countries. Lancet Psychiatry 2023;10:668-81.
19. Scott KM, de Jonge P, Stein DJ et al. Mental disorders around the world: facts and figures from the WHO World Mental Health surveys. Cambridge: Cambridge University Press, 2018.
20. Kessler RC, Heeringa SG, Pennell B-E et al. Methods of the World Mental Health Surveys. In: Bromet EJ, Karam EG, Koenen KC (eds). Trauma and posttraumatic stress disorder: global perspectives from the WHO World Mental Health Surveys. Cambridge: Cambridge University Press, 2018:13-42.
21. LKS Faculty of Medicine School of Public Health. The FAMILY Cohort. www.familycohort.sph.hku.hk.
22. Leung GM, Ni MY, Wong PT et al. Cohort profile: FAMILY Cohort. Int J Epidemiol 2017;46:e1.

23. De Guzman MLR, Barrios E, Crisostomo ME et al. Prevalence of mental, neurological and substance use (MNS) disorders in the Philippines. Submitted for publication.
24. Khaled SM, Amro I, Bader L et al. Qatar's National Mental Health Survey – The World Mental Health Qatar: sampling design, instrument adaptation, quality control, and fielding procedures. Int J Methods Psychiatr Res 2024;33:e2010.
25. American Association for Public Opinion Research. Standard definitions. www.aapor.org.
26. Kessler RC, Ustün TB. The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). Int J Methods Psychiatr Res 2004;13:93-121.
27. Harkness J, Pennell B, Villar A et al. Translation procedures and translation assessment in the World Mental Health Survey Initiative. In: Kessler RC, Üstun TB (eds). The WHO World Mental Health Surveys: global perspectives on the epidemiology of mental disorders. New York: Cambridge University Press, 2008:91-113.
28. Pennell BE, Mneimneh ZN, Bowers A et al. Implementation of the World Mental Health Surveys. In: Kessler R, Üstun T (eds). The WHO World Mental Health Surveys: global perspectives on the epidemiology of mental disorders. New York: Cambridge University Press, 2008:33-57.
29. Evans-Lacko S, Aguilar-Gaxiola S, Al-Hamzawi A et al. Socio-economic variations in the mental health treatment gap for people with anxiety, mood, and substance use disorders: results from the WHO World Mental Health (WMH) surveys. Psychol Med 2018;48:1560-71.
30. Scott KM, Al-Hamzawi AO, Andrade LH et al. Associations between subjective social status and DSM-IV mental disorders: results from the World Mental Health surveys. JAMA Psychiatry 2014;71:1400-8.
31. Lieb R, Meinlschmidt G, Araya R. Epidemiology of the association between somatoform disorders and anxiety and depressive disorders: an update. Psychosom Med 2007;69:860-3.
32. Schoenborn CA, Adams PF, Schiller JS. Summary health statistics for the U.S. population: National Health Interview Survey, 2000. Vital Health Stat 2003 (Series 10);214:1-83.
33. Centers for Disease Control and Prevention. Health, United States, 2004. Washington: US Department of Health and Human Services, 2004.
34. Knight M, Stewart-Brown S, Fletcher L. Estimating health needs: the impact of a checklist of conditions and quality of life measurement on health information derived from community surveys. J Public Health Med 2001;23:179-86.
35. Baker M, Stabile M, Deri Armstrong C. What do self-reported, objective, measures of health measure? Cambridge: US National Bureau of Economic Research, Inc, 2001.
36. Von Korff M, Crane PK, Alonso J et al. Modified WHODAS-II provides valid measure of global disability but filter items increased skewness. J Clin Epidemiol 2008;61:1132-43.
37. Kessler RC, Ormel J, Demler O et al. Comorbid mental disorders account for the role impairment of commonly occurring chronic physical disorders: results from the National Comorbidity Survey. J Occup Environ Med 2003;45:1257-66.
38. Revicki DA, Irwin D, Reblando J et al. The accuracy of self-reported disability days. Med Care 1994;32:401-4.
39. Kessler RC, Ames M, Hymel PA et al. Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. J Occup Environ Med 2004;46:S23-37.
40. SAS Institute. Statistical Analysis Software (SAS) User's Guide Version 9.4. Cary: SAS Institute, Inc, 2016.
41. Zubizarreta JR. Stable weights that balance covariates for estimation with incomplete outcome data. J Am Stat Assoc 2015;110:910-22.
42. Zubizarreta J, Li Y, Kim K et al. Stable balancing weights for causal inference and missing data. 1.1.9 version. https://cran.r-project.org.
43. Gureje O, Lewis-Fernandez R, Hall BJ et al. Systematic inclusion of culture-related information in ICD-11. World Psychiatry 2019;18:357-8.
44. Lee S, Leung CMC, Kwok KP et al. A community-based study of the relationship between somatic and psychological distress in Hong Kong. Transcult Psychiatry 2015;52:594-615.
45. Lee S, Guo WJ, Tsang A et al. Perceived barriers to mental health treatment in metropolitan China. Psychiatr Serv 2010;61:1260-2.
46. Zimmerman M. The value and limitations of self-administered questionnaires in clinical practice and epidemiological studies. World Psychiatry 2024;23:210-2.
47. Grabe HJ, Meyer C, Hapke U et al. Specific somatoform disorder in the general population. Psychosomatics 2003;44:304-11.
48. Lieb R, Pfister H, Mastaler M et al. Somatoform syndromes and disorders in a representative population sample of adolescents and young adults: prevalence, comorbidity and impairments. Acta Psychiatr Scand 2000;101:194-208.
49. Frances A. The new somatic symptom disorder in DSM-5 risks mislabeling many people as mentally ill. BMJ 2013;346:f1580.
50. Frances AJ, Nardo JM. ICD-11 should not repeat the mistakes made by DSM-5. Br J Psychiatry 2013;203:1-2.
51. Voigt K, Wollburg E, Weinmann N et al. Predictive validity and clinical utility of DSM-5 somatic symptom disorder: prospective 1-year follow-up study. J Psychosom Res 2013;75:358-61.
52. Creed F, Barsky A. A systematic review of the epidemiology of somatisation disorder and hypochondriasis. J Psychosom Res 2004;56:391-408.
53. Ziebold C, Goldberg DP, Reed GM et al. Dimensional analysis of depressive, anxious and somatic symptoms presented by primary care patients and their relationship with ICD-11 PHC proposed diagnoses. Psychol Med 2019;49:764-71.
54. Goldberg DP, Reed GM, Robles R et al. Multiple somatic symptoms in primary care: a field study for ICD-11 PHC, WHO's revised classification of mental disorders in primary care settings. J Psychosom Res 2016;91:48-54.
55. Aaron RV, Ravyts SG, Carnahan ND et al. Prevalence of depression and anxiety among adults with chronic pain: a systematic review and meta-analysis. JAMA Netw Open 2025;8:e250268.
56. Torres ME, Lowe B, Schmitz S et al. Suicide and suicidality in somatic symptom and related disorders: a systematic review. J Psychosom Res 2021;140:110290.
57. Frances A, Chapman S. DSM-5 somatic symptom disorder mislabels medical illness as mental disorder. Aust N Z J Psychiatry 2013;47:483-4.

DOI:10.1002/wps.21353

All-cause and cause-specific mortality in people with depression: a large-scale systematic review and meta-analysis of relative risk and aggravating or attenuating factors, including antidepressant treatment

Joe Kwun Nam Chan[1], Marco Solmi[2-5], Heidi Ka Ying Lo[1], Michelle Wai Yan Chan[1], Leo Long Tin Choo[1], Eric Tsz Him Lai[6], Corine Sau Man Wong[7], Christoph U. Correll[5,8-10], Wing Chung Chang[1,11]

[1]Department of Psychiatry, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong SAR, China; [2]SCIENCES lab, Department of Psychiatry, University of Ottawa, Ottawa, ON, Canada; [3]Regional Centre for the Treatment of Eating Disorders and On Track: Champlain First Episode Psychosis Program, Department of Mental Health, Ottawa Hospital, Ottawa, ON, Canada; [4]Clinical Epidemiology Program, Ottawa Hospital Research Institute, University of Ottawa, Ottawa, ON, Canada; [5]Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin, Berlin, Germany; [6]Department of Psychiatry, Kwai Chung Hospital, Hong Kong, Hong Kong SAR, China; [7]School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong SAR, China; [8]Department of Psychiatry, Zucker Hillside Hospital, Northwell Health, Glen Oaks, NY, USA; [9]Department of Psychiatry and Molecular Medicine, Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, USA; [10]German Center for Mental Health (DZPG), partner site Berlin, Berlin, Germany; [11]State Key Laboratory of Brain and Cognitive Sciences, University of Hong Kong, Hong Kong SAR, China

Depression has been reported to be associated with premature mortality. However, no meta-analysis has comprehensively examined all-cause and cause-specific mortality risk in people with this condition, focusing also on possible aggravating and attenuating factors, including antidepressant treatment. We conducted a systematic review and meta-analysis of cohort studies to synthesize mortality risk estimates associated with depression (major depressive disorder and dysthymia) due to any and specific causes, and when depression is accompanied by comorbid conditions. Effects of antidepressant medication and electroconvulsive therapy (ECT), and other potential moderators of mortality risk, were evaluated. We searched EMBASE, Medline and PsycINFO databases up to January 26, 2025, pooling mortality estimates using random-effect models. Publication bias, subgroup and meta-regression analyses, and quality assessment (Newcastle-Ottawa Scale) were performed. Across 268 studies, 10,842,094 individuals with depression and 2,837,933,536 control subjects were included. All-cause mortality was doubled in people with depression versus no depression/general population controls (relative risk, RR=2.10, 95% CI: 1.87-2.35, I^2=99.9%), being especially high for suicide (RR=9.89, 95% CI: 7.59-12.88, I^2=99.6%), but also elevated for natural causes (RR=1.63, 95% CI: 1.51-1.75, I^2=99.6%). Among individuals with versus without depression matched for comorbid conditions, the depression-associated mortality risk was also significantly elevated (RR=1.29, 95% CI: 1.21-1.37, I^2=99.9%). Depression with versus without psychotic symptoms (RR=1.61, 95% CI: 1.45-1.78, I^2=6.3%), and treatment-resistant versus non-treatment-resistant depression (RR=1.27, 95% CI: 1.16-1.39, I^2=85.3%), conferred an incremental mortality risk. Antidepressant use (versus no antidepressant use) was associated with significantly lower all-cause mortality in people with depression (RR=0.79, 95% CI: 0.68-0.93, I^2=99.2%). ECT use (versus no ECT use) was associated with reduced all-cause mortality (RR=0.73, 95% CI: 0.66-0.82, I^2=0%), natural-cause mortality (RR=0.76, 95% CI: 0.59-0.97, I^2=12.0%), and suicide (RR=0.67, 95% CI: 0.53-0.85, I^2=32.3%). Our results affirm heightened mortality risk in depression, identify clinically relevant patient subgroups with increased mortality risk, and highlight mortality-reducing effects of antidepressant treatment and ECT. Multipronged intervention approaches targeting physical health improvement and suicide risk alleviation, optimizing antidepressant treatment, and pursuing early identification and effective interventions for psychotic and treatment-resistant depression, could help reduce this mortality gap, which is still growing.

Key words: Depression, mortality, suicide, major depressive disorder, dysthymia, psychotic depression, treatment-resistant depression, antidepressant treatment, electroconvulsive therapy

(World Psychiatry 2025;24:404–421)

Depression is a potentially chronic[1,2] and treatable[3,4] mental disorder, with a lifetime prevalence of 15-18%[1,2], which represents one of the leading causes of global disease burden[1,5], involving substantial health care and societal costs. The disorder is also highly prevalent in people with a wide range of chronic physical diseases, with an average point prevalence of 25%[6,7]. Critically, accumulating data have shown that people with depression have an increased risk of premature mortality relative to the general population[8], with a reduced life expectancy of 13 years[9]. Despite markedly elevated risk of suicide, the excess death in individuals with depression is mainly attributable to natural causes[10,11]. Considering the persistent mortality gap associated with depression in recent decades[10,12], the health inequalities experienced by people with this condition represent a serious public health concern.

Several studies have investigated premature mortality patterns associated with depression, aiming to enhance the understanding of mechanisms underlying this excess mortality, as well as to identify modifiable factors that can inform policy formulation, resource allocation and health care enhancement. Some meta-analyses have been conducted in this respect[8], but they are hampered by significant methodological limitations.

First, a majority of the studies included in previous meta-analyses ascertained depression by self-report tools, which are actually intended to be used as a screening instrument for probable depression[12-15]. This procedure may increase the likelihood of misclassifying individuals with subthreshold depressive symptoms as having a psychiatric diagnosis of depression, potentially underestimating the excess mortality risk associated with depression. Misclassification bias may be more pronounced when self-rating instruments are used to ascertain comorbid depression among individuals with severe physical diseases, in whom physical symptoms can overlap with or mimic depressive symptoms[16]. Cross-study variations in cut-off scores used with the same tool introduce even greater heterogeneity, and further compromise accuracy in depression case ascertainment[17]. Thus far, there has been no meta-analysis only including studies which defined depression according to ICD or

DSM, based on diagnostic interviews or clinician-assigned diagnosis ascertained from health-record databases.

Second, evaluation of mortality risk was often restricted to a subgroup of patients with a specific physical morbidity, such as cardiovascular disease[18-20], cancer[21,22] or diabetes mellitus[23,24], comparing people with depression to those with the same physical morbidity but without depression. Third, prior analyses did not take into consideration the incident and prevalent depression status, precluding the investigation of the association between mortality risk and duration of depression[8,20,21,22]. Fourth, most prior meta-analyses focused on all-cause mortality risk, without a comprehensive evaluation of risk for cause-specific deaths in people with depression[12,15,18-20,23,25]. Fifth, evaluation of the relationships of mortality risk with subtypes of the condition, such as psychotic and treatment-resistant depression, is limited.

Notably, despite the mixed findings reported in the literature concerning the association of antidepressant or electroconvulsive therapy (ECT) use with excess mortality in people with depression [26-45], there has been no meta-analysis including the evaluation of the impact of these treatments on mortality risk.

To fill this research gap, we conducted the most comprehensive systematic review and meta-analysis to date examining the risk of all-cause and cause-specific mortality in people with depression versus those with no depression or the general population. We also evaluated mortality risk associated with depression in people with any or specific comorbid conditions. We only included studies that ascertained depression according to ICD or DSM, using diagnostic interviews or health-record database-derived clinician-assigned diagnosis. In addition, associations of antidepressant treatment (any antidepressant, drug classes, and individual agents) and ECT with mortality risk were assessed. To explore potential sources of heterogeneity and factors that may aggravate or attenuate mortality risk associated with depression, we performed subgroup and meta-regression analyses stratified by a range of study characteristics and depression-related factors, such as incident/prevalent sample, time intervals of observation after depression diagnosis, presence of psychotic symptoms, and treatment-resistant status.

METHODS

This study was conducted in accordance with the Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines and the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA 2020)[46]. The study protocol was registered with PROSPERO (CRD42023451258).

Search strategy and selection criteria

We searched EMBASE, Medline and PsycINFO for articles published from inception to January 26, 2025, without language restrictions. The search key words included terms related to depression, antidepressant treatment and mortality (see supplementary information for details). We also hand-searched references of all selected papers and relevant reviews to identify additional eligible studies. Two reviewers performed the search independently and compared the results. Disagreement was resolved by consensus.

Studies were selected if they: a) included patients of any age with depression (i.e., major depressive disorder or dysthymia) defined according to any version of ICD or DSM, based on a diagnostic interview or a clinician-assigned coded diagnosis derived from health-record databases; b) reported data on all-cause and cause-specific mortality; and c) were cohort studies. Publications that adopted non-cohort designs, such as case-control studies; reviews and meta-analyses; studies containing qualitative or non-meta-analyzable data, or restricted to population subgroups (e.g., homeless or incarcerated people), or with sample sizes <100 were excluded. Two authors independently screened titles and abstracts of relevant papers for inclusion, and disagreements were resolved through discussion with two other authors.

Outcomes, data extraction and assessment of study quality

The primary outcome was risk of all-cause mortality in individuals with depression. The secondary outcomes included mortality due to natural, unnatural and more specific causes. Analyses were performed in prevalent plus incident cohorts, where prevalent cases were individuals living with depression, regardless of diagnosis date, while incident cases included individuals with newly-diagnosed depression within the period of observation. Comparison groups included the general population, people without depression, and psychiatric controls. Individuals with any/specific comorbid conditions with and without depression were also compared.

To investigate mortality risk associated with antidepressant use and ECT, people with depression treated with any/specific antidepressants (drug classes or individual agents) or ECT were compared to those with depression not receiving treatment with antidepressants or ECT, respectively. Additional comparisons in relation to other depression-related characteristics, including dysthymia/no depression, various time intervals of observation after depression diagnosis, late-life depression/no depression, early-life depression/no depression, depression with/without psychotic symptoms, and treatment-resistant versus non-treatment-resistant depression, were performed.

Data were extracted independently by two authors using a predefined form, with discrepancies resolved by consensus. Since the current study focused on depression, studies pooling data of people with other psychiatric diagnoses (e.g., combining patients with depression and schizophrenia) were excluded, unless the study provided stratified analyses only for people with depression. If several adjusted risk estimates were reported, the one controlling for the most comprehensive set of covariates was chosen.

When studies presented findings graphically, we extracted the data from figures using WebPlotDigitizer, a web-based tool for numerical data extraction from plots and graph images. For studies that only reported data on point estimates without standard errors (SE) or 95% confidence intervals (CIs), we extrapolated the SE as the

mean from studies that reported SE. Following previous research[9,47], for studies using the general population as the reference with standardized mortality ratios (SMRs) for mortality risk, we estimated the sample size of the control group as the size of the general population in that country or region and in the age range of the depression group, based on census-based data for the median year of the study period.

Risk of bias was assessed independently by two reviewers using the Newcastle-Ottawa Scale[48], which covers the following three domains: a) selection (representativeness, selection of non-exposed cohort, ascertainment of exposure, outcome of interest not present at baseline); b) comparability (control for covariates); and c) outcome (assessment of mortality; follow-up duration ≥3 years, unless predefined time frame for investigation). Disagreements were resolved through consultation with other members of the research team.

Data analysis

Given the generally rare cumulative incidence of mortality in included studies (i.e., <10%)[49], SMRs, hazard ratios, odds ratios, risk ratios, and incidence rate ratios were treated as equivalent measures of risk, with an aim to give an overview of relative associations[47,50,51], and the term relative risk (RR) is then used thereafter. Random-effects meta-analytic models were applied to generate pooled estimates of RR for depression versus no depression/general population, depression versus no depression matched for any comorbid conditions, and major depressive disorder versus no depression/general population.

I^2 statistic was used to measure the total variation due to heterogeneity[52]. Additionally, Cochran's Q test was performed to assess the statistical significance of the heterogeneity across studies. Publication bias was assessed using Egger's test[53], with p values <0.1 considered significant. In case of publication bias, we also calculated the fail-safe number as the estimated number of studies needed to move the mortality risk from significant to non-significant, and performed the Duval and Tweedie's trim-and-fill procedure[54].

Aggravating or attenuating factors and sources of heterogeneity were explored with subgroup and meta-regression analyses. Subgroup analyses were stratified by: control group (general population and people with no depression); prevalent/incident depression sample; sex; age categories (<25 years, 25-60 years, and >60 years); diagnostic system (ICD and DSM); geographical location in terms of continents; source of study samples (health-system case registers, health-insurance databases, hospital/clinic samples or records, community surveys); population of people with depression (community, inpatient, outpatient, or inpatient and outpatient); other depression-related characteristics (dysthymia, time intervals of observation after depression diagnosis, late-life depression, early-life depression, depression with psychotic symptoms, and treatment-resistant depression); and use of antidepressants (any antidepressant, drug classes, individual agents) and ECT.

Random-effects meta-regression analyses were performed on potential moderators, including characteristics of the overall sample (median year of observation period, number of years in observation period, mean follow-up duration, number of adjusted covariates, human development index[55], socio-demographic index[56], and Newcastle-Ottawa Scale score); characteristics of the depression sample (sample size, and proportion of people with major depressive disorder, dysthymia and antidepressant treatment); and difference in characteristics between depression and non-depression samples (mean age, body mass index, proportion of people being female, White, current smoker, and married; and percentage of people with obesity, alcohol use disorders, substance use disorders, diabetes mellitus, cancers, and renal diseases).

Meta-analysis models were performed in R (version 4.1.2) with *metafor* package. For all analyses, except Egger's test, p values <0.05 were considered significant.

RESULTS

Search results

The PRISMA flow diagram describing the process of study identification and selection is shown in Figure 1. The literature search identified 18,056 papers (18,026 from database searching and 30 from manual search), of which 16,860 remained after removal of duplicates. Upon exclusion of irrelevant studies, we retrieved 653 full-text papers to be assessed for eligibility. Of these, 385 were excluded, mainly due to lack of relevant outcomes, no depression, or ascertainment of depression based on self-report measures or depressive symptoms (see supplementary information).

Altogether, 268 publications met inclusion criteria[10,30-36,43,45,57-314], comprising 10,842,094 individuals with depression and 2,837,933,536 control subjects. Comparisons included people with depression (N=1,900,317) versus the general population (N=2,650,612,526); individuals with depression (N=5,455,521) versus no depression (N=43,415,950); people with depression (N=5,881,116) versus no depression (N=40,284,386) matched for comorbid conditions; and individuals with depression (N=76,751) versus other mental disorders (N=37,421). Other mental disorders included schizophrenia (one study, N=861), bipolar disorder (three studies, N=5,192), adjustment disorder (one study, N=31), and alcohol use disorders (one study, N=31,337). Only data on depression versus bipolar disorder were sufficient for meta-analysis. The characteristics of individual studies are provided in the supplementary information.

Studies were conducted in the US (n=79), the UK (n=31), South Korea (n=24), Sweden (n=24), Taiwan (n=21), Denmark (n=17), Canada (n=12), The Netherlands (n=10), Finland (n=9), Germany (n=5), Spain (n=5), Australia (n=4), Hong Kong (n=4), Switzerland (n=4), Italy (n=3), France (n=2); Brazil, China, Ethiopia, Hungary, Israel, Japan, Lithuania, Norway, Portugal, Singapore and South Africa (n=1 each). Two studies were conducted using data from multiple countries worldwide, and one using data from multiple countries in Europe.

Data of study samples were mainly derived from health-system case registers (n=135). Other data sources included health-insurance databases (n=52), hospital/clinic samples or records (n=50),

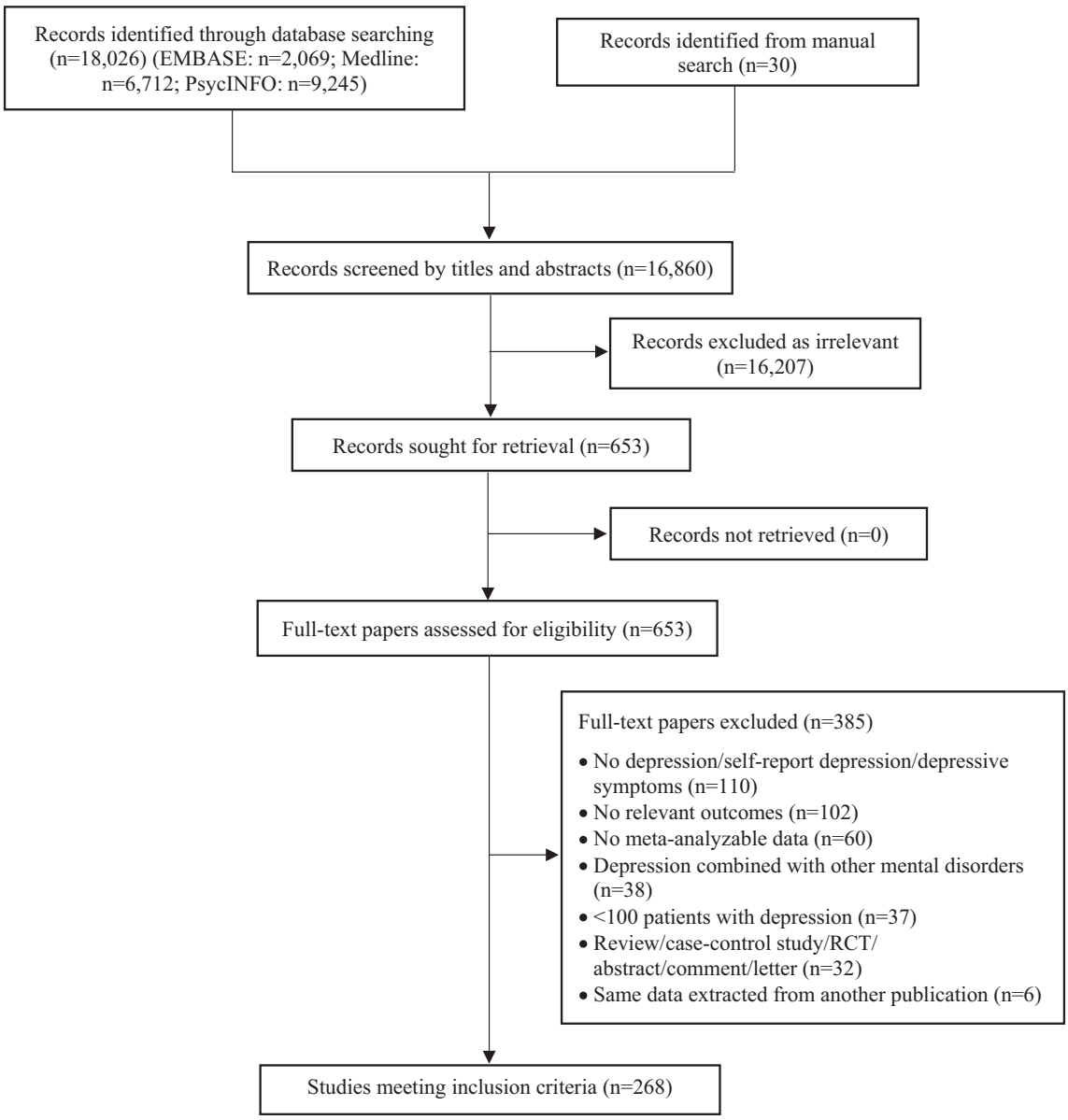

Figure 1 PRISMA 2020 flow chart. RCT – randomized controlled trial

and community surveys (n=31). The length of observation period ranged from 1 to 65 years. All studies examined cases following the ICD (n=200), DSM (n=62) or ICD/DSM (n=6). People with depression were identified in outpatient or inpatient/outpatient settings (n=187), inpatient settings (n=53), or the community (n=28).

Primary outcome: all-cause mortality

The pooled RR for all-cause mortality of individuals with depression versus no depression/general population was 2.10 (95% CI: 1.87-2.35; I^2=99.9%, n=128) (see Figure 2). The mortality risk was not significantly different in patients with incident (RR=2.04, 95% CI: 1.60-2.60, I^2=99.9%, n=20) versus prevalent (RR=2.05, 95% CI: 1.81-2.33, I^2=99.9%, n=110) depression (between-group p=0.974) (see Table 1). The pooled RRs for all-cause mortality of people with depression versus the general population (RR=2.38, 95% CI: 1.74-3.25, I^2=100.0%, n=37), and individuals with depression versus no depression (RR=2.01, 95% CI: 1.80-2.24, I^2=99.7%, n=92) were similar in magnitude (between-group p=0.782) (see supplementary information).

Among individuals with depression versus no depression matched for comorbid conditions, the depression-mortality association was significant (RR=1.29, 95% CI: 1.21-1.37, I^2=99.9%, n=98) (see Figure 3). The all-cause mortality risk was increased in people with depression versus no depression matched for comorbid alcohol/substance use disorders (RR=2.59, 95% CI: 1.71-3.93, I^2=99.8%, n=5); colorectal cancer (RR=1.80, 95% CI: 1.28-2.55, I^2=82.5%, n=2);

Death causes	n	Sample size for depression group/comparison group	RR (95% CI)	I^2 (%)
All causes	128	7,410,593/2,797,649,150	**2.10 (1.87-2.35)**	99.87
Natural causes	58	15,706,165/2,115,890,845	**1.63 (1.51-1.75)**	99.59
Unnatural causes	49	5,159,404/1,677,640,085	**5.81 (4.55-7.43)**	99.77
Suicide	44	2,687,171/1,487,047,309	**9.89 (7.59-12.88)**	99.59
Non-suicide unnatural causes	15	2,712,694/569,661,047	**2.06 (1.70-2.50)**	96.87
Accidents	8	994,802/462,323,299	**2.19 (1.38-3.47)**	98.81
Neurological diseases	6	2,104,035/16,994,013	**2.28 (1.81-2.88)**	96.20
Respiratory diseases	14	1,679,058/442,741,489	**2.34 (1.94-2.83)**	99.02
Endocrine diseases	5	1,764,197/121,544,138	**2.01 (1.34-3.01)**	99.36
Diabetes mellitus	3	1,700,885/113,086,772	**1.82 (1.12-2.96)**	99.58
Infectious diseases	7	1,210,291/19,795,495	**1.65 (1.21-2.24)**	96.41
Gastrointestinal diseases	9	2,928,690/1,400,830,974	**1.64 (1.25-2.15)**	98.42
Liver diseases	4	1,145,866/19,621,608	1.49 (0.98-2.27)	98.62
Cardiovascular diseases	36	3,849,896/844,677,849	**1.47 (1.35-1.60)**	98.89
Genitourinary diseases	6	1,866,232/565,557,974	**1.45 (1.19-1.76)**	93.95
Cancers	25	2,518,185/567,911,595	**1.35 (1.20-1.52)**	98.84
Cerebrovascular diseases	10	1,515,008/362,275,370	**1.27 (1.10-1.47)**	95.34
COVID-19	6	1981,456/26,917,222	1.13 (0.94-1.35)	95.68

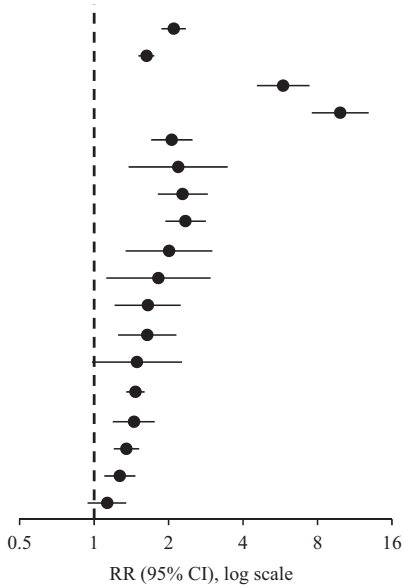

Figure 2 All-cause and cause-specific mortality risk in people with depression versus no depression/general population. RR – relative risk, COVID-19 – coronavirus 19 disease. Significant values are highlighted in bold prints.

peripheral vascular diseases (RR=1.42, 95% CI: 1.22-1.65, I^2=98.0%, n=3); myocardial infarction (RR=1.41, 95% CI: 1.16-1.71, I^2=97.3%, n=10); stroke (RR=1.40, 95% CI: 1.16-1.68, I^2=98.8%, n=7); prostate cancer (RR=1.38, 95% CI: 1.01-1.89, I^2=97.5%, n=3); diabetes mellitus (RR=1.33, 95% CI: 1.22-1.46, I^2=99.0%, n=11); any cardiovascular diseases (RR=1.32, 95% CI: 1.24-1.41, I^2=98.3%, n=35); chronic pulmonary diseases (RR=1.32, 95% CI: 1.09-1.61, I^2=98.7%, n=3); ischemic heart diseases (RR=1.29, 95% CI: 1.06-1.57, I^2=97.8%, n=13); non-ischemic cardiovascular diseases (RR=1.28; 95% CI: 1.18-1.39, I^2=95.9%, n=9); any cancers (RR=1.27, 95% CI: 1.16-1.39, I^2=98.0%, n=21); heart failure (RR=1.26, 95% CI: 1.16-1.37, I^2=95.9%, n=8); renal diseases (RR=1.20, 95% CI: 1.09-1.33, I^2=97.1%, n=5); respiratory diseases (RR=1.20, 95% CI: 1.06-1.36, I^2=97.8%, n=6); breast cancer (RR=1.20, 95% CI: 1.04-1.39, I^2=84.6%, n=8); and coronavirus disease 2019 (COVID-19) (RR=1.14, 95% CI: 1.02-1.28, I^2=99.8%, n=4) (see Figure 3).

In the analyses on people with major depressive disorder compared to no depression/general population, the pooled RR was 2.17 (95% CI: 1.69-2.79, I^2=99.7%, n=36), with evidence of publication bias (Egger's test p<0.001) (see Figure 4). The risk of all-cause mortality associated with major depressive disorder versus the general population (RR=2.41, 95% CI: 1.43-4.04, I^2=99.6%, n=8) was comparable to that versus no depression (RR=2.13, 95% CI: 1.61-2.83, I^2=99.7%, n=29) (between-group p=0.751). The magnitude of depression-associated all-cause mortality risk was significant in individuals matched for any comorbid conditions with major depressive disorder versus no depression (RR=1.33, 95% CI: 1.25-1.41, I^2=98.7%, n=25) (see supplementary information).

Individuals with dysthymia had an increased all-cause mortality risk compared to those with no depression (RR=1.40, 95% CI: 1.30-1.51, I^2=0%, n=3). All-cause mortality risk was markedly elevated in the 0-180 days after depression diagnosis (RR=10.80, 95% CI: 6.21-18.77, I^2=98.5, n=2); and lower but still significantly elevated in the observation periods of 180-365 days (RR=3.29, 95% CI: 1.51-7.17, I^2=98.0, Egger's test p<0.001, n=2), and 1-5 years (RR=4.23, 95% CI: 2.25-7.97, I^2=99.8, n=4) following depression diagnosis (see Figure 4).

Psychotic depression (versus non-psychotic depression: RR=1.61, 95% CI: 1.45-1.78, I^2=6.3%, n=2) and treatment-resistant depression (versus non-treatment-resistant depression: RR=1.27, 95% CI: 1.16-1.39, I^2=85.3%, Egger's test p<0.001, n=9) further increased depression-associated mortality risk. Both late-life depression (versus no depression: RR=2.11, 95% CI: 1.11-4.00, I^2=64.1%, n=3) and early-life depression (versus no depression: RR=1.73, 95% CI: 1.38-2.17, I^2=0.0%, n=2) were associated with increased mortality risk (see Figure 4).

Secondary outcomes: natural, unnatural and other cause-specific mortality

The RR for natural-cause mortality was 1.63 (95% CI: 1.51-1.75, I^2=99.6%, n=58) for depression relative to no depression/general population (see Figure 2). The depression-mortality association was consistent when compared to the general population (RR=1.74, 95% CI: 1.54-1.98, I^2=99.7%, n=19) and to individuals with no depression (RR=1.57, 95% CI: 1.43-1.73, I^2=99.4%, n=39) (between-group p=0.755). The association was also significant in individuals matched for any comorbid conditions (RR=1.20, 95% CI: 1.14-1.27, I^2=97.6%, n=22) (see supplementary information).

Table 1 Subgroup analyses on risk of all-cause mortality in patients with depression versus no depression/general population

Subgroup	n	Sample size for depression group/comparison group	RR (95% CI)	I^2 (%)	Between groups p	p for differences
Sex					0.615	
Male	65	4,990,780/1,946,058,373	2.37 (2.06-2.71)	99.9		Ref.
Female	58	5,267,202/2,162,929,399	2.27 (1.90-2.71)	99.9		0.615
Age					0.139	
<25 years	5	75,335/4,287,267	3.28 (1.79-6.02)	98.1		Ref.
25–60 years	8	2,697,902/61,509,449	3.54 (2.30-5.44)	99.9		0.942
>60 years	32	1,375,463/87,212,436	2.17 (1.67-2.83)	99.9		0.252
Depression sample nature					0.974	
Prevalent	110	4,964,453/2,457,348,522	2.05 (1.81-2.33)	99.9		Ref.
Incident	20	2,461,401/230,300,779	2.04 (1.60-2.60)	99.9		0.974
Diagnostic system					0.002	
ICD	75	7,268,131/995,528,374	2.41 (2.06-2.82)	99.9		Ref.
DSM	47	86,428/984,715,642	1.66 (1.47-1.87)	92.1		0.004
ICD/DSM	4	650/381,342,460	1.06 (0.69-1.64)	92.3		0.028
Continent					0.278	
Africa	2	143,614/838,526	2.00 (0.71-5.64)	92.0		Ref.
Asia	16	2,140,725/177,689,675	2.39 (1.53-3.74)	100.0		0.893
Australia	4	1,299/28,615,691	3.13 (0.91-10.84)	98.9		0.499
Europe	71	1,265,030/580,795,483	2.26 (1.97-2.59)	99.7		0.830
North America	30	1,586,594/1,573,503,906	1.51 (1.25-1.83)	99.8		0.703
South America	1	2,201,147/NA	2.35 (1.60-3.46)	99.6		0.784
Source of study samples					0.043	
Community surveys	30	61,195/302,131	1.58 (1.36-1.84)	90.3		Ref.
Health-system case registers	55	4,006,290/1,134,611,614	2.23 (1.94-2.56)	99.9		0.036
Health-insurance databases	15	2,218,351/188,888,363	2.80 (1.65-4.75)	100.0		0.041
Hospital/clinic samples or records	27	1,069,373/594,016,368	1.96 (1.49-2.59)	98.9		0.161
Population of depression sample					<0.001	
Community	28	60,917/578,747	1.57 (1.34-1.85)	91.1		Ref.
Outpatient or inpatient and outpatient	69	3,752,495/1,095,159,416	1.89 (1.65-2.16)	99.9		0.175
Inpatient	29	3,391,983/1,265,545,640	2.95 (2.31-3.76)	99.8		0.001

RR – relative risk, NA – not available

Depression was associated with increased unnatural-cause mortality risk relative to no depression/general population (RR=5.81, 95% CI: 4.55-7.43, I^2=99.8%, n=49) (see Figure 2). Depression-associated mortality risk estimates for unnatural causes were significantly higher when compared to the general population (RR=9.69, 95% CI: 6.02-15.59, I^2=99.9%, n=20) than compared to no depression (RR=4.36, 95% CI: 3.41-5.58, I^2=99.4%, n=28) (between-group p<0.001). The associations were also significant in individuals matched for any comorbid conditions (RR=2.57, 95% CI: 1.89-3.50, I^2=97.5%, n=9). Treatment-resistant status conferred an incremental effect on the depression-associated unnatural-cause mortality risk (RR=2.30, 95% CI: 1.68-3.14, I^2=92.6%, n=4), compared to non-treatment-resistant depression (see supplementary information).

Individuals with depression exhibited increased mortality risk compared to no depression/general population for suicide (RR=9.89, 95% CI: 7.59-12.88, I^2=99.6%, n=44); any non-suicide unnatural cause (RR=2.06, 95% CI: 1.70-2.50, I^2=96.9%, n=15); accidents (RR=2.19, 95% CI: 1.38-3.47, I^2=99.8%, n=8); neurological diseases (RR=2.28, 95% CI: 1.81-2.88, I^2=96.2%, n=6); respiratory diseases (RR=2.34, 95% CI: 1.94-2.83, I^2=99.0%, n=14); endocrine diseases (RR=2.01, 95% CI: 1.34-3.01, I^2=99.4%, n=5); diabetes mellitus (RR=1.82, 95% CI: 1.12-2.96, I^2=99.6%, n=3); infectious diseases (RR=1.65, 95% CI: 1.21-2.24, I^2=99.4%, n=7); gastrointestinal diseases (RR=1.64, 95% CI: 1.25-2.15, I^2=98.4%, n=9); cardiovascular

Comorbid conditions	n	Sample size for depression group/comparison group	RR (95% CI)	I^2 (%)
Any comorbid conditions	98	5,881,116/40,284,386	**1.29 (1.21-1.37)**	99.85
Alcohol/substance use disorders	5	67,540/538,232	**2.59 (1.71-3.93)**	99.84
Colorectal cancer	2	9,101/38,564	**1.80 (1.28-2.55)**	82.54
Peripheral vascular diseases	3	35,826/133,527	**1.42 (1.22-1.65)**	97.98
Myocardial infarction	10	18,952/171,040	**1.41 (1.16-1.71)**	97.27
Stroke	7	41,392/320,874	**1.40 (1.16-1.68)**	98.75
Prostate cancer	3	6,242/132,621	**1.38 (1.01-1.89)**	97.51
Angina	2	8,567/22,874	1.36 (0.76-2.43)	97.68
Diabetes mellitus	11	542,856/4,288,006	**1.33 (1.22-1.46)**	99.05
Cardiovascular diseases	35	224,187/1,455,452	**1.32 (1.24-1.41)**	98.32
Chronic pulmonary diseases	3	43,234/70,177	**1.32 (1.09-1.61)**	98.66
Ischemic heart diseases (IHD)	13	32,312/284,293	**1.29 (1.06-1.57)**	97.80
AIDS/HIV	3	9,360/38,944	1.29 (0.68-2.45)	98.12
Non-IHD cardiovascular diseases	9	36,195/469,815	**1.28 (1.18-1.39)**	95.85
Cancers	21	100,282/670,646	**1.27 (1.16-1.39)**	98.02
Heart failure	8	35,944/467,693	**1.26 (1.16-1.37)**	95.92
Infectious diseases	5	210,085/622,475	1.22 (0.96-1.55)	99.11
Renal diseases	5	58,972/893,327	**1.20 (1.09-1.33)**	97.12
Respiratory diseases	6	429,170/187,102	**1.20 (1.06-1.36)**	97.78
Breast cancer	8	20,378/243,339	**1.20 (1.04-1.39)**	84.60
Lung cancer	3	9,671/50,160	1.17 (0.98-1.40)	89.29
COVID-19	4	3,848,244/26,019,000	**1.14 (1.02-1.28)**	99.75
Neurological diseases	2	13,460/3,127	1.11 (0.94-1.31)	96.75
Gastrointestinal diseases	5	326,166/4,017,223	1.10 (0.92-1.31)	98.82
Liver diseases	4	11,266/22,584	1.03 (0.83-1.28)	96.41
Dementia	2	12,540/3,127	1.01 (0.89-1.14)	93.67
Other conditions	6	60,710/520,134	1.26 (0.84-1.89)	98.01

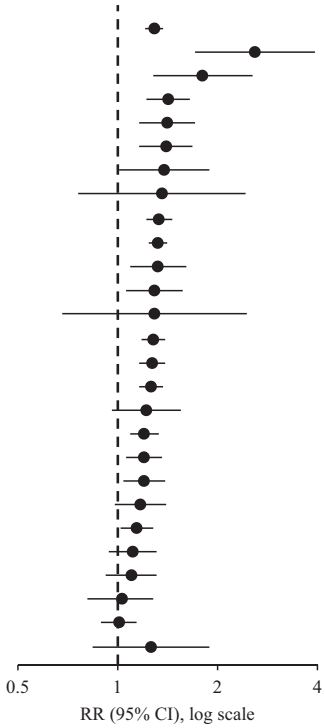

Figure 3 All-cause mortality risk in people with depression versus without depression matched for specific comorbid conditions. RR – relative risk, AIDS/HIV – acquired immunodeficiency syndrome / human immunodeficiency virus infection, COVID-19 – coronavirus 2019 disease. Significant values are highlighted in bold prints.

diseases (RR=1.47, 95% CI: 1.35-1.60, I^2=98.9%, n=36); genitourinary diseases (RR=1.45, 95% CI: 1.19-1.76, I^2=94.0%, n=6); cancers (RR=1.35, 95% CI: 1.20-1.52, I^2=98.8%, n=25); and cerebrovascular diseases (RR=1.27, 95% CI: 1.10-1.47, I^2=95.3%, n=10) (see Figure 2).

Subgroup and meta-regression analyses

Among individuals with depression (with or without any comorbid conditions), those treated with any antidepressant had a reduced risk of all-cause mortality (RR=0.79, 95% CI: 0.68-0.93, I^2=99.2%, n=16) compared to those without antidepressant use (see Figure 5). Moreover, while all-cause mortality risk was still increased in people with antidepressant-treated depression relative to no depression (RR=1.22, 95% CI: 1.10-1.37, I^2=98.9%, n=12), its magnitude in these people was significantly lower (p<0.001) than the all-cause mortality risk observed in the overall analysis for depression versus no depression (RR=2.01, 95% CI: 1.80-2.24, I^2=99.7%, n=92). Regarding antidepressant drug classes, use of serotonin and noradrenaline reuptake inhibitors (SNRIs) was associated with a decreased risk of all-cause mortality (versus no antidepressant use: RR=0.81, 95% CI: 0.65-0.99, I^2=96.7%, n=6), whereas the mortality risk was not decreased significantly with use of selective serotonin reuptake inhibitors (SSRIs) and tricyclic antidepressants (TCAs) (see Figure 5).

In individuals with depression and any comorbid physical conditions, use of any antidepressant (RR=0.69, 95% CI: 0.59-0.81, I^2=98.4%, n=9), of SSRIs (RR=0.75, 95% CI: 0.61-0.92, I^2=98.4%, n=4), of SNRIs (RR=0.74, 95% CI: 0.57-0.96, I^2=94.6%, n=4), and of TCAs (RR=0.78, 95% CI: 0.69-0.87, I^2=82.6%, n=4) was associated with reduced risk of all-cause mortality compared to no antidepressant use (see Figure 5). Individuals using SNRIs had an increased risk of suicide than those using SSRIs (RR=1.55, 95% CI: 1.08-2.22, I^2=5.9%, n=3) (see supplementary information).

Among individuals with depression (with or without any comorbid conditions), use of ECT (versus no ECT use) was associated with reduced mortality risk due to all causes (RR=0.73, 95% CI: 0.66-0.82, I^2=0%, n=6), natural causes (RR=0.76, 95% CI: 0.59-0.97, I^2=12.0%, n=4), and suicide (RR=0.67, 95% CI: 0.53-0.85, I^2=32.3%, n=4) (see Figure 5).

Subgroup analyses by sex did not show a significant difference in depression-associated all-cause mortality risk between men (RR=2.37, 95% CI: 2.06-2.71, I^2=99.9%, n=65) and women (RR=2.27, 95% CI: 1.90-2.71, I^2=99.9%, n=58) (between-group p=0.615). No significant difference in mortality risk for depression versus no depression/general population was detected for age categories (between-group p=0.139) (see Table 1). Subgroup analyses stratified by age and sex revealed a greatly increased mortality risk associated with depres-

Comparison	n	Sample size for depression group/comparison group	RR (95% CI)
All-cause mortality			
MDD vs. no depression/general population	36	869,555/487,657,309	**2.17 (1.69-2.79)**
Dysthymia vs. no depression	3	2,366/253,209	**1.40 (1.30-1.51)**
0-180 days after depression diagnosis	2	NA/NA	**10.80 (6.21-18.77)**
180-365 days after depression diagnosis	2	NA/NA	**3.29 (1.51-7.17)**
1-5 years after depression diagnosis	4	NA/NA	**4.23 (2.25-7.97)**
Late-life depression vs. no depression	3	394/45,200	**2.11 (1.11-4.00)**
Early-life depression vs. no depression	2	819/43,872	**1.73 (1.38-2.17)**
Psychotic depression vs. non-psychotic depression	2	19,174/92,587	**1.61 (1.45-1.78)**
TRD vs. non-TRD	9	127,981/1,025,266	**1.27 (1.16-1.39)**
Unnatural-cause mortality			
MDD vs. no depression/general population	13	381,162/468,337,155	**7.71 (4.72-12.59)**
Depression vs. bipolar disorder (reference: depression)	3	48,854/5,196	0.71 (0.37-1.37)
TRD vs. non-TRD	4	72,284/573,484	**2.30 (1.68-3.14)**
Suicide			
MDD vs. no depression/general population	10	327,886/301,755,306	**12.01 (7.18-20.10)**
Depression vs. bipolar disorder (reference: depression)	3	48,854/5,196	0.64 (0.28-1.50)

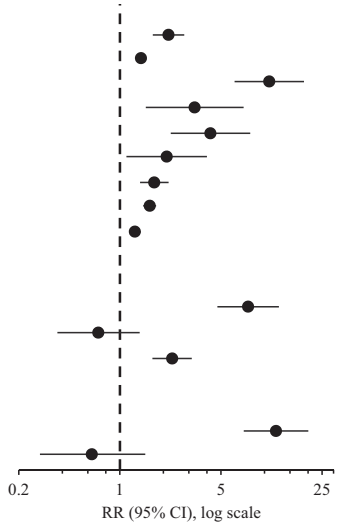

Figure 4 Other characteristics of depression associated with mortality. Regarding natural-cause mortality, no comparison pairs had sufficient number of studies for analyses. RR – relative risk, NA – not available, MDD – major depressive disorder, TRD – treatment-resistant depression. Significant values are highlighted in bold prints.

sion (versus no depression) in females aged <25 years (RR=6.15, 95% CI: 1.89-20.00, I^2=95.0%, n=2), and a substantially increased suicide risk associated with depression (versus no depression) in people aged <25 years (RR=9.91, 95% CI: 6.68-14.69, I^2=90.3%, n=3) and >60 years (RR=13.07, 95% CI: 7.87-21.71, I^2=95.8%, n=5) (see supplementary information).

Mortality risk associated with depression (versus no depression/general population) was higher when the source of study samples was health-system case registers (RR=2.23, 95% CI: 1.94-2.56, I^2=99.9%, n=55) (p for difference with community surveys = 0.036), or health-insurance databases (RR=2.80, 95% CI: 1.65-4.75, I^2=100%, n=15) (p for difference with community surveys = 0.041) (see Table 1).

All-cause mortality risk for depression versus no depression/general population was significantly higher (p<0.001) when people with depression were identified from inpatient settings (RR=2.95, 95% CI: 2.31-3.76, I^2=99.8%, n=29) than in the community (RR=1.57, 95% CI: 1.34-1.85, I^2=91.1%, n=28). Based on data from six continents, there was no significant regional difference in depression-associated mortality risk (versus no depression/general population; between-group p=0.278) (see Table 1).

In meta-regression analyses, depression-associated excess all-cause mortality (versus no depression/general population) decreased with increasing number of adjusted covariates (beta=–0.03, 95% CI: –0.05 to –0.01, p=0.001) and higher Newcastle Ottawa Scale scores (beta=–0.19, 95% CI: –0.35 to –0.02, p=0.026) (see Table 2). In comparisons between depression versus general population, the magnitude of excess all-cause mortality associated with depression increased with higher country/region human development index (beta=11.15, 95% CI: 1.14-21.17, p=0.029) (see supplementary information).

Depression-associated all-cause mortality risk (versus no depression matched for any comorbid condition) increased with higher socio-demographic index (beta=3.21, 95% CI: 1.20-5.22, p=0.002) (see Table 2). Excess natural-cause mortality associated with depression (versus no depression/general population, and versus stratified comparison groups) generally increased with more recent median year of observation period, higher human development index, and larger depression sample size, and decreased with longer observation period, greater number of adjusted covariates, and higher Newcastle Ottawa Scale score (see supplementary information).

DISCUSSION

This large-scale meta-analysis of 268 cohort studies, comparing 10.8 million people with depression versus about 2.8 billion controls, comprehensively quantifies the risk of excess mortality associated with depression. Specifically, we observed a two-fold increased all-cause mortality risk in people with depression versus no depression/general population controls (and in individuals with major depressive disorder versus no depression/general population), and a lower but still significantly 1.3-fold increased all-cause mortality risk versus comorbid condition-matched (mostly physical diseases) non-depression controls. People with depression displayed elevated risk of natural-cause (1.6-fold) and unnatural-cause (5.8-fold) mortality, as well as a 9.9-fold increased risk of suicide, relative to no

Comparison	n	Sample size for treatment group/comparison group	RR (95% CI)
All-cause mortality			
Any AD vs. no AD	16	636,340/688,746	**0.79 (0.68-0.93)**
SSRIs vs. no AD	11	495,609/299,141	0.94 (0.79-1.12)
SNRIs vs. no AD	6	212,573/376,193	**0.81 (0.65-0.99)**
TCAs vs. no AD	9	90,002/386,376	1.03 (0.81-1.31)
ECT vs. no ECT	6	25,841/51,876	**0.73 (0.66-0.82)**
All-cause mortality (samples matched for comorbid conditions)			
Any AD vs. no AD	9	271,358/287,621	**0.69 (0.59-0.81)**
SSRIs vs. no AD	4	90,504/80,173	**0.75 (0.61-0.92)**
SNRIs vs. no AD	4	29,428/184,232	**0.74 (0.57-0.96)**
TCAs vs. no AD	4	27,058/194,879	**0.78 (0.69-0.87)**
Natural-cause mortality			
Any AD vs. no AD	3	26,823/131,053	1.09 (0.99-1.20)
SSRIs vs. no AD	2	10,814/21,842	1.04 (0.90-1.20)
SNRIs vs. no AD	2	1,417/23,957	0.98 (0.86-1.13)
TCAs vs. no AD	2	12,024/22,121	0.92 (0.69-1.22)
ECT vs. no ECT	4	15,558/34,467	**0.76 (0.59-0.97)**
Suicide			
Any AD vs. no AD	3	330,772/211,728	1.13 (0.78-1.64)
ECT vs. no ECT	4	5,393/15,498	**0.67 (0.53-0.85)**

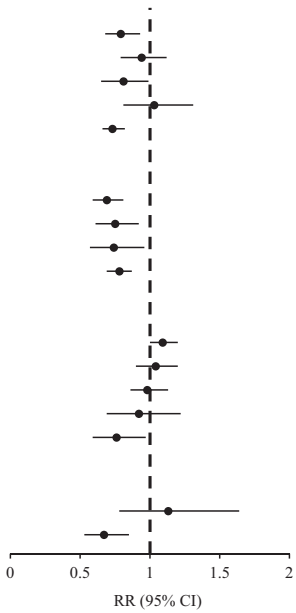

Figure 5 Risk of mortality associated with antidepressant (AD) treatment and electroconvulsive therapy (ECT) in patients with depression. RR – relative risk, SSRIs – selective serotonin reuptake inhibitors, SNRIs – serotonin and noradrenaline reuptake inhibitors, TCAs – tricyclic antidepressants. Significant values are highlighted in bold prints.

depression/general population controls. Dysthymia was also associated with excess mortality, while depression with psychotic symptoms and treatment-resistant depression conferred an incremental mortality risk. Antidepressant-treated patients exhibited decreased all-cause mortality risk versus untreated patients, both in the overall population and, especially, in the sub-populations of patients with depression matched for comorbidities. ECT was associated with a reduced mortality from all causes, natural causes and suicide in patients with depression.

Compared with previous meta-analyses[12,15] dating back to one decade ago, our review included a larger proportion of studies utilizing health-record databases (i.e., case registers, health-insurance databases, or clinic/hospital records), which identify people with depression who have received psychiatric outpatient and/or inpatient care, and who are generally more severely ill than those recruited in community surveys. Moreover, we only included studies defining depression according to ICD or DSM based on diagnostic interviews or a clinician-assigned coded diagnosis derived from health-record databases. This allowed us to avoid misclassification bias due to self-report screening measures, which tend to identify a significant proportion of people with milder or subthreshold symptoms who do not fulfill the clinical diagnosis of depression, resulting in an underestimation of the mortality risk associated with depression.

Our subgroup analyses found no significant differences in mortality risk in men versus women and across age categories. However, subgroup analyses further stratified by age and sex revealed a greatly increased depression-associated all-cause mortality risk in females aged <25 years, and a substantially increased suicide-specific mortality risk in people aged <25 years and >60 years. These represent specific groups requiring multi-component prevention and intervention strategies.

We observed excess depression-associated mortality risk across a broad spectrum of physical comorbidities, with a similar magnitude of risk estimates (RR range: 1.14-1.80). This similar degree of mortality risk may suggest that the association of depression with raised natural-cause deaths in the context of physical comorbidities is mostly attributable to general rather than disease-specific mechanisms, such as inflammatory processes, lifestyle risk factors (e.g., smoking, physical inactivity, unhealthy diet, alcohol use) and depression-related behavioral factors (e.g., poor self-management of health conditions, treatment non-adherence)[1]. Intriguingly, the magnitude of risk estimates associated with incident and prevalent depression was comparable in individuals with physical comorbidities, indicating that depression which occurs prior to the onset of physical diseases and depression emerging after the onset of these diseases may confer similar premature mortality risk, although the involved mechanisms may be different[315,316].

Notably, the risk of excess mortality was most pronounced within 180 days following depression diagnosis (10.8-fold increased risk), as compared to other post-depression time intervals (i.e., 180-365 days and 1-5 years). This finding indicates that the initial few months after depression diagnosis represent a critical period warranting comprehensive assessment, close monitoring and intensive

Table 2 Meta-regression analyses on risk of all-cause mortality in patients with depression

Moderators	Depression vs. no depression/general population			Depression vs. no depression (with comorbid conditions)		
	n	Sample size for depression group/comparison group	Beta (95% CI)	n	Sample size for depression group/comparison group	Beta (95% CI)
Characteristics of overall sample						
Median year of observation period	125	7,407,473/2,737,228,741	0.00 (–0.01 to 0.01)	98	5,881,116/40,284,386	–0.01 (–0.02 to 0.00)
Number of years of observation period	128	7,410,593/2,797,649,150	0.00 (–0.01 to 0.01)	98	5,881,116/40,284,386	0.00 (0.00-0.01)
Mean follow-up duration	71	3,005,306/1,725,191,966	–0.01 (–0.04 to 0.02)	75	3,589,033/28,728,634	**0.01 (0.00-0.02)**
Number of adjusted covariates	128	7,410,593/2,797,649,150	**–0.03 (–0.05** to **–0.01)**	98	5,881,116/40,284,386	0.00 (–0.01 to 0.00)
Human development index	111	7,364,768/1,275,537,198	0.29 (–1.19 to 1.76)	98	5,881,116/40,284,386	2.73 (–0.32 to 5.77)
Socio-demographic index	128	7,410,593/2,797,649,150	0.04 (–1.17 to 1.25)	98	5,881,116/40,284,386	**3.21 (1.20-5.22)**
Newcastle-Ottawa Scale score	128	7,410,593/2,797,649,150	**–0.19 (–0.35** to **–0.02)**	98	5,881,116/40,284,386	0.02 (–0.16 to 0.19)
Characteristics of depression sample						
Sample size	118	7,410,593/2,361,649,150	0.00 (0.00-0.00)	96	5,881,116/40,284,386	0.00 (0.00-0.00)
% with major depressive disorder	44	931,391/855,171,523	0.64 (–0.19 to 1.47)	31	697,579/979,804	0.05 (–0.26 to 0.35)
% with dysthymia	15	244,899/462,835,825	–0.65 (–1.45 to 0.16)	23	663,079/636,807	–0.15 (–0.60 to 0.30)
% with antidepressant exposure	11	967,816/15,846,099	0.04 (–0.27 to 0.34)	17	218,211/3,417,105	–0.18 (–0.45 to 0.10)
Difference between depression and non-depression samples						
% females	71	4,419,696/1,474,617,672	–0.32 (–0.74 to 0.09)	68	1,727,911/11,977,697	0.00 (0.00-0.00)
Mean age	46	3,161,286/934,272,763	0.01 (0.00-0.01)	60	1,065,200/11,009,157	0.00 (0.00-0.00)
% White ethnicity	18	1,930,793/650,325,201	–0.13 (–0.52 to 0.27)	31	800,646/5,758,516	–0.94 (–2.28 to 0.39)
Body mass index	5	45,789/437,289	-	15	523,662/3,391,730	0.01 (–0.07 to 0.09)
% with obesity	12	98,377/662,761	0.95 (–2.86 to 4.76)	14	354,947/4,818,722	0.15 (–0.26 to 0.56)
% with current smoker status	25	1,092,009/5,462,422	–0.21 (–1.54 to 1.11)	25	209,272/2,970,839	–0.2 (–1.02 to 0.62)
% with married status	14	67,153/77,637,512	0.07 (–0.53 to 0.67)	12	552,089/1,263,174	–0.28 (–2.22 to 1.67)
% with alcohol use disorder	24	2,111,327/128,710,342	–0.27 (–0.95 to 0.42)	18	1,008,032/5,892,071	–0.29 (–0.87 to 0.29)
% with substance use disorder	15	2,233,456/153,465,672	–1.57 (–6.49 to 3.36)	11	460,289/6,325,928	–0.3 (–1.37 to 0.77)
% with diabetes	27	2,047,125/73,283,440	–0.46 (–1.94 to 1.01)	45	1,528,503/9,203,064	**1.21 (0.08-2.33)**
% with cancers	14	1,135,834/71,005,967	2.35 (–0.84 to 5.54)	17	1,068,797/5,269,880	**2.49 (0.35-4.63)**
% with renal diseases	7	1,822,787/71,983,909	-	17	758,904/3,662,367	1.88 (–1.04 to 4.80)

Significant values are highlighted in bold prints

treatment to optimize illness outcome and reduce mortality risk, in particular from suicide.

We found that the presence of psychotic symptoms conferred an incremental mortality risk associated with depression. As these symptoms might not be readily identified in depression[317,318], careful evaluation is required to facilitate their early detection and effective management. Moreover, treatment-resistant depression, which affects at least 30% of depressed people[319], was associated with 27% higher risk for all-cause mortality and a 2.3-fold increased risk for unnatural deaths relative to non-treatment-resistant depression. Previous research suggested that this increased mortality risk is driven largely by suicide and other external causes[203,252]. However, common chronic physical comorbidities such as cardiovascular diseases and diabetes mellitus are also over-represented in patients with this condition[319,320]. Early identification of treatment-refractory status followed by provision of adequate management is therefore needed to reduce the disproportionate morbidity and mortality associated with this subtype of depression.

To our knowledge, this is the first meta-analysis comprehensively assessing mortality risk associated with antidepressant treatment in people with depression. In the overall analyses (i.e., including depression with and without comorbid conditions), we observed a significant mortality-reducing effect of any antidepressant and of SNRIs, relative to non-use of antidepressants. These data were reinforced by the observation that the magnitude of increased mortality in people with antidepressant-treated depression versus no depression was significantly lower (RR=1.22) than in the overall primary analysis of depression versus no depression (RR=2.01). Depressive symptom alleviation by antidepressant treatment might contribute to better physical health outcomes via enhanced self-management of physical conditions, improved treatment adherence, and increased engagement in healthy lifestyle behaviors[321]. Moreover, the observed protective effect of antidepressant treatment might also be due to factors such as improved glycemic control[322], reduction of pro-inflammatory state[323-325], and enhanced motor function[326].

Our analyses generally revealed comparable mortality risk between antidepressant drug classes. Nonetheless, SNRI use was associated with a higher suicide risk compared to SSRI use. A recent network meta-analysis based on randomized controlled trials (RCTs) reported venlafaxine as the only antidepressant linked to significantly increased risk of suicidal behavior or ideation compared to placebo and other antidepressants in children and adolescents[28]. However, our comparison analyses between venlafaxine and fluoxetine, which were based on only two studies, revealed no significant difference in suicide risk.

The US Food and Drug Administration issued a black-box warning in 2004 that antidepressants might have a differential effect on suicide risk across age groups, with an elevated risk in young people, no association in middle age, and a protective effect in the elderly[327]. Limited available research comparing the effect of antidepressant versus no antidepressant use on suicide risk in people with depression precluded us from investigating age-specific associations between suicide and antidepressant treatment.

Our pooled analyses demonstrated that use of ECT was associated with a reduced risk of all-cause, natural-cause and suicide-related deaths in people with depression, further supporting its critical role as an effective treatment for severe depression.

Large-scale research utilizing health-record databases with long observation periods would be required to better clarify the effect of antidepressant treatment and ECT on suicide risk in people with depression, which otherwise could unlikely be adequately captured (as a rare outcome event) and investigated in the context of RCTs.

None of the included studies compared mortality risk in people with depression who had received versus those who had not received psychotherapies or neuromodulation therapies, thereby precluding us from performing subgroup analyses to explore the associations between these treatment modalities and mortality risk associated with depression.

In line with a prior meta-analysis[12], our meta-regression models showed that an increasing number of adjusted covariates and higher study quality decreased the magnitude of elevated mortality risk in people with depression, suggesting that residual confounding might contribute to the reported excess mortality. This potential bias was partly addressed by our selection of the reported risk estimates adjusted for the most comprehensive set of covariates per included study into the pooled analyses. We also found that more recent median study year of investigation accentuated the excess natural-cause mortality risk in depression (versus no depression), indicating that people with depression have not benefited equally from recent enhancement of health care and life expectancy improvement compared to the general population.

Our results suggest that a higher human development index, which measures levels of social and economic development in a specific country/region[55], increases the risk of premature mortality in people with depression versus the general population. Despite better access to health services, it is recognized that individuals from regions with high social and economic development may be more likely to experience an escalated stress in relation to social exclusion, unemployment, working conditions, lack of family and social support, and violence, which are closely associated with suicidal behaviors and other non-communicable diseases[328]. One alternative explanation is that depression is more likely to be underdiagnosed and under-reported in less developed countries, resulting in apparently lower mortality risk in regions with low human development index.

Some limitations warrant consideration in interpreting our results. First, there was significant heterogeneity across studies regarding the mortality risk associated with depression. We attempted to assess the sources of heterogeneity via subgroup and meta-regression analyses. However, as data for other potentially relevant variables, such as socio-economic status and lifestyle risk factors, were not adequately captured in most included studies, sources of heterogeneity could not be further explored. Second, the included studies were observational in nature, and thus causality cannot be inferred regarding the moderating or aggravating factors that we identified. Third, although 268 studies were included in the meta-analysis, findings of some subgroup analyses (e.g., several specific physical comorbidities, some characteristics of depression, use of antidepressants and ECT) were based on few studies, and should be re-evaluated when more studies have been conducted in this respect.

Notwithstanding these limitations, this study is the most comprehensive meta-analysis to date quantifying the mortality risk associated with depression, encompassing a wide range of comorbid conditions, taking into account an array of potential aggravating and attenuating factors, and evaluating the protective effect of antidepressant treatment and ECT against excess mortality. The study findings thus facilitate formulation of relevant and actionable targets for clinicians and allied health professionals, researchers, health system administrators, policy makers, patients and caregivers, that can be leveraged to effectively reduce the avoidable mortality gap associated with depression.

ACKNOWLEDGEMENTS

C.U. Correll and W.C. Chang contributed equally to this work. Supplementary information on this study is available at https://osf.io/dsj2f/.

REFERENCES

1. Herrman H, Patel V, Kieling C et al. Time for united action on depression: a Lancet-World Psychiatric Association Commission. Lancet 2022;399:957-1022.
2. Malhi GS, Mann JJ. Depression. Lancet 2018;392:2299-312.
3. Cipriani A, Furukawa TA, Salanti G et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. Lancet 2018;391:1357-66.
4. Cuijpers P, Quero S, Noma H et al. Psychotherapies for depression: a network meta-analysis covering efficacy, acceptability and long-term outcomes of all main treatment types. World Psychiatry 2021;20:283-93.
5. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022;9:137-50.
6. Berk M, Köhler-Forsberg O, Turner M et al. Comorbidity between major depressive disorder and physical diseases: a comprehensive review of epidemiology, mechanisms and management. World Psychiatry 2023;22:366-87.
7. Stubbs B, Vancampfort D, Veronese N et al. Depression and physical health multimorbidity: primary data and country-wide meta-analysis of population data from 190593 people across 43 low- and middle-income countries. Psychol Med 2017;47:2107-17.
8. Machado MO, Veronese N, Sanches M et al. The association of depression and all-cause and cause-specific mortality: an umbrella review of systematic reviews and meta-analyses. BMC Med 2018;16:112.
9. Chan JKN, Correll CU, Wong CSM et al. Life expectancy and years of potential life lost in people with mental disorders: a systematic review and meta-analysis. EClinicalMedicine 2023;65:102294.
10. Pan YJ, Yeh LL, Chan HY et al. Excess mortality and shortened life expectancy in people with major mental illnesses in Taiwan. Epidemiol Psychiatr Sci 2020;29:e156.
11. Weye N, Momen NC, Christensen MK et al. Association of specific mental disorders with premature mortality in the Danish population using alternative measurement methods. JAMA Netw Open 2020;3:e206646.
12. Cuijpers P, Vogelzangs N, Twisk J et al. Comprehensive meta-analysis of excess mortality in depression in the general community versus patients with specific illnesses. Am J Psychiatry 2014;171:453-62.
13. Baxter AJ, Page A, Whiteford HA. Factors influencing risk of premature mortality in community cases of depression: a meta-analytic review. Epidemiol Res Int 2011;832945:1-13.
14. Cuijpers P, Smit F. Excess mortality in depression: a meta-analysis of community studies. J Affect Disord 2002;72:227-36.
15. Walker ER, McGee RE, Druss BG. Mortality in mental disorders and global disease burden implications: a systematic review and meta-analysis. JAMA Psychiatry 2015;72:334-41.
16. Miloyan B, Fried E. A reassessment of the relationship between depression and all-cause mortality in 3,604,005 participants from 293 studies. World Psychiatry 2017;162:219-20.
17. Levis B, Benedetti A, Levis AW et al. Selective cutoff reporting in studies of diagnostic test accuracy: a comparison of conventional and individual patient-data meta-analysis of the Patient Health Questionnaire-9 depression screening tool. Am J Epidemiol 2017;185:954-64.
18. Bartoli F, Di Brita C, Crocamo C et al. Early post-stroke depression and mortality: meta-analysis and meta-regression. Front Psychiatry 2018;9:530.
19. Fan H, Yu W, Zhang Q et al. Depression after heart failure and risk of cardiovascular and all-cause mortality: a meta-analysis. Prev Med 2014;63:36-42.
20. Wei J, Hou R, Zhang X et al. The association of late-life depression with all-cause and cardiovascular mortality among community-dwelling older adults: systematic review and meta-analysis. Br J Psychiatry 2019;215:449-55.
21. Low CE, Yau CE, Tan RY et al. Association of depression with all-cause and cancer-specific mortality in older adults with cancer: systematic review, meta-analysis, and meta-regression. J Geriatr Oncol 2024;15:101700.
22. Wang YH, Li JQ, Shi JF et al. Depression and anxiety in relation to cancer incidence and mortality: a systematic review and meta-analysis of cohort studies. Mol Psychiatry 2020;25:1487-99.
23. Park M, Katon WJ, Wolf FM. Depression and risk of mortality in individuals with diabetes: a meta-analysis and systematic review. Gen Hosp Psychiatry 2013;35:217-25.
24. van Dooren FE, Nefs G, Schram MT et al. Depression and risk of mortality in people with diabetes mellitus: a systematic review and meta-analysis. PLoS One 2013;8:e57058.
25. Brandão DJ, Fontenelle LF, da Silva SA et al. Depression and excess mortality in the elderly living in low- and middle-income countries: systematic review and meta-analysis. Int J Geriatr Psychiatry 2019;34:22-30.
26. Kato M, Hori H, Inoue T et al. Discontinuation of antidepressants after remission with antidepressant medication in major depressive disorder: a systematic review and meta-analysis. Mol Psychiatry 2021;26:118-33.
27. Kishi T, Ikuta T, Sakuma K et al. Antidepressants for the treatment of adults with major depressive disorder in the maintenance phase: a systematic review and network meta-analysis. Mol Psychiatry 2023;28:402-9.
28. Cipriani A, Zhou X, Del Giovane C et al. Comparative efficacy and tolerability of antidepressants for major depressive disorder in children and adolescents: a network analysis. Lancet 2016;388:881-90.
29. Stone M, Laughren T, Jones ML et al. Risk of suicidality in clinical trials of antidepressants in adults: analysis of proprietary data submitted to US Food and Drug Administration. BMJ 2009;339:b2880.
30. Suominen K, Haukka J, Valtonen HM et al. Outcome of patients with major depressive disorder after serious suicide attempt. J Clin Psychiatry 2009;70:1372-8.
31. Zivin K, Kim HM, Yosef M et al. Antidepressant medication treatment and risk of death. J Clin Psychopharmacol 2016;36:445-52.
32. Coupland C, Hill T, Morriss R et al. Antidepressant use and risk of adverse outcomes in people aged 20-64 years: cohort study using a primary care database. BMC Med 2018;16:36.
33. Coupland C, Hill T, Morriss R et al. Antidepressant use and risk of suicide and attempted suicide or self harm in people aged 20 to 64: cohort study using a primary care database. BMJ 2015;350:h517.
34. Taylor CB, Youngblood ME, Catellier D et al. Effects of antidepressant medication on morbidity and mortality in depressed patients after myocardial infarction. Arch Gen Psychiatry 2005;62:792-8.
35. Kim JH, Song YK, Jang HY et al. Major adverse cardiovascular events in antidepressant users within patients with ischemic heart diseases: a nationwide cohort study. J Clin Psychopharmacol 2020;40:475-81.
36. Kollhorst B, Jobski K, Krappweis J et al. Antidepressants and the risk of death in older patients with depression: a population-based cohort study. PLoS One 2019;14:e0215289.
37. Bansal N, Hudda M, Payne RA et al. Antidepressant use and risk of adverse outcomes: population-based cohort study. BJPsych Open 2022;8:e164.
38. Munk-Olsen T, Laursen TM, Videbech P et al. All-cause mortality among recipients of electroconvulsive therapy: register-based cohort study. Br J Psychiatry 2007;190:435-9.
39. Peltzman T, Shiner B, Watts BV. Effects of electroconvulsive therapy on short-term suicide mortality in a risk-matched patient population. J ECT 2020;36:187-92.
40. Watts BV, Peltzman T, Shiner B. Mortality after electroconvulsive therapy. Br J Psychiatry 2021;219:588-93.
41. Blumberger DM, Seitz DP, Herrmann N et al. Low medical morbidity and mortality after acute courses of electroconvulsive therapy in a population-based sample. Acta Psychiatr Scand 2017;136:583-93.
42. Carney S, Cowen P, Geddes J et al. Efficacy and safety of electroconvulsive therapy in depressive disorders: a systematic review and meta-analysis. Lancet 2003;361:799-808.
43. Kaster TS, Blumberger DM, Gomes T et al. Risk of suicide death following electroconvulsive therapy treatment for depression: a propensity score-weighted, retrospective cohort study in Canada. Lancet Psychiatry 2022;9:435-46.
44. Rhee TG, Sint K, Olfson M et al. Association of ECT with risks of all-cause mortality and suicide in older Medicare patients. Am J Psychiatry 2021;178:1089-97.
45. Rönnqvist I, Nilsson FK, Nordenskjöld A. Electroconvulsive therapy and the risk of suicide in hospitalized patients with major depressive disorder. JAMA Netw Open 2021;4:e2116589.
46. Page MJ, McKenzie JE, Bossuyt PM et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71.
47. Correll CU, Solmi M, Croatto G et al. Mortality in people with schizophrenia: a systematic review and meta-analysis of relative risk and aggravating or attenuating factors. World Psychiatry 2022;21:248-71.
48. Wells GA, Shea B, O'Connell D et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses. www.ohri.ca.
49. Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA 1998;280:1690-1.
50. Biazus TB, Beraldi GH, Tokeshi L et al. All-cause and cause-specific mortality among people with bipolar disorder: a large-scale systematic review and meta-analysis. Mol Psychiatry 2023;28:2508-24.
51. Chen D, Ejlskov L, Laustsen LM et al. The role of socioeconomic position in the association between mental disorders and mortality: a systematic review and meta-analysis. JAMA Psychiatry 2024;81:125-34.

52. Higgins JP, Thompson SG, Deeks JJ et al. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-60.
53. Egger M, Davey SG, Schneider M et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629-34.
54. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56:455-63.
55. United Nations Development Programme. Human development index. https://hdr.undp.org.
56. Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 (GBD 2019) Socio-Demographic Index (SDI) 1950-2019. Seattle: Institute for Health Metrics and Evaluation, 2020.
57. Abrams TE, Vaughan-Sarrazin M, Rosenthal GE. Influence of psychiatric comorbidity on surgical mortality. Arch Surg 2010;145:947-53.
58. Abrams TE, Vaughan-Sarrazin M, Rosenthal GE. Preexisting comorbid psychiatric conditions and mortality in nonsurgical intensive care patients. Am J Crit Care 2010;19:241-9.
59. Adelborg K, Schmidt M, Sundbøll J et al. Mortality risk among heart failure patients with depression: a nationwide population-based cohort study. J Am Heart Assoc 2016;5:e004137.
60. Aguayo E, Lyons R, Juo YY et al. Impact of new-onset postoperative depression on readmission outcomes after surgical coronary revascularization. J Surg Res 2019;233:50-6.
61. Ahmadi N, Moss L, Simon E et al. Efficacy and long-term clinical outcome of comorbid posttraumatic stress disorder and major depressive disorder after electroconvulsive therapy. Depress Anxiety 2016;33:640-7.
62. Albert NM, Fonarow GC, Abraham WT et al. Depression and clinical outcomes in heart failure: an OPTIMIZE-HF analysis. Am J Med 2009;122:366-73.
63. Allgulander C, Lavori PW. Causes of death among 936 elderly patients with 'pure' anxiety neurosis in Stockholm County, Sweden, and in patients with depressive neurosis or both diagnoses. Compr Psychiatry 1993;34:299-302.
64. Allgulander C. Suicide and mortality patterns in anxiety neurosis and depressive neurosis. Arch Gen Psychiatry 1994;51:708-12.
65. Almeida OP, Hankey GJ, Yeap BB et al. Mortality among people with severe mental disorders who reach old age: a longitudinal study of a community-representative sample of 37,892 men. PLoS One 2014;9:e111882.
66. Alotaibi R, Halbesma N, Wild SH et al. Severe depression and all-cause and cause-specific mortality in Scotland: 20 year national cohort study. BJPsych Open 2024;10:e28.
67. An H, Yang HW, Oh DJ et al. Mood disorders increase mortality mainly through dementia: a community-based prospective cohort study. Aust N Z J Psychiatry 2022;56:1017-24.
68. Angst J, Angst F, Gerber-Werder R et al. Suicide in 406 mood-disorder patients with and without long-term medication: a 40 to 44 years' follow-up. Arch Suicide Res 2005;9:279-300.
69. Angst J, Hengartner MP, Gamma A et al. Mortality of 403 patients with mood disorders 48 to 52 years after their psychiatric hospitalisation. Eur Arch Psychiatry Clin Neurosci 2013;263:425-34.
70. Arya S, Lee S, Zahner GJ et al. The association of comorbid depression with mortality and amputation in veterans with peripheral artery disease. J Vasc Surg 2018;68:536-45.
71. Bach L, Kalder M, Kostev K. Depression and sleep disorders are associated with early mortality in women with breast cancer in the United Kingdom. J Psychiatr Res 2021;143:481-4.
72. Baldessarini RJ, Tondo L. Suicidal risks in 12 DSM-5 psychiatric disorders. J Affect Disord 2020;271:66-73.
73. Balogun RA, Abdel-Rahman EM, Balogun SA et al. Association of depression and antidepressant use with mortality in a large cohort of patients with nondialysis-dependent CKD. Clin J Am Soc Nephrol 2012;7:1793-800.
74. Balogun RA, Omotoso BA, Xin W et al. Major depression and long-term outcomes of acute kidney injury. Nephron 2017;135:23-30.
75. Banta JE, Andersen RM, Young AS et al. Psychiatric comorbidity and mortality among veterans hospitalized for congestive heart failure. Mil Med 2010;175:732-41.
76. Black DW, Warrack G, Winokur G. The Iowa record-linkage study. I. Suicides and accidental deaths among psychiatric patients. Arch Gen Psychiatry 1985;42:71-5.
77. Black DW, Winokur G, Nasrallah A. Is death from natural causes still excessive in psychiatric patients? A follow-up of 1593 patients with major affective disorder. J Nerv Ment Dis 1987;175:674-80.
78. Boschesi Barros V, Fortti Vianna Schmidt F, Chiavegatto Filho ADP. Mortality, survival, and causes of death in mental disorders: comprehensive prospective analyses of the UK Biobank cohort. Psychol Med 2023;53:3480-9.
79. Boyd CA, Benarroch-Gampel J, Sheffield KM et al. The effect of depression on stage at diagnosis, treatment, and survival in pancreatic adenocarcinoma. Surgery 2012;152:403-13.
80. Brådvik L, Berglund M. Late mortality in severe depression. Acta Psychiatr Scand 2001;103:111-6.
81. Brådvik L, Mattisson C, Bogren M et al. Long-term suicide risk of depression in the Lundby cohort 1947-1997 – severity and gender. Acta Psychiatr Scand 2008;117:185-91.
82. Brandel MG, Hirshman BR, McCutcheon BA et al. The association between psychiatric comorbidities and outcomes for inpatients with traumatic brain injury. J Neurotrauma 2017;34:1005-16.
83. Brodaty H, MacCuspie-Moore CM, Tickle L et al. Depression, diagnostic subtype and death: a 25 year follow-up study. J Affect Disord 1997;46:233-42.
84. Brouwers C, Christensen SB, Damen NL et al. Antidepressant use and risk for mortality in 121,252 heart failure patients with or without a diagnosis of clinical depression. Int J Cardiol 2016;203:867-73.
85. Busby J, Mills K, Zhang SD et al. Selective serotonin reuptake inhibitor use and breast cancer survival: a population-based cohort study. Breast Cancer Res 2018;20:4.
86. Butnoriene J, Bunevicius A, Saudargiene A et al. Metabolic syndrome, major depression, generalized anxiety disorder, and ten-year all-cause and cardiovascular mortality in middle aged and elderly patients. Int J Cardiol 2015;190:360-6.
87. Byers AL, Covinsky KE, Barnes DE et al. Dysthymia and depression increase risk of dementia and mortality among older veterans. Am J Geriatr Psychiatry 2012;20:664-72.
88. Cabello M, Borges G, Lara E et al. The relationship between all-cause mortality and depression in different gender and age groups of the Spanish population. J Affect Disord 2020;266:424-8.
89. Callaghan RC, Veldhuizen S, Jeysingh T et al. Patterns of tobacco-related mortality among individuals diagnosed with schizophrenia, bipolar disorder, or depression. J Psychiatr Res 2014;48:102-10.
90. Carney RM, Blumenthal JA, Catellier D et al. Depression as a risk factor for mortality after acute myocardial infarction. Am J Cardiol 2003;92:1277-81.
91. Carney RM, Howells WB, Blumenthal JA et al. Heart rate turbulence, depression, and survival after acute myocardial infarction. Psychosom Med 2007;69:4-9.
92. Carney RM, Freedland KE, Steinmeyer B et al. Depression and five year survival following acute myocardial infarction: a prospective study. J Affect Disord 2008;109:133-8.
93. Carney RM, Steinmeyer B, Freedland KE et al. Nighttime heart rate and survival in depressed patients post acute myocardial infarction. Psychosom Med 2008;70:757-63.
94. Chan VK, Cheung EC, Chan SS et al. Mortality-causing mechanisms and healthcare resource utilisation of treatment-resistant depression: a six-year population-based cohort study. Lancet Reg Health West Pac 2022;22:100426.
95. Chang HJ, Lin HC, Lee HC et al. Risk of mortality among depressed younger patients: a five-year follow-up study. J Affect Disord 2009;113:255-62.
96. Che SE, Geun GY, Lee JY et al. Trends in patient suicide rate after psychiatric discharge in Korea from 2010 to 2018: a nationwide population-based study. J Affect Disord 2023;323:860-5.
97. Chen YH, Lee HC, Lin HC. Mortality among psychiatric patients in Taiwan – results from a universal National Health Insurance programme. Psychiatry Res 2010;178:160-5.
98. Chen R, Hu Z, Wei L et al. Socioeconomic status and survival among older adults with dementia and depression. Br J Psychiatry 2014;204:436-40.
99. Chen SJ, Chang CH, Chen KC et al. Association between depressive disorders and risk of breast cancer recurrence after curative surgery. Medicine 2016;95:e4547.
100. Chen HM, Hung TH, Chou SY et al. Three-year mortality rate of suicide attempters in consultation-liaison service. Int J Psychiatry Clin Pract 2016;20:254-9.
101. Chen HM, Yang YH, Chen KJ et al. Antidepressants reduced risk of mortality in patients with diabetes mellitus: a population-based cohort study in Taiwan. J Clin Endocrinol Metab 2019;104:4619-25.
102. Chen C, Chan HY, Yeh LL et al. Longitudinal factors associated with mortality in older patients with mood disorders. J Affect Disord 2021;278:12500.
103. Cheon DY, Park YM, Park MS et al. Depression and risk of stroke and mortality after percutaneous coronary intervention: a nationwide population study. J Intern Med 2024;296:468-80.
104. Chierzi F, Stivanello E, Musti MA et al. Cancer mortality in common mental disorders: a 10-year retrospective cohort study. Soc Psychiatry Psychiatr Epidemiol 2023;58:309-18.
105. Chiu M, Vigod S, Rahman F et al. Mortality risk associated with psychological

distress and major depression: a population-based cohort study. J Affect Disord 2018;234:117-23.
106. Choi JW, Lee KS, Kim TH et al. Suicide risk after discharge from psychiatric care in South Korea. J Affect Disord 2019;251:287-92.
107. Choi JW, Lee SG, Kim TH et al. Poststroke suicide risk among older adults in South Korea: a retrospective longitudinal cohort study. Int J Geriatr Psychiatry 2020;35:282-9.
108. Christensen GT, Maartensson S, Osler M. The association between depression and mortality – a comparison of survey- and register-based measures of depression. J Affect Disord 2017;210:111-4.
109. Chung AN, Su SS, Tsai SY et al. Sex-specific incidences and risk profiles of suicide mortality in people with alcohol dependence in Taiwan. Addiction 2022;117:3058-68.
110. Ciompi L, Medvecka J. Etude comparative de la mortalité à long terme dans les maladies mentales. Schweiz Arch Neurol Neurochir Psychiatr 1976;118:111-35.
111. Connerney I, Sloan RP, Shapiro PA et al. Depression is associated with increased mortality 10 years after coronary artery bypass surgery. Psychosom Med 2010;72:874-81.
112. Coryell W, Noyes R, Clancy J. Excess mortality in panic disorder. A comparison with primary unipolar depression. Arch Gen Psychiatry 1982;39:701-3.
113. Crump C, Sundquist K, Sundquist J et al. Sociodemographic, psychiatric and somatic risk factors for suicide: a Swedish national cohort study. Psychol Med 2014;44:279-89.
114. Cully JA, Johnson M, Moffett ML et al. Depression and anxiety in ambulatory patients with heart failure. Psychosomatics 2009;50:592-8.
115. Damián J, Pastor-Barriuso R, Valderrama-Gama E et al. Association of detected depression and undetected depressive symptoms with long-term mortality in a cohort of institutionalised older people. Epidemiol Psychiatr Sci 2017;26:189-98.
116. Dao TK, Chu D, Springer J et al. Clinical depression, posttraumatic stress disorder, and comorbid depression and posttraumatic stress disorder as risk factors for in-hospital mortality after coronary artery bypass grafting surgery. J Thorac Cardiovasc Surg 2010;140:606-10.
117. Das-Munshi J, Chang CK, Schofield P et al. Depression and cause-specific mortality in an ethnically diverse cohort from the UK: 8-year prospective study. Psychol Med 2019;49:1639-51.
118. Davidson IA, Dewey ME, Copeland JR. The relationship between mortality and mental disorder: evidence from the Liverpool longitudinal study. Int J Geriatr Psychiatry 1988;3:95-8.
119. Davydow DS, Ribe AR, Pedersen HS et al. The association of unipolar depression with thirty-day mortality after hospitalization for infection: a population-based cohort study in Denmark. J Psychosom Res 2016;89:32-8.
120. Dickens C, McGowan L, Percival C et al. Depression is a risk factor for mortality after myocardial infarction: fact or artifact? J Am Coll Cardiol 2007;49:1834-40.
121. Dickerson F, Origoni A, Rowe K et al. Risk factors for natural cause mortality in a cohort of 1494 persons with serious mental illness. Psychiatry Res 2021; 298:113755.
122. Didden C, Egger M, Folb N et al. The contribution of noncommunicable and infectious diseases to the effect of depression on mortality: a longitudinal causal mediation analysis. Epidemiology 2025;36:88-98.
123. Döme P, Kunovszki P, Takács P et al. Clinical characteristics of treatment-resistant depression in adults in Hungary: real-world evidence from a 7-year-long retrospective data analysis. PLoS One 2021;16:e0245510.
124. Evans NJ, Whitlock FA. Mortality and late-onset affective disorder. J Affect Disord 1983;5:297-304.
125. Fang SC, Hung CC, Huang CY et al. Influence of baseline psychiatric disorders on mortality and suicide and their associations with gender and age in patients with methamphetamine use disorder. Int J Ment Health Addiction 2025;23:263-80.
126. Fernando DT, Clapperton A, Berecki-Gisolf J. Suicide following hospital admission for mental health conditions, physical illness, injury and intentional self-harm in Victoria, Australia. PLoS One 2022;17:e0271341.
127. Findley PA, Banerjea R, Sambamoorthi U. Excess mortality associated with mental illness and substance use disorders among veteran clinic users with spinal cord injury. Disabil Rehabil 2011;33:1608-15.
128. Fleetwood K, Wild SH, Smith DJ et al. Association of severe mental illness with stroke outcomes and process-of-care quality indicators: nationwide cohort study. Br J Psychiatry 2022;221:394-401.
129. Fleetwood KJ, Wild SH, Licence KAM et al. Severe mental illness and type 2 diabetes outcomes and complications: a nationwide cohort study. Diabetes Care 2023;46:1363-71.
130. Freedland KE, Hesseler MJ, Carney RM et al. Major depression and long-term survival of patients with heart failure. Psychosom Med 2016;78:896-903.
131. Fuller-Thomson E, Lung Y, West KJ et al. Suboptimal baseline mental health associated with 4-month premature all-cause mortality: findings from 18 years of follow-up of the Canadian National Population Health Survey. J Psychosom Res 2020;136:110176.
132. Gale CR, Batty GD, Osborn DP et al. Association of mental disorders in early adulthood and later psychiatric hospital admissions and mortality in a cohort study of more than 1 million men. Arch Gen Psychiatry 2012;69:823-31.
133. Gallo JJ, Bogner HR, Morales KH et al. Depression, cardiovascular disease, diabetes, and two-year mortality among older, primary-care patients. Am J Geriatr Psychiatry 2005;13:748-55.
134. Gasse C, Laursen TM, Baune BT. Major depression and first-time hospitalization with ischemic heart disease, cardiac procedures and mortality in the general population: a retrospective Danish population-based cohort study. Eur J Prev Cardiol 2014;1:532-40.
135. Gibbs A, Maripuu M, Öhlund L et al. COVID-19-associated mortality in individuals with serious mental disorders in Sweden during the first two years of the pandemic – a population-based register study. BMC Psychiatry 2024;24:189.
136. Goodwin JS, Zhang DD, Ostir GV. Effect of depression on diagnosis, treatment, and survival of older women with breast cancer. J Am Geriatr Soc 2004;52:106-11.
137. Graversen SB, Pedersen HS, Ribe AR et al. The significance of depression for short-term readmission and mortality after a pneumonia admission. Med Care 2021;59:872-80.
138. Gronemann FH, Jørgensen MB, Nordentoft M et al. Treatment-resistant depression and risk of all-cause mortality and suicidality in Danish patients with major depression. J Psychiatr Res 2021;135:197-202.
139. Gu WJ, Zhang LM, Wang CM et al. Pre-intensive care unit use of selective serotonin reuptake inhibitors and mortality in critically ill adults with mental disorders: analysis from the MIMIC-IV database. Transl Psychiatry 2023;13:187.
140. Guan NC, Termorshuizen F, Laan W et al. Cancer mortality in patients with psychiatric diagnoses: a higher hazard of cancer death does not lead to a higher cumulative risk of dying from cancer. Soc Psychiatry Psychiatr Epidemiol 2013;48:1289-95.
141. Gustafsson TT, Taipale H, Lähteenvuo M et al. Cause-specific mortality in treatment-resistant major depression: population-based cohort study. J Affect Disord 2025;368:136-42.
142. Hagnell O, Lanke J, Rorsman B. Suicide rates in the Lundby study: mental illness as a risk factor for suicide. Neuropsychobiology 1981;7:248-53.
143. Hamer M, Stamatakis E, Steptoe A. Psychiatric hospital admissions, behavioral risk factors, and all-cause mortality: the Scottish Health Survey. Arch Intern Med 2008;168:2474-9.
144. Han X, Hou C, Yang H et al. Disease trajectories and mortality among individuals diagnosed with depression: a community-based cohort study in UK Biobank. Mol Psychiatry 2021;26:6736-46.
145. Hassan L, Peek N, Lovell K et al. Disparities in COVID-19 infection, hospitalisation and death in people with schizophrenia, bipolar disorder, and major depressive disorder: a cohort study of the UK Biobank. Mol Psychiatry 2022; 27:1248-55.
146. Hassan L, Sawyer C, Peek N et al. Heightened COVID-19 mortality in people with severe mental illness persists after vaccination: a cohort study of Greater Manchester residents. Schizophr Bull 2023;49:275-84.
147. Hayano J, Carney RM, Watanabe E et al. Interactive associations of depression and sleep apnea with adverse clinical outcomes after acute myocardial infarction. Psychosom Med 2012;74:832-9.
148. Hedström AK, Bellocco R, Hössjer O et al. The relationship between nightmares, depression and suicide. Sleep Med 2021;77:1-6.
149. Henriksson M, Nyberg J, Schiöler L et al. Cause-specific mortality in Swedish males diagnosed with non-psychotic mental disorders in late adolescence: a prospective population-based study. J Epidemiol Community Health 2018;72: 582-8.
150. Herweijer E, Wang J, Hu K et al. Overall and cervical cancer survival in patients with and without mental disorders. JAMA Netw Open 2023;6:e2336213.
151. Ho CSh, Jin A, Nyunt MS et al. Mortality rates in major and subthreshold depression: 10-year follow-up of a Singaporean population cohort of older adults. Postgrad Med 2016;128:642-7.
152. Ho MTH, Chan JKN, Lo HKY F et al. Risk of mortality and complications in people with depressive disorder and diabetes mellitus: a 20-year population-based propensity score-matched cohort study. Eur Neuropsychopharmacol 2024;92:10-8.
153. Holmstrand C, Bogren M, Mattisson C et al. Long-term suicide risk in no, one or more mental disorders: the Lundby Study 1947-1997. Acta Psychiatr Scand 2015;132:459-69.
154. Holwerda TJ, Schoevers RA, Dekker J et al. The relationship between general-

ized anxiety disorder, depression and mortality in old age. Int J Geriatr Psychiatry 2007;22:241-9.
155. Hong M, Lee SM, Han KM et al. Suicide death and other-cause mortality in psychiatric patients: a South Korean study using nationwide claims data. J Affect Disord 2024;352:288-95.
156. Høye A, Nesvåg R, Reichborn-Kjennerud T et al. Sex differences in mortality among patients admitted with affective disorders in North Norway: a 33-year prospective register study. Bipolar Disord 2016;18:272-81.
157. Huang CL, Tsai IJ, Lin WC et al. Reduced mortality in patients with extended duration of methadone maintenance treatment: a five-year retrospective nationwide study. Psychol Med 2023;53:722-30.
158. Jayadevappa R, Malkowicz SB, Chhatre S et al. The burden of depression in prostate cancer. Psychooncology 2012;21:1338-45.
159. Jeong HG, Lee JJ, Lee SB et al. Role of severity and gender in the association between late-life depression and all-cause mortality. Int Psychogeriatr 2013;25:677-84.
160. Jeong JH, Um YH, Ko SH et al. Depression and mortality in people with type 2 diabetes mellitus, 2003 to 2013: a nationwide population-based cohort study. Diabetes Metab J 2017;41:296-302.
161. Jeuring HW, Stek ML, Huisman M et al. A six-year prospective study of the prognosis and predictors in patients with late-life depression. Am J Geriatr Psychiatry 2018;26:985-97.
162. Jørgensen TS, Wium-Andersen IK, Wium-Andersen MK et al. Incidence of depression after stroke, and associated risk factors and mortality outcomes, in a large cohort of Danish patients. JAMA Psychiatry 2016;73:1032-40.
163. Joukamaa M, Heliövaara M, Knekt P et al. Mental disorders and cause-specific mortality. Br J Psychiatry 2001;179:498-502.
164. Jubran A, Lawm G, Kelly J et al. Depressive disorders during weaning from prolonged mechanical ventilation. Intensive Care Med 2010;36:828-35.
165. Kalesan B, Galea S. The relation of depression to in-hospital outcomes among adults hospitalized for firearm-related injury. J Affect Disord 2015;183:166-72.
166. Kao LT, Liu SP, Lin HC et al. Poor clinical outcomes among pneumonia patients with depressive disorder. PLoS One 2014;9:e116436.
167. Kaster TS, Vigod SN, Gomes T et al. Risk of serious medical events in patients with depression treated with electroconvulsive therapy: a propensity score-matched, retrospective cohort study. Lancet Psychiatry 2021;8:686-95.
168. Kawamura T, Shioiri T, Takahashi K et al. Survival rate and causes of mortality in the elderly with depression: a 15-year prospective study of a Japanese community sample, the Matsunoyama-Niigata suicide prevention project. J Investig Med 2007;55:106-14.
169. Kilbourne AM, Morden NE, Austin K et al. Excess heart-disease-related mortality in a national study of patients with mental disorders: identifying modifiable risk factors. Gen Hosp Psychiatry 2009;31:555-63.
170. Kim W, Jang SY, Lee TH et al. Association between continuity of care and subsequent hospitalization and mortality in patients with mood disorders: results from the Korea National Health Insurance cohort. PLoS One 2018;13:e0207740.
171. Kim C, Duan L, Phan DQ et al. Frequency of utilization of beta blockers in patients with heart failure and depression and their effect on mortality. Am J Cardiol 2019;124:746-50.
172. Kim GE, Jo MW, Shin YW. Increased prevalence of depression in South Korea from 2002 to 2013. Sci Rep 2020;10:16979.
173. Kim H, Lee KN, Shin DW et al. Association of comorbid mental disorders with cardiovascular disease risk in patients with type 2 diabetes: a nationwide cohort study. Gen Hosp Psychiatry 2022;79:33-41.
174. Kim S, Kim G, Cho SH et al. Impact of mental disorders on the all-cause mortality and cardiovascular disease outcomes in adults with new-onset type 1 diabetes: a nationwide cohort study. Psychiatry Res 2024;342:116228.
175. Kinder LS, Bradley KA, Katon WJ et al. Depression, posttraumatic stress disorder, and mortality. Psychosom Med 2008;70:20-6.
176. Kingsbury M, Sucha E, Horton NJ et al. Lifetime experience of multiple common mental disorders and 19-year mortality: results from a Canadian population-based cohort. Epidemiol Psychiatr Sci 2019;29:e18.
177. Ko A, Kim K, Sik Son J et al. Association of pre-existing depression with all-cause, cancer-related, and noncancer-related mortality among 5-year cancer survivors: a population-based cohort study. Sci Rep 2019;9:18334.
178. Konings SRA, Mierau JO, Visser E et al. Life years lost for users of specialized mental healthcare. Acta Psychiatr Scand 2023;148:338-46.
179. Kostev K, Hagemann-Goebel M, Gessler N et al. Is there an association between depression, anxiety disorders and COVID-19 severity and mortality? A multicenter retrospective cohort study conducted in 50 hospitals in Germany. J Psychiatr Res 2023;157:192-6.
180. Kouzis A, Eaton WW, Leaf PJ. Psychopathology and mortality in the general population. Soc Psychiatry Psychiatr Epidemiol 1995;30:165-70.
181. Koyanagi A, Köhler-Forsberg O, Benros ME et al. Mortality in unipolar depression preceding and following chronic somatic diseases. Acta Psychiatr Scand 2018;138:500-8.
182. Kumar A, Sloane D, Aiken L et al. Hospital nursing factors associated with decreased odds of mortality in older adult Medicare surgical patients with depression. BMC Geriatr 2022;22:665.
183. Kuruppu S, Ghani M, Pritchard M et al. A prospective investigation of depression and adverse outcomes in patients undergoing vascular surgical interventions: a retrospective cohort study using a large mental health database in South London. Eur Psychiatry 2021;64:e13.
184. Laan W, Termorshuizen F, Smeets HM et al. A comorbid anxiety disorder does not result in an excess risk of death among patients with a depressive disorder. J Affect Disord 2011;135:284-91.
185. Lasserre AM, Marti-Soler H, Strippoli MP et al. Clinical and course characteristics of depression and all-cause mortality: a prospective population-based study. J Affect Disord 2016;189:17-24.
186. Laursen TM, Musliner KL, Benros ME et al. Mortality and life expectancy in persons with severe unipolar depression. J Affect Disord 2016;193:203-7.
187. Lawrence WR, Kuliszewski MG, Hosler AS et al. Association between preexisting mental illnesses and mortality among Medicaid-insured women diagnosed with breast cancer. Soc Sci Med 2021;270:113643.
188. Lee H, Myung W, Lee C et al. Clinical epidemiology of long-term suicide risk in a nationwide population-based cohort study in South Korea. J Psychiatr Res 2018;100:47-55.
189. Lee J, Cho Y, Oh J et al. Analysis of anxiety or depression and long-term mortality among survivors of out-of-hospital cardiac arrest. JAMA Netw Open 2023;6:e237809.
190. Lemogne C, Nabi H, Melchior M et al. Mortality associated with depression as compared with other severe mental disorders: a 20-year follow-up study of the GAZEL cohort. J Psychiatr Res 2013;47:851-7.
191. Leone M, Kuja-Halkola R, Leval A et al. Association of youth depression with subsequent somatic diseases and premature death. JAMA Psychiatry 2021;78:302-10.
192. Levola J, Laine R, Pitkänen T. In-patient psychiatric care and non-substance-related psychiatric diagnoses among individuals seeking treatment for alcohol and substance use disorders: associations with all-cause mortality and suicide. Br J Psychiatry 2022;221:386-93.
193. Lewis G, Werbeloff N, Hayes JF et al. Diagnosed depression and sociodemographic factors as predictors of mortality in patients with dementia. Br J Psychiatry 2018;213:471-6.
194. Li G, Fife D, Wang G et al. All-cause mortality in patients with treatment-resistant depression: a cohort study in the US population. Ann Gen Psychiatry 2019;18:23.
195. Li HJ, Kuo CC, Li YC et al. Depression may not be a risk factor for mortality in stroke patients with nonsurgical treatment: a retrospective case-controlled study. Medicine 2019;98:e15753.
196. Lindström C, Siersma V, Kriegbaum M et al. Time trends in mortality for people with severe mental illness in Denmark 2000-2018. Nord J Psychiatry 2025;79:79-85.
197. Lloyd S, Baraghoshi D, Tao R et al. Mental health disorders are more common in colorectal cancer survivors and associated with decreased overall survival. Am J Clin Oncol 2019;42:355-62.
198. Lo HKY, Chan JKN, Wong CSM et al. Excess mortality and life-years lost in people diagnosed with depression: a 20-year population-based cohort study of 126,573 depressed individuals followed for 1,139,073 persons-years. Eur Neuropsychopharmacol 2025;91:1-8.
199. Lozupone M, Castellana F, Sardone R et al. Late-onset depression but not early-onset depression may increase the risk of all-cause mortality in older age: 8-year follow-up of the Salus in Apulia Study. J Am Med Dir Assoc 2023;24:679-87.
200. Lundberg J, Cars T, Lööv SÅ et al. Association of treatment-resistant depression with patient outcomes and health care resource utilization in a population-wide study. JAMA Psychiatry 2023;80:167-75.
201. Lundin A, Lundberg I, Allebeck P et al. Psychiatric diagnosis in late adolescence and long-term risk of suicide and suicide attempt. Acta Psychiatr Scand 2011;124:454-61.
202. Lundin A, Modig K, Halldin J et al. Mental disorder and long-term risk of mortality: 41 years of follow-up of a population sample in Stockholm, Sweden. Epidemiol Psychiatr Sci 2016;25:384-92.
203. Madsen KB, Plana-Ripoll O, Musliner KL et al. Cause-specific life years lost in individuals with treatment-resistant depression: a Danish nationwide register-based cohort study. J Affect Disord 2021;280:250-7.
204. Markkula N, Härkänen T, Perälä J et al. Mortality in people with depressive, anxiety and alcohol use disorders in Finland. Br J Psychiatry 2012;200:143-9.

205. Marrie RA, Walld R, Bolton JM et al. Psychiatric comorbidity increases mortality in immune-mediated inflammatory diseases. Gen Hosp Psychiatry 2018; 53:65-72.
206. Martin RL, Cloninger CR, Guze SB et al. Mortality in a follow-up of 500 psychiatric outpatients. I. Total mortality. Arch Gen Psychiatry 1985;42:47-54.
207. Martin RL, Cloninger CR, Guze SB et al. Mortality in a follow-up of 500 psychiatric outpatients. II. Cause-specific mortality. Arch Gen Psychiatry 1985;42:58-66.
208. Martin-Subero M, Kroenke K, Diez-Quevedo C et al. Depression as measured by PHQ-9 versus clinical diagnosis as an independent predictor of long-term mortality in a prospective cohort of medical inpatients. Psychosom Med 2017;79:273-82.
209. McCusker J, Cole M, Ciampi A et al. Does depression in older medical inpatients predict mortality? J Gerontol A Biol Sci Med Sci 2006;61:975-81.
210. Meier SM, Mattheisen M, Mors O et al. Increased mortality among people with anxiety disorders: total population study. Br J Psychiatry 2016;209:216-21.
211. Meller I, Fichter MM, Schröppel H. Mortality risk in the octo- and nonagenerians: longitudinal results of an epidemiological follow-up community study. Eur Arch Psychiatry Clin Neurosci 1999;249:180-9.
212. Melo APS, Dippenaar IN, Johnson SC et al. All-cause and cause-specific mortality among people with severe mental illness in Brazil's public health system, 2000-15: a retrospective study. Lancet Psychiatry 2022;9:771-81.
213. Mogga S, Prince M, Alem A et al. Outcome of major depression in Ethiopia: population-based study. Br J Psychiatry 2006;189:241-6.
214. Molnar MZ, Streja E, Sumida K et al. Pre-ESRD depression and post-ESRD mortality in patients with advanced CKD transitioning to dialysis. Clin J Am Soc Nephrol 2017;12:1428-37.
215. Morriss R, Tyrer F, Zaccardi F et al. Safety of antidepressants in a primary care cohort of adults with obesity and depression. PLoS One 2021;16:e0245722.
216. Moustgaard H, Joutsenniemi K, Sihvo S et al. Alcohol-related deaths and social factors in depression mortality: a register-based follow-up of depressed inpatients and antidepressant users in Finland. J Affect Disord 2013;148:278-85.
217. Müller-Oerlinghausen B, Ahrens B, Grof E et al. The effect of long-term lithium treatment on the mortality of patients with manic-depressive and schizoaffective illness. Acta Psychiatr Scand 1992;86:218-22.
218. Newman SC, Bland RC. Suicide risk varies by subtype of affective disorder. Acta Psychiatr Scand 1991;83:420-6.
219. Nordenskjöld A, Güney P, Nordenskjöld AM. Major adverse cardiovascular events following electroconvulsive therapy in depression: a register-based nationwide Swedish cohort study with 1-year follow-up. J Affect Disord 2022;296: 298-304.
220. Novak M, Mucsi I, Rhee CM et al. Increased risk of incident chronic kidney disease, cardiovascular disease, and mortality in patients with diabetes with comorbid depression. Diabetes Care 2016;39:1940-7.
221. Nurutdinova D, Chrusciel T, Zeringue A et al. Mental health disorders and the risk of AIDS-defining illness and death in HIV-infected veterans. AIDS 2012;26:229-34.
222. Oduyale OK, Eltahir AA, Stem M et al. What does a diagnosis of depression mean for patients undergoing colorectal surgery? J Surg Res 2021;260:454-61.
223. Oh TK, Park HY, Song IA. Depression and long-term survival among South Korean sepsis survivors: a nationwide cohort study from 2011 to 2014. Crit Care Med 2021;49:1470-80.
224. Oh TK, Song IA, Park HY et al. Depression and mortality after craniotomy for brain tumor removal: a nationwide cohort study in South Korea. J Affect Disord 2021;295:291-7.
225. Oh TK, Park HY, Song IA. Association between depression and mortality in patients with pain conditions: a South Korean nationwide cohort study. Yonsei Med J 2023;64:481-8.
226. Olfson M, Wall M, Wang S et al. Short-term suicide risk after psychiatric hospital discharge. JAMA Psychiatry 2016;73:1119-26.
227. Olin B, Jayewardene AK, Bunker M et al. Mortality and suicide risk in treatment-resistant depression: an observational study of the long-term impact of intervention. PLoS One 2012;7:e48002.
228. Olusunmade M, Qadir T, Akyar S et al. Incremental hospital utilization and mortality associated with co-morbid depression in pediatric hospitalizations. J Affect Disord 2019;251:270-3.
229. Osby U, Brandt L, Correia N et al. Excess mortality in bipolar and unipolar disorder in Sweden. Arch Gen Psychiatry 2001;58:844-50.
230. Ouk M, Wu CY, Colby-Milley J et al. Depression and diabetes mellitus multimorbidity is associated with loss of independence and dementia poststroke. Stroke 2020;51:3531-40.
231. Paljärvi T, Tiihonen J, Lähteenvuo M et al. Mortality in psychotic depression: 18-year follow-up study. Br J Psychiatry 2023;222:37-43.
232. Paljärvi T, Tiihonen J, Lähteenvuo M et al. Psychotic depression and deaths due to suicide. J Affect Disord 2023;321:28-32.
233. Palladino R, Chataway J, Majeed A et al. Interface of multiple sclerosis, depression, vascular disease, and mortality: a population-based matched cohort study. Neurology 2021;97:e1322-33.
234. Pan YJ, Yeh LL. Associations between mortality and exposure to psychotropic medication: a population-based cohort study for depressive disorders. Aust N Z J Psychiatry 2023;57:1253-62.
235. Pan YJ, Yeh LL, Kuo KH. Psychotropic medications and mortality from cardiovascular disease and suicide for individuals with depression in Taiwan. Asian J Psychiatr 2024;98:104091.
236. Park S, Rim SJ, Jo M et al. Comorbidity of alcohol use and other psychiatric disorders and suicide mortality: data from the South Korean National Health Insurance cohort, 2002 to 2013. Alcohol Clin Exp Res 2019;43:842-9.
237. Park J, Park S, Kim YG et al. Pre-existing depression in patients with coronary artery disease undergoing percutaneous coronary intervention. Sci Rep 2021;11:8600.
238. Paro A, Hyer JM, Pawlik T. Association of depression with in-patient and post-discharge disposition and expenditures among Medicare beneficiaries undergoing resection for cancer. Ann Surg Oncol 2021;28:6525-34.
239. Patten SB, Williams JV, Lavorato D et al. Mortality associated with major depression in a Canadian community cohort. Can J Psychiatry 2011;56:658-66.
240. Patten SB, Williams JV, Bulloch AG. Major depressive episodes and mortality in the Canadian household population. J Affect Disord 2019;242:165-71.
241. Pedersen JK, Wang L, Risbo N et al. Mortality in patients with incident rheumatoid arthritis and depression: a Danish cohort study of 11071 patients and 55355 comparators. Rheumatology 2024;63:680-8.
242. Penninx BW, Geerlings SW, Deeg DJ et al. Minor and major depression and the risk of death in older persons. Arch Gen Psychiatry 1999;56:889-95.
243. Phillips AC, Batty GD, Gale CR et al. Generalized anxiety disorder, major depressive disorder, and their comorbidity as predictors of all-cause and cardiovascular mortality: the Vietnam experience study. Psychosom Med 2009;71:395-403.
244. Pitman A, McDonald K, Logeswaran Y et al. The role of depression and use of alcohol and other drugs after partner suicide in the association between suicide bereavement and suicide: cohort study in the Danish population. Psychol Med 2024;54:2273-82.
245. Prasad SM, Eggener SE, Lipsitz SR et al. Effect of depression on diagnosis, treatment, and mortality of men with clinically localized prostate cancer. J Clin Oncol 2014;32:2471-8.
246. Prigge R, Wild SH, Jackson CA. Depression, diabetes, comorbid depression and diabetes and risk of all-cause and cause-specific mortality: a prospective cohort study. Diabetologia 2022;65:1450-60.
247. Pulska T, Pahkala K, Laippala P et al. Six-year survival of depressed elderly Finns: a community study. Int J Geriatr Psychiatry 1997;12:942-50.
248. Qian J, Simoni-Wastila L, Langenberg P et al. Effects of depression diagnosis and antidepressant treatment on mortality in Medicare beneficiaries with chronic obstructive pulmonary disease. J Am Geriatr Soc 2013;61:754-61.
249. Qian J, Simoni-Wastila L, Rattinger GB et al. Associations of depression diagnosis and antidepressant treatment with mortality among young and disabled Medicare beneficiaries with COPD. Gen Hosp Psychiatry 2013;35:612-8.
250. Rajan S, McKee M, Rangarajan S et al. Association of symptoms of depression with cardiovascular disease and mortality in low-, middle-, and high-income countries. JAMA Psychiatry 2020;77:1052-63.
251. Ranger TA, Clift AK, Patone M et al. Preexisting neuropsychiatric conditions and associated risk of severe COVID-19 infection and other acute respiratory infections. JAMA Psychiatry 2023;80:57-65.
252. Reutfors J, Andersson TM, Brenner P et al. Mortality in treatment-resistant unipolar depression: a register-based cohort study in Sweden. J Affect Disord 2018; 238:674-9.
253. Ribeiro-Carvalho F, Gonçalves-Pinho M, Bergantim R et al. Trend of depression and its association with sociodemographic and clinical factors among multiple myeloma hospitalizations: a Portuguese nationwide study from 2000 to 2015. Psychooncology 2020;29:1587-94.
254. Richardson LK, Egede LE, Mueller M. Effect of race/ethnicity and persistent recognition of depression on mortality in elderly men with type 2 diabetes and depression. Diabetes Care 2008;31:880-1.
255. Ried LD, Jia H, Feng H et al. Selective serotonin reuptake inhibitor treatment and depression are associated with poststroke mortality. Ann Pharmacother 2011;45:888-97.
256. Rogal SS, Mankaney G, Udawatta V et al. Pre-transplant depression is associated with length of hospitalization, discharge disposition, and survival after liver transplantation. PLoS One 2016;11:e0165517.
257. Rubino A, Roskell N, Tennis P et al. Risk of suicide during treatment with ven-

lafaxine, citalopram, fluoxetine, and dothiepin: retrospective cohort study. BMJ 2007;334:242.
258. Ryan J, Carriere I, Ritchie K et al. Late-life depression and mortality: influence of gender and antidepressant use. Br J Psychiatry 2008;192:12-8.
259. Saint Onge JM, Krueger PM, Rogers RG. The relationship between major depression and nonsuicide mortality for U.S. adults: the importance of health behaviors. J Gerontol B Psychol Sci Soc Sci 2014;69:622-32.
260. Salinero-Fort MA, Gómez-Campelo P, Cárdenas-Valladolid J et al. Effect of depression on mortality in type 2 diabetes mellitus after 8 years of follow-up. The DIADEMA study. Diabetes Res Clin Pract 2021;176:108863.
261. Saz P, Launer LJ, Día JL et al. Mortality and mental disorders in a Spanish elderly population. Int J Geriatr Psychiatry 1999;14:1031-8.
262. Scherrer JF, Garfield LD, Lustman PJ et al. Antidepressant drug compliance: reduced risk of MI and mortality in depressed patients. Am J Med 2011;124:318-24.
263. Scherrer JF, Chrusciel T, Garfield LD et al. Treatment-resistant and insufficiently treated depression and all-cause mortality following myocardial infarction. Br J Psychiatry 2012;200:137-42.
264. Schneider B, Müller MJ, Philipp M. Mortality in affective disorders. J Affect Disord 2001;65:263-74.
265. Schoevers RA, Geerlings MI, Beekman AT et al. Association of depression and gender with mortality in old age. Results from the Amsterdam Study of the Elderly (AMSTEL). Br J Psychiatry 2000;177:336-42.
266. Schultebraucks K, Blekic W, Basaraba C et al. The impact of preexisting psychiatric disorders and antidepressant use on COVID-19 related outcomes: a multicenter study. Mol Psychiatry 2023;28:2462-8.
267. Seoud T, Syed A, Carleton N et al. Depression before and after a diagnosis of pancreatic cancer: results from a national, population-based study. Pancreas 2020;49:1117-22.
268. Shah AJ, Veledar E, Hong Y et al. Depression and history of attempted suicide as risk factors for heart disease mortality in young individuals. Arch Gen Psychiatry 2011;68:1135-42.
269. Shaheen AA, Kaplan GG, Almishri W et al. The impact of depression and antidepressant usage on primary biliary cholangitis clinical outcomes. PLoS One 2018;13:e0194839.
270. Shaheen AA, Kaplan GG, Sharkey KA et al. Impact of major depression and antidepressant use on alcoholic and non-alcoholic fatty liver disease: a population-based study. Liver Int 2021;41:2308-17.
271. Shaheen AA, Kaplan GG, Sharkey KA et al. Impact of depression and antidepressant use on clinical outcomes of hepatitis B and C: a population-based study. Hepatol Commun 2023;7:e0062.
272. Shim EJ, Lee JW, Cho J et al. Association of depression and anxiety disorder with the risk of mortality in breast cancer: a National Health Insurance Service study in Korea. Breast Cancer Res Treat 2020;179:491-8.
273. Shoval G, Balicer RD, Feldman B et al. Adherence to antidepressant medications is associated with reduced premature mortality in patients with cancer: a nationwide cohort study. Depress Anxiety 2019;36:921-9.
274. So-Armah K, Gupta SK, Kundu S et al. Depression and all-cause mortality risk in HIV-infected and HIV-uninfected US veterans: a cohort study. HIV Med 2019;20:317-29.
275. Song IA, Park HY, Oh TK. Effect of preoperative psychiatric morbidity on postoperative outcomes of lung cancer surgery: a nationwide cohort study in South Korea. J Psychosom Res 2022;161:111002.
276. Stenman M, Holzmann MJ, Sartipy U. Relation of major depression to survival after coronary artery bypass grafting. Am J Cardiol 2014;114:698-703.
277. Strömberg R, Backlund LG, Johansson SE et al. Mortality in depressed and non-depressed primary care Swedish patients: a 12-year follow-up cohort study. Fam Pract 2013;30:514-9.
278. Su JA, Chang CC, Wang HM et al. Antidepressant treatment and mortality risk in patients with dementia and depression: a nationwide population cohort study in Taiwan. Ther Adv Chronic Dis 2019;10:2040622319853719.
279. Su KP, Lu N, Tang CH et al. Comparisons of the risk of medication noncompliance and suicidal behavior among patients with depressive disorders using different monotherapy antidepressants in Taiwan: a nationwide population-based retrospective cohort study. J Affect Disord 2019;250:170-7.
280. Subramaniapillai M, Chen VC, McIntyre RS et al. Added burden of major depressive disorder on cardiovascular morbidity and mortality among patients with cardiovascular disease and the modifying effects of antidepressants: a national retrospective cohort study. J Affect Disord 2021;294:580-5.
281. Sullivan DR, Ganzini L, Duckart JP et al. Treatment receipt and outcomes among lung cancer patients with depression. Clin Oncol (R Coll Radiol) 2014;26:25-31.
282. Sumino K, O'Brian K, Bartle B et al. Coexisting chronic conditions associated with mortality and morbidity in adult patients with asthma. J Asthma 2014;51:306-14.
283. Sundbøll J, Schmidt M, Adelborg K et al. Impact of pre-admission depression on mortality following myocardial infarction. Br J Psychiatry 2017;210:356-61.
284. Suppli NP, Johansen C, Kessing LV et al. Survival after early-stage breast cancer of women previously treated for depression: a nationwide Danish cohort study. J Clin Oncol 2017;35:334-42.
285. Surtees PG, Wainwright NW, Luben RN et al. Depression and ischemic heart disease mortality: evidence from the EPIC-Norfolk United Kingdom prospective cohort study. Am J Psychiatry 2008;165:515-23.
286. Szpakowski N, Bennell MC, Qiu F et al. Clinical impact of subsequent depression in patients with a new diagnosis of stable angina: a population-based study. Circ Cardiovasc Qual Outcomes 2016;9:731-9.
287. Thomson W. Lifting the shroud on depression and premature mortality: a 49-year follow-up study. J Affect Disord 2011;130:60-5.
288. Thomson W. Rate of stroke death after depression: a 40-year longitudinal study extension of Chichester/Salisbury Catchment area study. J Stroke Cerebrovasc Dis 2014;23:1837-42.
289. Ting RZ, Lau ES, Ozaki R et al. High risk for cardiovascular disease in Chinese type 2 diabetic patients with major depression – a 7-year prospective analysis of the Hong Kong Diabetes Registry. J Affect Disord 2013;149:129-35.
290. Tsuang MT, Woolson RF. Mortality in patients with schizophrenia, mania, depression and surgical conditions. A comparison with general population mortality. Br J Psychiatry 1977;130:162-6.
291. Tsuang MT, Woolson RF. Excess mortality in schizophrenia and affective disorders. Do suicides and accidental deaths solely account for this excess? Arch Gen Psychiatry 1978;35:1181-5.
292. van den Berg KS, Wiersema C, Hegeman JM et al. Clinical characteristics of late-life depression predicting mortality. Aging Ment Health 2021;25:476-83.
293. Veien KT, Videbæk L, Schou M et al. High mortality among heart failure patients treated with antidepressants. Int J Cardiol 2011;146:64-7.
294. Vin-Raviv N, Akinyemiju TF, Galea S et al. Depression and anxiety disorders among hospitalized women with breast cancer. PLoS One 2015;10:e0129169.
295. Vogt T, Pope C, Mullooly J et al. Mental health status as a predictor of morbidity and mortality: a 15-year follow-up of members of a health maintenance organization. Am J Public Health 1994;84:227-31.
296. Walker J, Mulick A, Magill N et al. Major depression and survival in people with cancer. Psychosom Med 2021;83:410-6.
297. Wang SM, Park SS, Park SH et al. Pre-transplant depression decreased overall survival of patients receiving allogeneic hematopoietic stem cell transplantation: a nationwide cohort study. Sci Rep 2020;10:15265.
298. Wedegaertner F, Arnhold-Kerri S, Sittaro NA et al. Depression- and anxiety-related sick leave and the risk of permanent disability and mortality in the working population in Germany: a cohort study. BMC Public Health 2013;13:145.
299. Whitlock FA, Sisking M. Depression and cancer: a follow-up study. Psychol Med 1979;9:747-52.
300. Williams LS, Ghose SS, Swindle RW. Depression and other mental health diagnoses increase mortality risk after ischemic stroke. Am J Psychiatry 2004;161:1090-5.
301. Wolfe HL, Boyer TL, Henderson ER et al. The association of depression with all-cause and cardiovascular disease mortality risk among transgender and gender diverse and cisgender patients. J Affect Disord 2024;366:364-9.
302. Wu YH, Chen YH, Chang MH et al. Depression in Parkinson's disease: a case-control study. PLoS One 2018;13:e0192050.
303. Wu PH, Lin MY, Huang TH et al. Depression amongst patients commencing maintenance dialysis is associated with increased risk of death and severe infections: a nationwide cohort study. PLoS One 2019;14:e0218335.
304. Wu CS, Hsu LY, Wang SH. Association of depression and diabetes complications and mortality: a population-based cohort study. Epidemiol Psychiatr Sci 2020;29:e96.
305. Wu YT, Kralj C, Acosta D et al. The association between depression, anxiety, and mortality in older people across eight low- and middle-income countries: results from the 10/66 cohort study. Int J Geriatr Psychiatry 2020;35:29-36.
306. Xu H, Li S, Mehta HB et al. Excess deaths from COVID-19 among Medicare beneficiaries with psychiatric diagnoses: community versus nursing home. J Am Geriatr Soc 2023;71:167-77.
307. Yang W, Wang Z, Li X et al. Association of depression with mortality in nationwide twins: the mediating role of dementia. Eur Psychiatry 2022;65:e63.
308. Yin J, Ma T, Li J et al. Association of mood disorder with cardiometabolic multimorbidity trajectory and life expectancy, a prospective cohort study. J Affect Disord 2022;312:1-8.
309. Zhang C, Zafari Z, Slejko JF et al. Impact of undertreatment of depression on suicide risk among children and adolescents with major depressive disorder: a

microsimulation study. Am J Epidemiol 2023;192:929-38.
310. Zhang GQ, Canner JK, Prince EJ et al. History of depression is associated with worsened postoperative outcomes following colectomy. Colorectal Dis 2021;23:2559-66.
311. Zhu N, Xu H, Lagerberg T et al. Comparative safety of antidepressants in adults with CKD. Clin J Am Soc Nephrol 2024;19:178-88.
312. Zielke T, Korepta L, Wesolowski M et al. The association of comorbid depression with mortality and amputation risk in patients with chronic limb-threatening ischemia. J Vasc Surg 2024;79:96-101.e1.
313. Zivin K, Yosef M, Miller EM et al. Associations between depression and all-cause and cause-specific risk of death: a retrospective cohort study in the Veterans Health Administration. J Psychosom Res 2015;78:324-31.
314. Zuidersma M, Conradi HJ, van Melle JP et al. Self-reported depressive symptoms, diagnosed clinical depression and cardiac morbidity and mortality after myocardial infarction. Int J Cardiol 2013;167:2775-80.
315. Parker GB, Hilton TM, Walsh WF et al. Timing is everything: the onset of depression and acute coronary syndrome outcome. Biol Psychiatry 2008;64:660-6.
316. Momen NC, Østergaard SD, Heide-Jorgensen U et al. Associations between physical diseases and subsequent mental disorders: a longitudinal study in a population-based cohort. World Psychiatry 2024;23:421-31.
317. Dubovsky SL, Ghosh BM, Serotte JC et al. Psychotic depression: diagnosis, differential diagnosis, and treatment. Psychother Psychosom 2021;90:160-77.
318. Rothschild AJ, Winer J, Flint AJ et al. Missed diagnosis of psychotic depression at 4 academic medical centers. J Clin Psychiatry 2008;69:1293-6.
319. McIntyre RS, Alsuwaidan M, Baune BT et al. Treatment-resistant depression: definition, prevalence, detection, management, and investigational interventions. World Psychiatry 2023;22:394-412.
320. McIntyre RS, Soczynska JK, Konarski JZ et al. Should depressive syndromes be reclassified as "metabolic syndrome type II"? Ann Clin Psychiatry 2007;19:257-64.
321. Dragioti E, Radua J, Solmi M et al. Impact of mental disorders on clinical outcomes of physical diseases: an umbrella review assessing population attributable fraction and generalized impact fraction. World Psychiatry 2023;22:86-104.
322. Khapre M, Kant R, Sharma D et al. Antidepressant use and glycemic control in diabetic population: a meta-analysis. Indian J Endocrinol Metab 2020;24:295-300.
323. Köhler CA, Freitas TH, Maes M et al. Peripheral cytokine and chemokine alterations in depression: a meta-analysis of 82 studies. Acta Psychiatr Scand 2017;135:373-87.
324. Köhler CA, Freitas TH, Stubbs B et al. Peripheral alterations in cytokine and chemokine levels after antidepressant drug treatment for major depressive disorder: systematic review and meta-analysis. Mol Neurobiol 2018;55:4195-206.
325. Sealock JM, Chen G, Davis LK. Anti-inflammatory action of antidepressants: investigating the longitudinal effect of antidepressants on white blood cell count. Complex Psychiatry 2023;9:1-10.
326. Elzib H, Pawloski J, Ding Y et al. Antidepressant pharmacotherapy and post-stroke motor rehabilitation: a review of neurophysiologic mechanisms and clinical relevance. Brain Circ 2019;5:62-7.
327. Friedman RA. Antidepressants' black-box warning – 10 years later. N Engl J Med 2014;371:1666-8.
328. Kirkbride JB, Anglin DM, Colman J et al. The social determinants of mental health and disorder: evidence, prevention and recommendations. World Psychiatry 2024;23:58-90.

DOI:10.1002/wps.21354

A pragmatic randomized controlled trial of cognitive therapy for post-traumatic stress disorder in children and adolescents exposed to multiple traumatic stressors: the DECRYPT trial

Richard Meiser-Stedman[1], Leila Allen[1], Polly-Anna Ashford[2], Ella Beeson[3], Sarah Byford[4], Andrea Danese[5-7], Annie Farr[8], Jack Finn[1], Ben Goodall[9], Lauren Grainger[1], Matthew Hammond[2], Rebecca Harmston[8], Ayla Humphrey[10], Dorothy King[11], Katie Lofthouse[1], Gerwyn Mahoney-Davies[12], Sarah Miles[7,11], Jessica Moore[1], Nicola Morant[13], Sarah Robertson[1], Lee Shepstone[2], Erika Sims[2], Paul Stallard[14], Annie Swanepoel[9], David Trickey[15], Katie Trigg[1], Ramesh Vishwakarma[2], Jon Wilson[16], Tim Dalgleish[17], Patrick Smith[7,10]

[1]Department of Clinical Psychology and Psychological Therapies, Norwich Medical School, University of East Anglia, Norwich, UK; [2]Norwich Clinical Trials Unit, Norwich Medical School, University of East Anglia, Norwich, UK; [3]Hertfordshire Partnership University NHS Foundation Trust, Hatfield, UK; [4]King's Health Economics, King's College London, London, UK; [5]Department of Child and Adolescent Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; [6]Social, Genetic and Developmental Psychiatry Centre, King's College London, London, UK; [7]National and Specialist CAMHS Clinic for Trauma, Anxiety and Depression, South London and Maudsley NHS Foundation Trust, London, UK; [8]Public and Patient Involvement representative; [9]North East London NHS Foundation Trust, Rainham, UK; [10]Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, UK; [11]Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; [12]Cardiff and Vale University Health Board, Cardiff, UK; [13]Division of Psychiatry, Faculty of Brain Sciences, University College London, London, UK; [14]Department for Health, University of Bath, Bath, UK; [15]Department of Clinical, Educational and Health Psychology, University College London, London, UK; [16]Norfolk and Suffolk NHS Foundation Trust, Norwich, UK; [17]MRC Cognition and Brain Sciences Unit, Cambridge, UK

Trauma-focused cognitive-behavioral therapies (TF-CBTs) are efficacious in children and adolescents with post-traumatic stress disorder (PTSD). However, there is limited evidence in youth exposed to multiple traumas, especially in real-world settings. This paper reports on a pragmatic randomized controlled trial (RCT) evaluating whether one form of TF-CBT, cognitive therapy for PTSD (CT-PTSD), was effective for PTSD following multiple trauma exposure in 8-17 year-olds attending UK mental health services, relative to treatment-as-usual (TAU). Youth with PTSD (N=120) following multiple traumas were randomly allocated to receive CT-PTSD or TAU. At baseline, complex PTSD diagnosis was common (55.0% of cases), and a large proportion of youth had comorbid mental disorders. The primary outcome was the score on the Child Revised Impact of Event Scale, 8-item version (CRIES-8) at post-treatment. Secondary outcomes included the CRIES-8 score at 11 months post-randomization, and several measures of PTSD, anxiety, depression, suicidal ideation, affect regulation, irritability, and general functioning at post-treatment and 11 months post-randomization. CT-PTSD was not found to be significantly superior to TAU on the CRIES-8 at post-treatment (adjusted difference: –3.80, 95% CI: –7.56 to –0.06, p=0.095; Hedges' g=–0.37, 95% CI: –0.78 to 0.03), but it was superior to TAU when patients who had received TF-CBT were excluded from that arm (adjusted difference: –4.60, 95% CI: –8.36 to –0.81, p=0.047; g=–0.46, 95% CI: –0.89 to –0.04). CT-PTSD was also superior to TAU on the CRIES-8 at 11 months (adjusted difference: –5.38, 95% CI: –8.88 to –1.87, p=0.003; g=–0.46, 95% CI: –0.90 to –0.02), and in a mixed-effect model incorporating all time points (p=0.007). Evidence of superiority for CT-PTSD was observed on parent-reported emotional difficulties at post-treatment and 11 months; and on child-reported total anxiety and depression, total anxiety, panic and separation anxiety, and parent-reported affect dysregulation and irritability at 11 months. Treatment withdrawal rate was low. Despite high baseline levels of comorbidity and impairment not seen in previous trials, CT-PTSD was not associated with significant deterioration or adverse events. This pragmatic trial is likely to contribute to the optimization of psychological intervention in youth with PTSD following multiple traumas, accompanied by severe comorbid mental disorders, in routine settings.

Key words: Post-traumatic stress disorder (PTSD), children and adolescents, multiple traumas, trauma-focused cognitive-behavioral therapies, cognitive therapy for PTSD (CT-PTSD), pragmatic trial

(World Psychiatry 2025;24:422–434)

Post-traumatic stress disorder (PTSD) is a deeply distressing and disabling psychiatric condition. In youth, it is usually comorbid with other psychiatric conditions[1], and may persist for years or even decades if untreated[2,3]. A recent epidemiological study suggested that over 7% of UK youth will have developed PTSD at some point by the age of 18 years[1]. Some systematic reviews attest to the efficacy of psychological therapies for the treatment of PTSD in children and adolescents[4,5], particularly trauma-focused cognitive-behavioral therapies (TF-CBTs), which are endorsed by treatment guidelines[6] as a first-line treatment for PTSD.

While this evidence is promising, several important issues remain. First, few trials to date have been pragmatic, which reduces the relevance of their findings to routine settings. In particular, few trials have compared an experimental treatment with treatment-as-usual (TAU) delivered by a mental health service. Several early trials used supportive counselling as a control condition, but network meta-analyses suggest that this treatment is no more efficacious than waiting list[4]. Although many trials have targeted youth identified in child protection/social service settings or schools, few have focused on youth referred to mental health services. Moreover, many trials have used highly trained research therapists rather than frontline clinicians. Finally, whether youth with complex PTSD[7] and severe comorbid psychiatric conditions stemming from multiple trauma exposure benefit from psychological therapies for PTSD is poorly understood.

In this context, it is important to consider the effectiveness of cognitive therapy for PTSD (CT-PTSD), a particular form of TF-CBT. Pronounced increases in treatment efficacy for adults with PTSD have been achieved through careful individual formulation and enhanced use of theoretically derived techniques in CT-PTSD[8,9]. This therapy may translate well to "frontline" clinical settings, as it employs a formulation-based approach (i.e., clinicians are able to tailor session content to a patient's needs), which may be particularly helpful when treating people with more complex

histories and more comorbidity.

The present randomized controlled trial (RCT) was designed to be pragmatic and focused on the most severe yet common PTSD presentations in child and adolescent mental health services (CAMHS) and on youth who had experienced multiple traumas. The primary objective of the study was to evaluate whether CT-PTSD is an effective treatment for PTSD in youth aged 8-17 years who have been exposed to multiple traumatic stressors, relative to TAU, in UK National Health Service (NHS) CAMHS.

Our primary hypothesis was that CT-PTSD would be superior to TAU on a routine outcome measure of PTSD severity. Our secondary hypothesis was that CT-PTSD would be superior to TAU with respect to other measures of PTSD, complex PTSD, anxiety, depression, general functioning, and parent-rated mental health in youth with a diagnosis of PTSD. We further assessed potential adverse events and harms.

METHODS

Trial design

The DECRYPT (Delivery of Cognitive Therapy for Young People after Trauma) trial employed a multicentre, pragmatic, single-blind, superiority study design[10]. All procedures complied with the ethical standards of the relevant national and institutional committees on human experimentation, as well as with the Helsinki Declaration of 1975, as revised in 2013. All procedures involving human subjects/patients were approved by the East of England - Cambridge South Research Ethics Committee (16/EE/0233).

Participants

Youth aged 8-17 years with a diagnosis of PTSD following exposure to multiple traumatic stressors were recruited from CAMHS clinics and youth mental health services across six mental health trusts in England and Wales (Cambridgeshire, Cardiff, Hertfordshire, North East London, Norfolk and Suffolk, and South London). Youth were eligible for inclusion if they: a) met the criteria for a diagnosis of PTSD according to the DSM-5[11], as ascertained by the Child PTSD Symptom Scale for DSM-5, interviewer version (CPSS-I-5)[12]; b) scored 17 or more on the Child Revised Impact of Event Scale, 8-item version (CRIES-8)[13]; and c) had been exposed to multiple traumatic stressors, assessed through the interview administration of the Child and Adolescent Trauma Screen (CATS)[14].

Exclusion criteria were: change of prescribed psychiatric medication in the previous two months; pervasive developmental disorder or neurodevelopmental disorder (except attention deficit hyperactivity disorder, ADHD); intellectual disability; another primary psychiatric diagnosis or clinical need that warranted treatment ahead of PTSD; inability to speak English; ongoing exposure to threat or safeguarding issues; strong likelihood of being unable to complete treatment (e.g., foster placement move); or history of organic brain damage.

Youth aged 16 and older gave their own written informed consent. Caregivers/parents gave written informed consent for participants aged under 16, while the children themselves gave their assent.

Interventions

CT-PTSD

CT-PTSD is a structured, fully manualized[15] psychological treatment delivered in an individual format for children and adolescents. In this trial, the suggested number of sessions was up to 15, lasting 60-90 min each. While previous trials (conducted in youth exposed to single traumas) suggested that up to 10 sessions be offered, we increased this number to allow for the extra complexities associated with PTSD following multiple traumas.

CT-PTSD includes several core elements: psychoeducation, with an emphasis on the role of cognitive processes in the onset and maintenance of PTSD; timelines, narrative work and imaginal reliving to help develop a coherent trauma narrative; cognitive restructuring (to reframe the meanings and interpretations associated with trauma and its aftermath), and coping management (e.g., addressing maladaptive strategies such as thought suppression, rumination, and safety-seeking behaviors). Up to three sessions were allowed for addressing other comorbid conditions and difficulties (e.g., depression or self-harm).

CT-PTSD was delivered by NHS CAMHS/youth mental health service therapists with an appropriate professional qualification, who had completed a CT-PTSD training by a member of the trial team (see also supplementary information). Therapists were discouraged from changing medication in this arm, but such changes were not prohibited or considered a breach of trial protocol. Therapists were asked to inform the trial team of any medication changes.

Treatment-as-usual (TAU)

TAU involved any active treatment selected by the mental health professionals in charge of the patients randomized to this arm. Since TF-CBTs are a recommended treatment for PTSD in the UK, therapists in this arm were not prevented from delivering these interventions. Changes to medication were not discouraged in this arm.

Treatment adherence

Supervision for CT-PTSD therapists was provided by a trial team clinical psychologist. Therapists delivering TAU received supervision according to their usual practice. To assess treatment fidelity and quality in the CT-PTSD arm, participants were asked to consent to therapy sessions being recorded, although this was not mandated. Clinical psychologists with extensive experience of delivering CT-PTSD rated these recordings, using a scale adapted from a trial

of CT-PTSD in adults[9] (see also supplementary information).

Therapists in each arm were asked to provide information at post-treatment about their own professional experience and training, the nature of the treatment they had provided, and the use of specific therapy techniques. For a planned sensitivity analysis, when psychoeducation, cognitive restructuring and some form of trauma memory work (e.g., narrative work, imaginal reliving) or narrative exposure therapy were endorsed, a participant randomized to TAU was deemed to have received TF-CBT. Eye movement desensitization and reprocessing (EMDR) was not regarded as a form of TF-CBT.

Outcomes

The primary outcome was self-reported PTSD symptoms at post-treatment (i.e., 5 to 6 months post-randomization), as measured with the CRIES-8. The CRIES-8, a validated self-report questionnaire[13], is the routine outcome monitoring tool for PTSD in children and adolescents endorsed by the UK Children and Young People's Improving Access to Psychological Therapies programme, and is recommended by the International Consortium for Health Outcomes Measurement[16]. The CRIES-8 was also completed at baseline, as well as at 2.5 months (mid-treatment) and 11 months post-randomization (secondary outcomes).

Further secondary outcomes (assessed at mid-treatment, post-treatment, and 11 months post-randomization) were: PTSD diagnosis and symptoms, using the CPSS-I-5[12] (with additional items assessing complex PTSD symptoms); DSM-5 PTSD symptom self-reported severity, using the CATS questionnaire[14] (with additional items assessing dissociation symptoms); disturbances in self-organization symptoms of complex PTSD, using the 12-item Child Complex PTSD Checklist questionnaire[17]; trauma-related misappraisals, using the Children's Post-Traumatic Cognitions Inventory[18]; anxiety and depression, using the Revised Child Anxiety and Depression Scale (RCADS)[19]; suicidal ideation, using four items from the Mood and Feelings Questionnaire[20]; and affect regulation and irritability, using the Affective Reactivity Index (ARI) (child and parent/caregiver report)[21]. Clinician-rated functioning was assessed using the Children's Global Assessment Scale[22]. Parent/caregiver-rated mental health and well-being were indexed by the Strengths and Difficulties Questionnaire (SDQ)[23], and emotional instability by the McLean Screening Instrument for Borderline Personality Disorder, caregiver version (MSI-BPD-C)[24].

If the onset of treatment was delayed after randomization, the timings of mid-treatment, post-treatment and 11-month assessments were also postponed, up to an additional three months.

Sample size

Our power calculation was based on a meta-analysis[25] which considered trials of TF-CBT in youth with PTSD, many of which used active control treatments, and obtained a pooled effect size of 0.67. In order to have 90% power to detect a between-group effect size of 0.67 (two-tailed t-test, 0.05 significance level), a combined sample size of 96 (48 participants per group) was required. In order to account for dropout (estimated at 20%), 120 participants were recruited.

Randomization and blinding

Trial data collection, randomization, blinding and data analysis were overseen by the Norwich Clinical Trials Unit. Patients were randomly assigned (1:1), via a web-based randomization service, to receive either CT-PTSD or TAU. Randomization was performed by the trial coordinator, with stratification by baseline CRIES-8 score (17-28 vs. 29-40) and site (i.e., recruiting NHS Trust). Allocation was by preset lists of permuted blocks with randomly distributed block sizes.

Participants and clinicians were aware of group allocation. Trained assessors who collected post-treatment and follow-up interview data were blinded to group allocation. Following allocation, all participants in the study and their clinical team were asked not to reveal the group to which they were randomized to the assessor.

At post-treatment, blind assessors administering the structured interviews were asked to guess their interviewee's randomization status. There was no relationship between these guesses and the actual arm to which a participant had been randomized (X^2=1.41, p=0.50).

Statistical analysis

The primary analysis was conducted on an intention-to-treat basis, i.e. all participants were followed up, and their data were analyzed according to group allocation rather than intervention received or adherence. The statistical analysis plan was approved by the Trial Steering Committee and the Trial Management Group. A general linear mixed effects model (assuming that the CRIES-8 has a normal distribution) was used for the primary efficacy analysis. This model included the stratification factor of site as a random factor. The CRIES-8 baseline score (also used for stratification of randomization) was included as a covariate, being a probable prognostic variable. Treatment arm was added as a fixed effect. The effect of therapist was not included as a random effect in the model, since very few therapists had more than one participant. Statistical significance was set at 0.05 (two-sided).

The analyses of secondary outcomes (including CRIES-8 score at 11 months) followed an analogous approach. In each case, an appropriate linear mixed effects model with inclusion of the stratification factors (i.e., site), measure at baseline (if available) and treatment arm was constructed. Between-group differences at each time point were analyzed (i.e., time by arm interaction) as well as overall treatment effect. This approach was also followed for the CRIES-8 (i.e., including mid-treatment, post-treatment and 11-month data). Diagnostic status (i.e., the presence of DSM-5 PTSD, and ICD-11 PTSD and complex PTSD) was analyzed using a logistic regression model with the stratification factors, measure at base-

line (if available) and treatment arm as explanatory variables.

In addition to reporting effect sizes, we also calculated several indices of clinical improvement for the PTSD outcomes, i.e. reliable improvement or deterioration[26], 50% improvement, and being below clinical thresholds where known. These were based on observed data only (see supplementary information for details). A separate committee, including three independent researchers, monitored adverse events and severe adverse events (see supplementary information for definitions).

Analyses were conducted using Stata (version 12.0). All models used full information maximum likelihood estimation to handle missing data. The trial was registered with an International Standard Randomized Controlled Trial Number (https://doi.org/10.1186/ISRCTN12077707).

RESULTS

Participant flow and recruitment

Of 304 youth identified by services as being potential participants, 35 (11.5%) were not eligible for trial entry. In 149 (49.0%) cases, the young persons or their families did not consider the trial further, their team did not advance the referral, or trial entry was not possible. Between February 7, 2017 and July 21, 2021, 58 participants were randomized to CT-PTSD and 62 to TAU (see Figure 1).

Three participants were not offered treatment sessions by their clinical team. Treatment data were missing for two participants. Six patients withdrew from the trial by post-treatment, and a further two withdrew by 11 months. Our recruitment and follow-up target of N=96 was met for the CRIES-8 at post-treatment.

Demographics and baseline data

Demographic characteristics of the participants, their trauma history, and aspects of their baseline psychopathology are reported in Table 1 (see also supplementary information). The sample mainly comprised adolescents (mean age: 14.9 years), and females (72.5%). Participants had typically been seen by the mental health service for several months (mean: 10.0) prior to trial entry. Ten participants (8.3%) were in a foster care placement.

The mean number of trauma types reported by youth was 4.7 (SD=2.2), while the median number of traumatic events reported was 14 (interquartile range: 4 to 45). The most common forms of trauma were physical abuse or attacks (within or outside the family), witnessing physical abuse or attacks, and sudden bereavement. Some form (either direct exposure or witnessing) of intrafamilial abuse was very common (75.0% of cases). Sexual abuse or violence was reported by 40.8% of participants.

The ICD-11 requirements for complex PTSD were fulfilled by 55.0% of the participants. A large proportion had clinically significant depression (74.2%), anxiety (50.8%) or emotional instability (48.0%); and a significant minority reported hearing voices (24.2%) (see Table 1). A minority of participants (26.7%) were taking medication for their mental health problems.

Therapist and treatment characteristics

Thirty-nine therapists delivered CT-PTSD and 52 therapists delivered TAU. Therapists were from a wide range of professional backgrounds. CT-PTSD arm participants (relative to TAU) were more likely to receive treatment from CBT therapists, while TAU arm participants (relative to CT-PTSD) were more likely to receive treatment from child psychotherapists. CT-PTSD arm therapists were more likely to have a cognitive-behavioral orientation, while TAU arm therapists were more likely to have a family/systemic or "other" approach (see Table 2 and supplementary information).

No differences were apparent between arms in terms of total sessions, but the distribution of sessions was different: CT-PTSD participants were more likely to have more than 8 sessions, while TAU participants were more likely to have more than 18 sessions (see Table 2 and supplementary information).

Therapists in the CT-PTSD arm received more supervision than those in the TAU arm. There were no differences between the trial arms with respect to participant-rated treatment credibility and therapeutic alliance at mid- or post-treatment. CT-PTSD was rated as quite credible (mean = 31.5 at post-treatment on a 4-40 scale) (see Table 2). The mean overall rating of treatment fidelity and quality in recorded sessions was 4.0 (i.e., "good") (see also supplementary information).

CT-PTSD arm therapists were more likely to report using behavioral activation, cognitive restructuring, trauma discussion, and trauma narrative work. TAU therapists were more likely to report using EMDR, mindfulness exercises, parent work, psychodynamic work, relaxation exercises, and supportive work (see Table 3 and supplementary information). One TAU case was mistakenly assigned a therapist who had been trained in CT-PTSD, contrary to our protocol, but the therapist reported delivering EMDR rather than CT-PTSD.

Assessments were conducted at an average of 199 (SD=64) days post-randomization for post-treatment, and 389 (SD=76) days post-randomization for 11 months. There were no between-arm differences.

Medication changes were reported for five participants in the CT-PTSD arm (four started an antidepressant; one switched from an antidepressant to a mood stabilizer), and two participants in the TAU arm (one started an antidepressant, and another an anti-ADHD medication).

Primary outcome

Our primary intention-to-treat analysis considered CRIES-8 score at post-treatment, finding a non-significant adjusted difference of –3.80 (95% CI: –7.56 to –0.06, p=0.095; Hedges' g=–0.37, 95% CI: –0.78 to 0.03) between the two arms (see Table 4).

Two sensitivity analyses were undertaken. In seven cases (all in the CT-PTSD arm), post-treatment CRIES-8 scores were taken

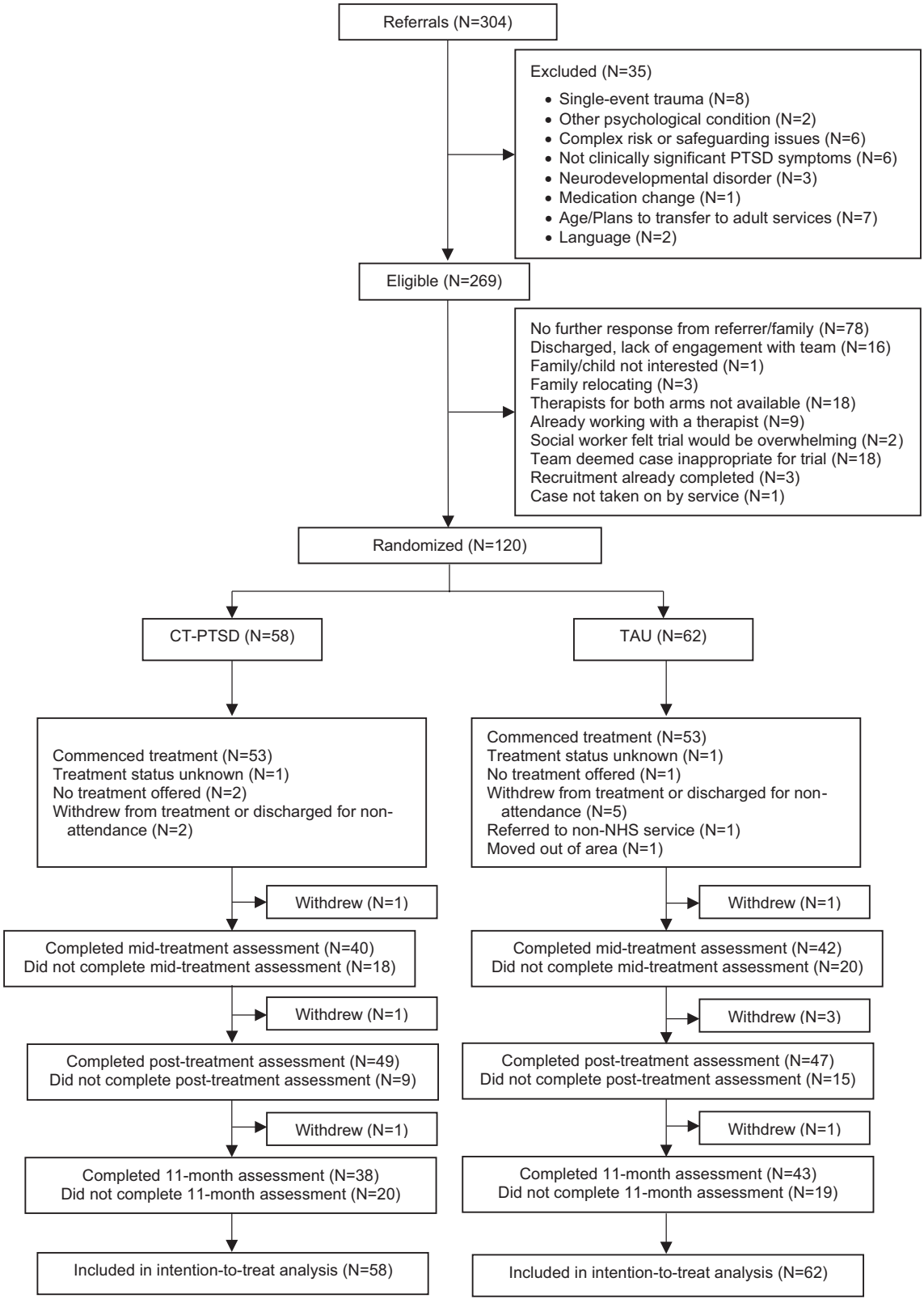

Figure 1 Study flow diagram. PTSD – post-traumatic stress disorder, CT-PTSD – cognitive therapy for PTSD, TAU – treatment as usual, NHS – National Health Service.

Table 1 Participant baseline characteristics

	CT-PTSD (N=58)	TAU (N=62)	Total (N=120)
Age (years), mean±SD	15.4±2.0	14.5±2.8	14.9±2.5
Female sex, N (%)	47 (81.0)	40 (64.5)	87 (72.5)
Race or ethnicity, N (%)			
Asian	0	3 (4.8)	3 (2.5)
Black	4 (6.9)	5 (8.1)	9 (7.5)
Mixed	7 (12.1)	7 (11.3)	14 (11.7)
White British	45 (77.6)	46 (74.2)	91 (75.8)
White other	2 (3.4)	1 (1.6)	3 (2.5)
Family status, N (%)			
Single	10 (21.7)	11 (21.6)	21 (21.6)
Relationship, not cohabiting	1 (2.2)	4 (7.8)	5 (5.2)
Cohabiting	8 (17.4)	6 (11.8)	14 (14.4)
Married	16 (34.8)	17 (33.3)	33 (34.0)
Separated/divorced	9 (19.6)	13 (25.5)	22 (22.7)
Widow/widower	2 (4.3)	0	2 (2.1)
Trauma history, N (%)			
Serious natural disaster	2 (3.4)	1 (1.6)	3 (2.5)
Serious accident/injury	8 (13.8)	26 (41.9)	34 (28.3)
Robbed by threat, force or weapon	5 (8.6)	5 (8.1)	10 (8.3)
Physical abuse/attack, family	34 (58.6)	23 (37.1)	57 (47.5)
Physical abuse/attack, non-family	24 (41.4)	33 (53.2)	57 (47.5)
Witness physical abuse/attack, family	33 (56.9)	33 (53.2)	66 (55.0)
Witness physical abuse/attack, non-family	25 (43.1)	28 (45.9)	53 (44.5)
Someone older touching private parts	20 (35.1)	16 (25.8)	36 (30.3)
Someone forcing or pressuring sex	17 (29.3)	13 (21.7)	30 (25.4)
Sudden/violent death of someone close	26 (44.8)	29 (46.8)	55 (45.8)
Attacked, stabbed, shot at, or hurt badly	2 (3.4)	11 (17.7)	13 (10.8)
Witness attack/stabbing/shooting/killing	13 (22.4)	22 (35.5)	35 (29.2)
Medical procedure	16 (27.6)	13 (21.3)	29 (24.4)
Other	38 (65.5)	45 (72.6)	83 (69.2)
Any sexual abuse, N (%)	27 (46.6)	22 (35.5)	49 (40.8)
Any intrafamilial trauma/abuse, N (%)	42 (72.4)	48 (77.4)	90 (75.0)
Trauma types, mean±SD	4.5±2.2	4.8±2.1	4.7±2.2
Trauma total number, median (IQR)	14 (4 to 42)	17 (6 to 46)	14 (4 to 45)
Use of psychotropic medication, N (%)			
Anti-ADHD	2 (3.4)	2 (3.2)	4 (3.3)
Antipsychotic	2 (3.4)	2 (3.2)	4 (3.3)
Antidepressant	15 (25.9)	11 (17.7)	26 (21.7)
Benzodiazepine	1 (1.7)	0	1 (0.8)
Any	19 (31.0)	13 (21.0)	32 (26.7)
Time under care before trial entry (months), mean±SD	9.9±7.6	10.1±8.2	10.0±7.9

Table 1 Participant baseline characteristics *(continued)*

	CT-PTSD (N=58)	TAU (N=62)	Total (N=120)
Psychopathology			
Complex PTSD, N (%)	38 (65.5)	28 (45.2)	66 (55.0)
Elevated emotional instability (MSI-BPD-C ≥7), N (%)	22 (46.8)	25 (49.0)	47 (48.0)
Hears voices (not re-experiencing), N (%)	14 (24.1)	15 (24.2)	29 (24.2)
RCADS clinical range (T-score ≥70), N (%)			
Total	36 (62.1)	39 (62.9)	75 (62.5)
Depression	43 (74.1)	46 (74.2)	89 (74.2)
Anxiety total	32 (55.2)	29 (46.8)	61 (50.8)
Generalized anxiety disorder	20 (34.5)	20 (32.3)	40 (33.3)
Obsessive-compulsive disorder	17 (29.3)	16 (25.8)	33 (27.5)
Panic disorder	38 (65.5)	31 (50.0)	69 (57.5)
Separation anxiety disorder	35 (60.3)	30 (48.4)	65 (54.2)
Social anxiety disorder	17 (29.3)	18 (29.0)	35 (29.2)

PTSD – post-traumatic stress disorder, CT-PTSD – cognitive therapy for PTSD, TAU – treatment as usual, ADHD – attention deficit hyperactivity disorder, MSI-BPD-C – McLean Screening Instrument for Borderline Personality Disorder, caregiver version, IQR – interquartile range, RCADS – Revised Child Anxiety and Depression Scale

from clinical notes (as the trial team questionnaire battery was not completed); when these data were excluded (in an analysis recommended by the Trial Steering Committee), the difference between the two arms was also non-significant (p=0.360). When excluding cases from the TAU arm that were confirmed to have involved TF-CBT (N=9), CT-PTSD was superior to TAU (adjusted difference: –4.60, 95% CI: –8.36 to –0.81, p=0.047; g=–0.46, 95% CI: –0.89 to –0.04).

Secondary outcomes

Linear mixed effect models that incorporated data at each time-point (i.e., mid-treatment, post-treatment and 11-month follow-up assessment) were derived for each secondary outcome. A significant treatment effect was found for the CRIES-8 score across time points (p=0.007), and for CRIES-8 score at the 11-month follow-up considered alone (adjusted difference: –5.38, 95% CI: –8.88 to –1.87, p=0.003; g=–0.46, 95% CI: –0.90 to –0.02). Pre-post effect sizes for each arm were large, and maintained at 11-month follow-up (see Table 5 and supplementary information).

No significant effects were detected for the self-reported DSM-5 PTSD questionnaire (the CATS), the interview-based measure of DSM-5 PTSD severity (the CPSS-I-5), and self-reported disturbances in self-organization, dissociation, or negative trauma-related appraisals. Pre-post effect sizes were large for the CATS and the CPSS-I-5 in both trial arms and were maintained at 11-month follow-up. No between-arm differences were apparent for PTSD diagnoses (i.e., DSM-5 PTSD, ICD-11 PTSD, or ICD-11 complex PTSD) at post-treatment or 11 months (see Table 5).

For several subscales of the RCADS, CT-PTSD was superior to TAU at 11 months: total anxiety and depression score (adjusted difference: –9.37, 95% CI: –18.28 to –0.46, p=0.039); total anxiety (adjusted difference: –7.44; 95% CI: –14.55 to –0.33, p=0.041); panic (adjusted difference: –2.57, 95% CI: –4.69 to –0.46, p=0.017), and separation anxiety (adjusted difference: –1.85, 95% CI: –3.26 to –0.44, p=0.011).

CT-PTSD was superior to TAU on the SDQ subscale for parent-reported emotional difficulties at post-treatment (p<0.001) and 11 months (p=0.002); on MSI-BPD-C emotional instability at post-treatment (p=0.014); and on parent-reported affect dysregulation and irritability at 11 months (p=0.015). Significant treatment effects across time points were found for parent-reported affect dysregulation and irritability (p=0.013), emotional instability (p=0.022), SDQ emotional difficulties (p<0.001), and SDQ total score (p=0.043) (see Table 5).

These analyses were repeated excluding cases from the TAU arm that were confirmed to have involved TF-CBT. The same pattern of results was found for secondary outcomes, but with further between-arm differences: CT-PTSD was superior to TAU on parent-reported affect dysregulation, SDQ total difficulties, and overall functioning at post-treatment, and on emotional instability at 11 months (see supplementary information).

Clinical improvement

Reliable improvement on the CRIES-8 (decreased by 11.92 or more) was reported by 42.9% of CT-PTSD and 29.8% of TAU participants at post-treatment, and by 47.4% of CT-PTSD and 27.9% of TAU participants at 11-month follow-up. These differences were not statistically significant (see also supplementary information).

Reliable improvement on the CATS (decreased by 14.89 or more) was reported by 39.0% of CT-PTSD and 24.4% of TAU participants at post-treatment (non-significant difference), and by 56.8% of CT-

Table 2 Therapist and therapy characteristics

	CT-PTSD (N=58)	TAU (N=62)	p
Therapist profession, N (%)			<0.001
CBT therapist	20 (37.0)*	3 (5.1)	
Child psychotherapist	0	9 (15.3)*	
Clinical psychologist	20 (37.0)	21 (35.6)	
Counselling psychologist	3 (5.6)	0	
Family therapist	1 (1.9)	6 (10.2)	
Nurse	4 (7.4)	6 (10.2)	
Social worker	1 (1.9)	3 (5.1)	
Other	5 (9.2)	11 (18.6)	
Therapist orientation, N (%)			<0.001
Cognitive-behavioral	38 (79.2)*	12 (27.3)	
EMDR	1 (2.1)	3 (6.8)	
Family/systemic	2 (4.2)	12 (27.3)*	
Psychodynamic	0	3 (6.8)	
Other	7 (14.6)	14 (31.8)*	
Number of therapy sessions, mean±SD	10.7±4.2	11.1±9.7	0.774
8+ sessions, N (%)	43 (75.4)	35 (56.5)	0.029
12+ sessions, N (%)	26 (45.6)	24 (28.7)	0.446
18+ sessions, N (%)	2 (3.5)	13 (21.0)	0.004
Supervision (minutes), mean±SD	159±103	90±109	0.005
Patients' therapy ratings			
Credibility, mid; mean±SD	30.0±8.36	28.4±8.79	0.421
Credibility, post; mean±SD	31.5±9.37	29.1±9.90	0.264
Therapeutic alliance, mid; mean±SD	36.9±7.86	37.0±7.20	0.941
Therapeutic alliance, post; mean±SD	37.6±7.94	38.0±6.97	0.853

Asterisks indicate a significant post-hoc difference between the two groups. CT-PTSD – cognitive therapy for PTSD, TAU – treatment as usual, CBT – cognitive-behavioral therapy, EMDR – eye movement desensitization and reprocessing, mid – mid-treatment, post – post-treatment.

PTSD and 30.2% of TAU participants at 11-month follow-up (p= 0.017). An improvement of at least 50% on the CATS was reported by 29.3% of CT-PTSD and 17.8% of TAU participants at post-treatment (non-significant difference), and by 37.8% of CT-PTSD and 14.0% of TAU participants at 11-month follow-up (p=0.014) (see also supplementary information).

Deterioration and adverse events

Our protocol pre-specified that we would report the frequency with which participants experienced a deterioration of 7 or more from baseline on the CRIES-8. At mid-treatment, three CT-PTSD participants (7.5%) and one TAU participant (2.4%) had experienced this degree of deterioration. At post-treatment, no CT-PTSD participants and two TAU participants (4.3%) had experienced it. This was also the case at the 11-month follow-up assessment (see also supplementary information).

Four adverse events were reported in the CT-PTSD arm, that were classified, respectively, as "decline in mental state", "any untoward increase in extent of self-harm or suicidal ideation", "overdose of medication without signs or symptoms", and "other" (the participant was thought to be beginning to experience a psychotic episode). Three adverse events were reported in the TAU arm, classified as "decline in mental state" (two cases) and "any untoward

Table 3 Treatment components reported by therapists (%)

	CT-PTSD (N=58)	TAU (N=62)	p
Behavioral activation	67.6	38.1	0.032
Art therapy	0	5.3	0.413
Cognitive restructuring	91.7	54.5	0.002
Trauma discussion	100	70.8	<0.001
EMDR	0	35.0	0.001
Medical review	14.3	11.8	1.00
Mindfulness exercises	28.1	60.9	0.015
Parent work	36.4	70.0	0.018
Play therapy	3.6	22.2	0.069
Problem solving	46.7	61.9	0.283
Psychoeducation	100.0	100.0	-
Psychodynamic work	0.0	20.0	0.037
Relaxation exercises	40.6	76.2	0.011
Reliving	70.6	55.6	0.278
Supportive work	74.3	100.0	0.005
Trauma narrative work	88.9	41.2	<0.001

CT-PTSD – cognitive therapy for PTSD, TAU – treatment as usual, EMDR – eye movement desensitization and reprocessing

Table 4 Adjusted mean differences between CT-PTSD and TAU for the primary and sensitivity analyses (CRIES-8) at post-treatment

	Adjusted difference (95% CI)	p	Hedges' g (95% CI)
All available data	–3.80 (–7.56 to –0.06)	0.095	–0.37 (–0.78 to 0.03)
Excluding scores taken from clinical notes	–2.20 (–6.05 to 1.74)	0.360	–0.20 (–0.62 to 0.22)
Removing TAU cases treated with TF-CBT	–4.60 (–8.36 to –0.81)	0.047	–0.46 (–0.89 to –0.04)

CT-PTSD – cognitive therapy for PTSD, TAU – treatment as usual, TF-CBT – trauma-focused cognitive behavioral therapy, CRIES-8 – Child Revised Impact of Event Scale, 8-item version

Table 5 Comparisons between CT-PTSD and TAU on the primary and secondary outcome measures (intention-to-treat analysis)

	Adjusted difference (95% CI)	Treatment by time point (p value)	Treatment effect across all time points (p value)	Standardized between-group effect size (95% CI)	Standardized within-group effect size (95% CI)	
					CT-PTSD	TAU
PTSD severity (CRIES-8)			0.007			
Mid-treatment	−3.434 (−6.923 to 0.055)	0.054		−0.352 (−0.789 to 0.084)	1.264 (0.751 to 1.776)	0.636 (0.178 to 1.094)
Post-treatment	−3.099 (−6.402 to 0.205)*	0.066		−0.343 (−0.747 to 0.060)	2.137 (1.661 to 2.613)	1.437 (1.012 to 1.861)
11 months	−5.377 (−8.883 to -1.871)	0.003		−0.459 (−0.901 to −0.017)	2.155 (1.533 to 2.777)	1.225 (0.720 to 1.730)
PTSD severity (CATS)			0.360			
Mid-treatment	−1.197 (−5.549 to 3.156)	0.589		−0.153 (−0.587 to 0.281)	0.618 (0.191 to 1.044)	0.330 (−0.077 to 0.738)
Post-treatment	0.493 (−3.748 to 4.735)	0.819		−0.026 (−0.447 to 0.395)	1.148 (0.717 to 1.579)	1.039 (0.633 to 1.444)
11 months	−4.230 (−8.628 to 0.169)	0.059		−0.270 (−0.712 to 0.171)	1.498 (0.989 to 2.008)	1.017 (0.561 to 1.472)
PTSD severity, assessor (CPSS-I-5)			0.366			
Post-treatment	−1.092 (−6.376 to 4.192)	0.684		0.081 (−0.359 to 0.521)	1.208 (0.761 to 1.655)	0.977 (0.564 to 1.391)
11 months	−4.869 (−10.761 to 1.022)	0.105		−0.071 (−0.566 to 0.424)	1.541 (1.032 to 2.051)	1.127 (0.652 to 1.601)
PTSD diagnosis, DSM-5 (CPSS-I-5)			-			
Post-treatment	OR=0.6 (0.22-1.84)	0.405				
11 months	OR=1.7 (0.59-4.89)	0.324				
PTSD diagnosis, ICD-11 (CPSS-I-5)			0.110			
Post-treatment	OR=1.4 (0.47-4.42)	0.479				
11 months	OR=1.1 (0.54-2.10)	0.835				
Complex PTSD diagnosis (CPSS-I-5)			0.522			
Post-treatment	OR=0.9 (0.45-1.60)	0.568				
11 months	OR=0.4 (0.10-1.80)	0.210				
Disturbances in self-organization (CCPC)			0.616			
Mid-treatment	0.309 (−2.590 to 3.208)	0.834		−0.050 (−0.483 to 0.383)	0.266 (−0.226 to 0.758)	0.120 (−0.363 to 0.602)
Post-treatment	−0.638 (−3.489 to 2.214)	0.660		−0.032 (−0.456 to 0.392)	0.771 (0.354 to 1.188)	0.645 (0.254 to 1.036)
11 months	−1.469 (−4.384 to 1.446)	0.322		−0.099 (−0.536 to 0.338)	0.956 (0.390 to 1.522)	0.740 (0.199 to 1.281)
Dissociation			0.992			
Mid-treatment	0.068 (−0.811 to 0.947)	0.879		−0.073 (−0.506 to 0.360)	−0.032 (−1.05 to 0.990)	−0.093 (−1.09 to 0.905)
Post-treatment	0.023 (−0.817 to 0.863)	0.958		−0.030 (−0.451 to 0.391)	0.465 (0.060 to 0.870)	0.451 (0.065 to 0.837)
11 months	−0.064 (−0.946 to 0.819)	0.888		0.013 (−0.423 to 0.450)	0.370 (−0.807 to 1.547)	0.402 (−0.781 to 1.585)
Trauma-related appraisals (CPTCI)			0.949			
Mid-treatment	0.451 (−5.469 to 6.370)	0.881		−0.011 (−0.449 to 0.427)	0.102 (−0.248 to 0.452)	0.101 (−0.239 to 0.441)
Post-treatment	−0.752 (−6.705 to 5.201)	0.804		−0.064 (−0.497 to 0.370)	0.606 (0.188 to 1.024)	0.541 (0.145 to 0.937)

Table 5 Comparisons between CT-PTSD and TAU on the primary and secondary outcome measures (intention-to-treat analysis) *(continued)*

	Adjusted difference (95% CI)	Treatment by time point (p value)	Treatment effect across all time points (p value)	Standardized between-group effect size (95% CI)	Standardized within-group effect size (95% CI)	
					CT-PTSD	TAU
Total anxiety and depression (RCADS)			0.402			
Post-treatment	1.597 (−7.126 to 1.319)	0.719		−0.034 (−0.460 to 0.392)	0.788 (0.371 to 1.206)	0.669 (0.275 to 1.063)
11 months	−9.369 (-18.282 to -0.456)	0.039		−0.243 (−0.681 to 0.195)	0.923 (0.543 to 1.303)	0.515 (0.166 to 0.865)
Depression (RCADS)			0.656			
Post-treatment	0.972 (−1.188 to 3.133)	0.376		0.050 (−0.376 to 0.476)	0.548 (0.138 to 0.958)	0.656 (0.262 to 1.049)
11 months	−1.860 (−4.071 to 0.352)	0.099		−0.275 (−0.714 to 0.163)	0.801 (0.208 to 1.394)	0.496 (−0.053 to 1.044)
Anxiety total (RCADS)			0.371			
Post-treatment	0.605 (−6.333 to 7.544)	0.864		−0.059 (−0.485 to 0.367)	0.797 (0.379 to 1.215)	0.617 (0.225 to 1.010)
11 months	-7.440 (−14.555 to -0.326)	0.041		−0.223 (−0.661 to 0.214)	0.884 (0.478 to 1.290)	0.479 (0.104 to 0.853)
Generalized anxiety (RCADS)			0.314			
Post-treatment	−0.170 (−1.608 to 1.269)	0.817		−0.182 (−0.609 to 0.245)	0.726 (0.310 to 1.141)	0.534 (0.143 to 0.924)
11 months	−1.325 (−2.795 to 0.146)	0.077		−0.260 (−0.699 to 0.178)	0.873 (0.135 to 1.610)	0.564 (−0.116 to 1.245)
Obsessive-compulsive (RCADS)			0.453			
Post-treatment	−0.032 (−1.311 to 1.247)	0.961		−0.223 (−0.651 to 0.204)	0.701 (0.286 to 1.115)	0.591 (0.199 to 0.982)
11 months	−0.816 (−2.131 to 0.499)	0.223		−0.275 (−0.713 to 0.164)	0.688 (−0.189 to 1.565)	0.505 (−0.286 to 1.296)
Panic (RCADS)			0.529			
Post-treatment	1.066 (−0.985 to 3.116)	0.307		0.159 (−0.267 to 0.586)	0.524 (0.115 to 0.934)	0.456 (0.067 to 0.844)
11 months	−2.575 (−4.690 to -0.461)	0.017		−0.140 (−0.576 to 0.297)	0.638 (−0.023 to 1.300)	0.248 (−0.376 to 0.872)
Separation anxiety (RCADS)			0.094			
Post-treatment	−0.466 (−1.834 to 0.902)	0.503		−0.074 (−0.500 to 0.352)	0.626 (0.213 to 1.038)	0.346 (−0.040 to 0.733)
11 months	−1.847 (−3.259 to -0.436)	0.011		−0.236 (−0.674 to 0.202)	0.650 (−0.160 to 1.459)	0.164 (−0.572 to 0.900)
Social anxiety (RCADS)			0.884			
Post-treatment	0.207 (−1.971 to 2.386)	0.852		−0.049 (−0.475 to 0.377)	0.581 (0.170 to 0.992)	0.487 (0.098 to 0.877)
11 months	−0.789 (−3.026 to 1.449)	0.488		−0.115 (−0.551 to 0.322)	0.633 (0.046 to 1.220)	0.456 (−0.109 to 1.021)
Affect regulation and irritability (ARI), child report			0.773			
Post-treatment	0.916 (−0.300 to 2.133)	0.139		−0.009 (−0.440 to 0.421)	0.184 (−0.219 to 0.588)	0.363 (−0.030 to 0.755)
11 months	−0.326 (−1.572 to 0.920)	0.607		−0.247 (−0.685 to 0.191)	0.445 (−0.424 to 1.313)	0.368 (−0.429 to 1.165)
Affect regulation and irritability (ARI), parent report			0.013			
Post-treatment	−1.315 (−2.716 to 0.086)	0.066		−0.398 (−0.861 to 0.066)	0.514 (0.066 to 0.962)	0.210 (−0.216 to 0.636)
11 months	−1.897 (−3.428 to -0.366)	0.015		−0.374 (−0.869 to 0.120)	0.669 (−0.293 to 1.631)	0.382 (−0.442 to 1.207)

Table 5 Comparisons between CT-PTSD and TAU on the primary and secondary outcome measures (intention-to-treat analysis) *(continued)*

	Adjusted difference (95% CI)	Treatment by time point (p value)	Treatment effect across all time points (p value)	Standardized between-group effect size (95% CI)	Standardized within-group effect size (95% CI)	
					CT-PTSD	TAU
Suicidality (MFQ-4)			0.334			
Post-treatment	0.392 (−0.314 to 1.098)	0.274		0.117 (−0.314 to 0.548)	0.440 (0.033 to 0.848)	0.433 (0.040 to 0.827)
Emotional instability (MSI-BPD-C)			0.022			
Post-treatment	−1.106 (−1.983 to -0.230)	0.014		−0.252 (−0.728 to 0.224)	0.880 (0.406 to 1.354)	0.426 (0.002 to 0.851)
11 months	−0.871 (−1.832 to 0.091)	0.076		−0.045 (−0.570 to 0.480)	1.076 (0.016 to 2.136)	0.850 (−0.184 to 1.884)
Functioning (CGAS)			0.845			
Post-treatment	3.863 (−1.638 to 9.365)	0.168		0.210 (−0.227 to 0.648)	−1.275 (−1.721 to −0.827)	−0.663 (−1.062 to −0.264)
11 months	−1.603 (−7.680 to 4.475)	0.604		−0.169 (−0.652 to 0.315)	−1.358 (−1.755 to −0.962)	−1.354 (−1.738 to −0.969)
Parent-reported emotional difficulties (SDQ)			<0.001			
Post-treatment	−1.486 (−2.290 to -0.683)	<0.001		−0.214 (−0.674 to 0.246)	0.892 (0.430 to 1.353)	0.128 (−0.298 to 0.553)
11 months	−1.338 (−2.197 to -0.478)	0.002		−0.129 (−0.619 to 0.362)	1.045 (0.132 to 1.959)	0.382 (−0.479 to 1.244)
Parent-reported total difficulties (SDQ)			0.043			
Post-treatment	−1.404 (−2.976 to 0.167)	0.080		−0.168 (−0.627 to 0.292)	0.647 (0.195 to 1.099)	0.305 (−0.122 to 0.733)
11 months	−1.687 (−3.389 to 0.016)	0.052		−0.166 (−0.657 to 0.325)	0.781 (0.224 to 1.337)	0.418 (−0.107 to 0.943)

*Values are different from those given in Table 4 as these analyses include data from all four time points (i.e., not just baseline and post-treatment). PTSD – post-traumatic stress disorder, CT-PTSD – cognitive therapy for PTSD, TAU – treatment as usual, ARI – Affective Reactivity Index, CATS – Child and Adolescent Trauma Screen questionnaire, CCPC – Child Complex PTSD Checklist, CGAS – Children's Global Assessment Scale, CPSS-I-5 – Child PTSD Symptom Scale for DSM-5, interviewer version, CPTCI – Children's Post-Traumatic Cognitions Inventory, CRIES-8 – Child Revised Impact of Event Scale, 8-item version, MFQ – Mood and Feelings Questionnaire, MSI-BPD-C – McLean Screening Instrument for Borderline Personality Disorder, caregiver version, RCADS – Revised Child Anxiety and Depression Scale, SDQ – Strengths and Difficulties Questionnaire.

increase in extent of self-harm or suicidal ideation" (one case). One severe adverse event was reported in each arm, classified respectively as "overdose of medication without signs or symptoms" in the CT-PTSD arm, and "any untoward increase in extent of self-harm or suicidal ideation" in the TAU arm.

DISCUSSION

In this pragmatic trial of CT-PTSD versus TAU for children and adolescents with PTSD following multiple trauma exposure, in whom significant comorbid mental health conditions at baseline were common, CT-PTSD was not significantly superior to TAU on the CRIES-8 at post-treatment, but showed a significant superiority at 11-month follow-up (p=0.003). CT-PTSD was significantly superior to TAU on the CRIES-8 at post-treatment when patients who had received TF-CBT were excluded from that arm (p=0.047), and in a mixed-effect model incorporating all time points (p=0.007).

CT-PTSD was superior to TAU at post-treatment with respect to parent-reported emotional instability (p=0.014) and emotional difficulties (p<0.001). At the 11-month assessment, CT-PTSD demonstrated superiority with respect to total anxiety and depression (p=0.039), total anxiety (p=0.041), panic symptoms (p=0.017) and separation anxiety symptoms (p=0.011), as well as to parent-reported emotional difficulties (p=0.002) and affect dysregulation and irritability (p=0.015). No differences were apparent for other outcomes, including PTSD diagnoses.

Several features of the DECRYPT trial underline the robust nature of the evaluation undertaken, and particularly the relevance of the trial's findings to routine clinical practice.

First, it is noteworthy how severe participants' mental health problems were at baseline. All participants met full DSM-5 PTSD criteria, unlike many earlier trials which included youth with subsyndromal PTSD. Moreover, a considerable proportion fulfilled ICD-11 requirements for a complex PTSD diagnosis (55.0%), and a large proportion had clinically significant depression (74.2%), anxiety (50.8%), or emotional instability (48.0%). A significant minority (24.2%) also reported hearing voices. The degree of impairment (measured by the Children's Global Assessment Scale) and in particular the degree of negative trauma-related appraisals (measured by the Children's Post-Traumatic Cognitions Inventory) were much more severe than in comparable European trials[27,28]. Future TF-CBT treatment protocols should consider how to address the degree of impairment and comorbidity identified here.

Second, the therapists in this trial were highly diverse, with a wide range of professional backgrounds and therapeutic orientations represented. The CT-PTSD arm comprised more therapists with a cognitive-behavioral orientation than the TAU arm, but the amount of specific training that they received during their participation in the present trial was minimal. While CT-PTSD therapists received a greater amount of supervision, TAU therapists were free to use whatever treatments they had already been using in clinical practice, including TF-CBT. Treatment fidelity and quality were deemed to be satisfactory in the recorded sessions, although there is certainly scope for improvement.

The number of completed sessions in the treatment phase was fewer than expected, and this was not attributable to treatment dropout. Services were free to continue to offer psychological treatment after the post-treatment assessment. It may be that there was an inadequate dose of therapy delivered in the main treatment phase for CT-PTSD to have its full effect, and the extension of therapy beyond the post-treatment assessment led to between-group differences at the 11-month assessment.

The trial's pattern of findings – with some features of the post-traumatic stress response (e.g., trauma-related appraisals and disturbances in self-organization) only showing a modest shift in the treatment phase – suggests that clinicians may need to consider extending the number of sessions offered to these complex cases to address issues relating to self-concept and identity, and possibly using multiple assessment measures to track progress over time.

It is noteworthy that, despite the severity of the needs of the recruited youth, there were very few adverse events, and little evidence of deterioration in PTSD severity over the course of treatment. Moreover, suicidal ideation showed some improvement in each group. This is important, given clinician concerns around the delivery of TF-CBT to youth with PTSD, including lack of confidence and fears around "retraumatization"[29]. Furthermore, the large pre-post effect sizes for each group with respect to PTSD severity and overall functioning (that compare favorably to earlier trials) emphasize the value of offering treatment to children and adolescents with PTSD following multiple trauma exposure, despite the complexity of their needs. An embedded qualitative study[30] found that youth who received CT-PTSD derived self-defined benefits, including feeling more able to talk about trauma and improved abilities to cope.

This trial has some limitations. In attempting to understand the full range of difficulties that youth experienced, we may have increased participant burden excessively, generating a relatively high number of missing data. The trial arms were unbalanced with respect to some treatment characteristics (e.g., more CBT practitioners and supervision in the CT-PTSD condition). While we attempted to assess treatment fidelity, only a limited number of therapy recordings were available.

It might be argued that TF-CBT should not have been allowed as a treatment option in the TAU arm. However, this has been the recommended first-line treatment for PTSD in children and adolescents since a UK treatment guideline recommendation in 2005, and so we deemed that formally excluding this option may have been unethical. Moreover, as a pragmatic trial, the TAU arm represented the range of interventions available in routine mental health settings.

In conclusion, although we did not find that CT-PTSD was superior to TAU on our primary outcome at post-treatment, we did show that CT-PTSD was superior at the 11-month follow-up, and in a mixed-effect model incorporating all time points. Both CT-PTSD and TAU were acceptable to youth and not associated with significant deterioration or adverse events, despite the significant range of comorbid mental health problems experienced by these youth. This pragmatic trial is likely to add significantly to the optimization of psychological intervention in youth with PTSD following multiple traumas, accompanied by severe comorbid mental health problems, in routine settings.

ACKNOWLEDGEMENTS

The DECRYPT trial was funded by a UK National Institute for Health and Care Research (NIHR) Career Development Fellowship to R. Meiser-Stedman (CDF-2015-08-073). The views expressed here are those of the authors and not necessarily those of the UK National Health System, the NIHR, or the UK Department of Health and Social Care. The authors gratefully acknowledge the support of the Trial Steering Committee (S. Cartwright-Hatton, P. Fonagy, S. Reynolds, and the late W. Yule) and the Data Monitoring Committee (B. Dunn, I. Goodyer, V. Vickerstaff), as well as the contribution of the late M.-S. Man. T. Dalgleish and P. Smith are joint senior authors of the paper. Supplementary information on this study is available at https://osf.io/jkbvr/.

REFERENCES

1. Lewis SJ, Arseneault L, Caspi A et al. The epidemiology of trauma and post-traumatic stress disorder in a representative cohort of young people in England and Wales. Lancet Psychiatry 2019;6:247-56.
2. Morgan L, Scourfield J, Williams D et al. The Aberfan disaster: 33-year follow-up of survivors. Br J Psychiatry 2003;182:532-6.
3. Yule W, Bolton D, Udwin O et al. The long-term psychological effects of a disaster experienced in adolescence: I: The incidence and course of PTSD. J Child Psychol Psychiatry 2000;41:503-11.
4. Mavranezouli I, Megnin-Viggars O, Daly C et al. Research Review: Psychological and psychosocial treatments for children and young people with post-traumatic stress disorder: a network meta-analysis. J Child Psychol Psychiatry 2020;61:18-29.
5. Hoppen TH, Meiser-Stedman R, Jensen TK et al. Efficacy of psychological interventions for post-traumatic stress disorder in children and adolescents exposed to single versus multiple traumas: meta-analysis of randomised controlled trials. Br J Psychiatry 2023;222:196-203.
6. National Institute for Health and Care Excellence. Guideline 116: Post-traumatic stress disorder. London: National Institute for Health and Care Excellence, 2018.
7. World Health Organization. International classification of diseases, 11th revision. Geneva: World Health Organization, 2022.
8. Ehlers A, Clark DM, Hackmann A et al. A randomized controlled trial of cognitive therapy, a self-help booklet, and repeated assessments as early interventions for posttraumatic stress disorder. Arch Gen Psychiatry 2003;60:1024-32.
9. Ehlers A, Hackmann A, Grey N et al. A randomized controlled trial of 7-day intensive and standard weekly cognitive therapy for PTSD and emotion-focused supportive therapy. Am J Psychiatry 2014;171:294-304.
10. Allen L, Ashford PA, Beeson E et al. DECRYPT trial: study protocol for a phase II randomised controlled trial of cognitive therapy for post-traumatic stress disorder (PTSD) in youth exposed to multiple traumatic stressors. BMJ Open 2021;11:e047600.
11. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 5th ed. Arlington: American Psychiatric Association, 2013.
12. Foa EB, McLean CP, Capaldi S et al. Prolonged exposure vs supportive counseling for sexual abuse-related PTSD in adolescent girls: a randomized clinical trial. JAMA 2013;310:2650-7.
13. Perrin S, Meiser-Stedman R, Smith P. The Children's Revised Impact of Event Scale (CRIES): validity as a screening instrument for PTSD. Behav Cogn Psychother 2005;33:487-98.
14. Sachser C, Berliner L, Holt T et al. International development and validation of the Child and Adolescent Trauma Screening Questionnaire (CATS). Dallas: International Society for Traumatic Stress Studies, 2016.
15. Smith P, Perrin S, Yule W et al. Post traumatic stress disorder: cognitive therapy with children and young people. London: Routledge, 2009.
16. Krause KR, Chung S, Adewuya AO et al. International consensus on a standard set of outcome measures for child and youth anxiety, depression, obsessive-compulsive disorder, and post-traumatic stress disorder. Lancet Psychiatry 2021;8:76-86.
17. Hiller RM, Meiser-Stedman R, Elliott E et al. A longitudinal study of cognitive predictors of (complex) post-traumatic stress in young people in out-of-home care. J Child Psychol Psychiatry 2021;62:48-57.
18. Meiser-Stedman R, Smith P, Bryant R et al. Development and validation of the Child Post-Traumatic Cognitions Inventory (CPTCI). J Child Psychol Psychiatry 2009;50:432-40.
19. Chorpita BF, Moffitt CE, Gray J. Psychometric properties of the Revised Child Anxiety and Depression Scale in a clinical sample. Behav Res Ther 2005;43:309-22.
20. Hammerton G, Zammit S, Potter R et al. Validation of a composite of suicide items from the Mood and Feelings Questionnaire (MFQ) in offspring of recurrently depressed parents. Psychiatry Res 2014;216:82-8.
21. Stringaris A, Goodman R, Ferdinando S et al. The Affective Reactivity Index: a concise irritability scale for clinical and research settings. J Child Psychol Psychiatry 2012;53:1109-17.
22. Shaffer D, Gould MS, Brasic J et al. A Children's Global Assessment Scale (CGAS). Arch Gen Psychiatry 1983;40:1228-31.
23. Goodman R. Psychometric properties of the Strengths and Difficulties Questionnaire. J Am Acad Child Adolesc Psychiatry 2001;40:1337-45.
24. Goodman M, Patil U, Triebwasser J et al. Parental burden associated with borderline personality disorder in female offspring. J Pers Disord 2011;25:59-74.
25. Cary CE, McMillen JC. The data behind the dissemination: a systematic review of trauma-focused cognitive behavioral therapy for use with children and youth. Child Youth Serv Rev 2012;34:748-57.
26. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991;59:12-9.
27. Goldbeck L, Muche R, Sachser C et al. Effectiveness of trauma-focused cognitive behavioral therapy for children and adolescents: a randomized controlled trial in eight German mental health clinics. Psychother Psychosom 2016;85:159-70.
28. Jensen TK, Holt T, Ormhaug SM et al. A randomized effectiveness study comparing trauma-focused cognitive behavioral therapy with therapy as usual for youth. J Clin Child Adolesc Psychol 2014;43:356-69.
29. Purnell L, Chiu K, Bhutani GE et al. Clinicians' perspectives on retraumatisation during trauma-focused interventions for post-traumatic stress disorder: a survey of UK mental health professionals. J Anxiety Disord 2024;106:102913.
30. Allen L, Fox A, Copello A et al. A qualitative study exploring children and young people's experiences of cognitive therapy for PTSD delivered within the context of a randomised controlled trial. Psychol Psychother 2025;98:74-89.

DOI:10.1002/wps.21355

Beyond treating mental disorders: the broad impact of acceptance and commitment training

Psychotherapists know things that the world needs to know. Although research on psychotherapy has for the last half a century largely targeted psychiatric syndromes, an extensive body of knowledge has also been accumulated on how change happens when psychotherapy is successful. It turns out that these "processes of change" are also helpful in virtually every other area of human functioning.

Psychotherapy works when people learn to be more cognitively and emotionally open, and develop a deeper or more spiritual sense of self, consciously focusing their attention on what is of importance to them in the present. They then need to use these "mindfulness" skills to create values-based behavioral habits that bring meaning and purpose into their lives, while also extending these skills to their relationships and taking care of their body[1].

Virtually every replicated mediator that empirically explains the impact of any form of psychotherapy on mental health outcomes can fit comfortably within the above two sentences. This set of processes is an extended version of the "psychological flexibility model"[2]. It was first championed and targeted by acceptance and commitment therapy or training ("ACT" in either case, said as a single word)[3], but we now know that other successful evidence-based forms of psychotherapy commonly work by altering elements of this same set[1].

ACT was never just a form of psychotherapy, which is why the word "training" soon became necessary to describe the use of its methods in contexts outside of the treatment of mental disorders, such as in prevention, physical and behavioral health, social wellness and justice, and education or performance areas. There are now over 1,100 randomized controlled trials on ACT, and a review of that vast literature shows that less than one third of the existing studies have focused on alleviating existing psychiatric syndromes[4].

The broad impact of ACT training outside of the domain of mental disorders suggests that other forms of evidence-based psychotherapy also might be repurposed to solve problems and to promote human prosperity in a much wider range of use cases.

Here we mention a handful of areas as examples of the impact of ACT training. In all of them it is known both that ACT training can be helpful and that the processes of change that explain its impact fit within the extended psychological flexibility model.

One very important area is the use of ACT to help patients step up to physical and behavioral health challenges, such as in chronic pain and terminal illness. Multiple meta-analyses in these areas suggest that ACT training can be helpful[5].

An impressive recent example is a study on motor neuron disease[6], a terminal illness with no known medical treatments, from which 300,000 people suffer and 40,000 persons a year die. When ACT training was used to increase psychological flexibility, the quality of life of patients improved substantially at 6 and 9 months post-randomization, compared to usual care alone. The authors stated that improving quality of life is "vital given the progressive nature of the condition" and, while previously "guidelines have not been able to recommend evidenced psychological interventions for this population", "this trial provides definitive evidence for one such intervention". Similar findings have been reported in a wide range of other physical diseases, as well as traumatic injuries[5].

Another set of high-quality studies on ACT training have addressed the sequelae of war. When war happens, almost all forms of human functioning are disrupted: relationships, sleep, self-care, eating, physical health, mental health, and so on. When the World Health Organization (WHO) identified the need for a method to address the broad range of human problems occasioned by war, it settled on ACT training as a possible option. An ACT-based guided self-help book in graphic novel or "cartoon book" form was created, called "Doing What Matters in Times of Stress". This is generally used in combination with audio tapes (available in 32 different languages), with the help of non-professionals who have received training in how to present the material in small groups (viewing the book, listening to the tapes, and discussing the material).

In a carefully done randomized trial published in this journal[7], it was found that, compared to an enhanced treatment as usual condition, this scalable intervention reduced by 47% the future development of mental health problems among war survivors who were not yet showing a diagnosable disorder.

As an indication of its perceived relevance to the public, the ACT-based book is now the most frequently downloaded document from the WHO website, and is being actively disseminated in Ukraine. Based on multiple studies, the WHO website itself says that "the guide is for anyone who experiences stress, wherever they live and whatever their circumstances". As an indication of that assessment, when the COVID-19 pandemic occurred and no validated psychological intervention was yet available, the WHO referred people to this guide.

The use of worksite-based training to decrease stress and to foster work effectiveness is another example with positive evidence[8]. Simple employee workshops can be used to train people in how to apply ACT methods to their own lives to enhance their psychological flexibility, reduce stress, increase job satisfaction, and improve objective measures of performance such as the ability to learn new tasks. This is particularly likely to occur if the worksite itself affords employees the ability to have control over how to accomplish work-related tasks. The formula is simple: flexible workers in a flexible workplace predict both employer and employee success.

A final example is addressing stigma. Across a wide range of specific forms, in groups who are the recipients of enacted stigma – ethnic and racial minorities; persons who are neurodiverse; lesbian, gay, bisexual, transgender, queer or questioning (LGBTQ+) popula-

tions; religious minorities; persons with high body weight; and so on – ACT training has been found to be both psychologically helpful and to empower steps needed to create social and environmental change and to promote social justice[9].

The breadth of application of ACT and the apparently near-universal relevance of the processes it targets raises the intriguing possibility that the core of what evidence-based psychotherapies of many forms target could be helpful beyond the treatment of mental disorders *per se*. The size of the database on successful studies of processes of change in ACT and ACT training[1], the expansion of ACT into lower and middle income countries[4], and the range of studies on applied areas outside of psychiatric syndromes[4] are large compared to most other forms of evidence-based psychotherapy, but the takeaway point may be relevant to these other forms as well.

Psychotherapists know things that the world needs to know. In the context of the ongoing and worldwide crises of climate change, immigration, economic disparities, rapid social change, religious conflicts, political division, and war, that is a message of cautious hope for the world.

Steven C. Hayes
University of Nevada, Reno, NV, USA

1. Hayes SC, Ciarrochi J, Hofmann SG et al. Behav Res Ther 2022;156:104155.
2. Ong CW, Ciarrochi J, Hofmann SG et al. J Context Behav Sci 2024;32:100734.
3. Hayes SC, Strosahl K, Wilson KG. Acceptance and commitment therapy: the process and practice of mindful change, 2nd ed. New York: Guilford, 2012.
4. Sahdra B, King G, Payne J et al. Behav Ther 2024;55:1348-63.
5. Association for Contextual Behavioral Science. Meta-analyses and systematic, scoping, or narrative reviews of the ACT evidence base. https://contextualscience.org.
6. Gould RL, McDermott CJ, Thompson BJ et al. Lancet 2024;403:2381-94.
7. Acarturk C, Uygun E, Ilkkursun Z et al. World Psychiatry 2022;21:88-95.
8. Flaxman PE, Bond FW, Livheim F. The mindful and effective employee: an acceptance and commitment therapy training manual for improving well-being and performance. Oakland: New Harbinger, 2013.
9. Krafft J, Ferrell J, Levin ME et al. J Context Behav Sci 2018;7:15-28.

DOI:10.1002/wps.21358

Has the time come to stop using control groups in trials of psychosocial interventions?

Control conditions are an essential element in evidence-based mental health. In randomized trials, these conditions are needed to make sure that the improvements found in people who receive an intervention are not caused by other factors, such as spontaneous recovery. But control conditions in mental health research trials have many problems.

Drug trials are relatively simple, because they almost always can make use of pill placebo control conditions. This allows to examine the exact contribution to the effects of the specific substance in the pill, compared to the same pill without the active substance, in double-blind designs. However, the outcomes of these trials can also be biased. For example, adverse effects of medications in psychiatry are common, and often reveal who receives the active drug in a randomized trial, thus breaking the masking. The outcomes may also be affected, for example, by sponsorship bias and the use of "placebo run-in" methods[1]. Furthermore, pill placebo cannot be used in trials examining psychosocial interventions.

One of the most used control conditions in these latter trials is waitlist, in which participants receive the intervention after the experimental group has finished[2,3]. One advantage is that in the end all participants receive the intervention. A disadvantage is, however, that waitlist controlled trials result in larger effects of an intervention than, for example, usual care control groups, and may overestimate the true effect of the intervention[4,5]. It is not clear why the effects are larger – maybe because of increased expectations, but that is not certain[2].

Usual care is another control condition that is often chosen[6]. Trials with such control groups have the advantage that they can show what an intervention can add to the care that is already available. The main disadvantage, however, is that usual care can vary considerably. It can be delivered in primary care, in specialized care, in general medical settings, or in other settings. The care delivered within each of these settings also varies considerably across countries and regions. For example, usual care delivered in a primary care setting can be a relatively extensive treatment in high-income countries, but in low- and middle-income countries it typically means no treatment at all. To solve this, trials sometimes use "enhanced usual care", in which the care is delivered as it should be in ideal circumstances. However, this does not solve the issue of the heterogeneity of usual care across settings and countries, making the results of trials difficult to generalize to other settings and communities than where the trial was conducted.

Another type of control condition is "no treatment", in which participants are not offered care at all. However, every person has the right to access treatment, whether or not he/she participates in a trial. This means that everyone always has access to "usual care". "No treatment" may, therefore, be better considered as a specific type of usual care. It suffers from the same heterogeneity as other usual care conditions, because it varies considerably across countries and communities.

"Psychological placebos" have in principle the possibility to mask participants for the condition they have been assigned to[2,3]. However, when a psychological placebo is delivered in such a way that it is a credible treatment, it can be quite effective in itself. For example, cognitive behavioural therapy for depression does not significantly differ from non-directive counseling, which is often used as a psychological placebo condition[7]. If, on the other hand, psychological placebo is not delivered as a credible treatment, then patients know that they are in the control group, the masking is broken, and expectancies will be reduced. In this case, pa-

tients in the control group may be disappointed and demoralized, and the effects that are found in the trial can be more an artefact than a true outcome of the intervention. "Psychological placebos" should, therefore, preferably be avoided in trials, or at least be used very cautiously.

So, there is not really a satisfactory control condition that can be used when examining the effects of psychosocial interventions. But, over the past decades, an increasing number of randomized controlled trials have examined the effects of many psychosocial interventions. For example, in the field of psychotherapies for depression, more than 400 randomized trials have been conducted with all kinds of control conditions. Overall, these therapies have comparable effects, can be administered in all kinds of formats and in many different target groups[5]. At some point new controlled trials add very little new knowledge to what is already available. In psychotherapies for depression that seems to be the case already, but also in other research areas there are several dozens and sometimes hundreds of trials available on psychosocial interventions, such as in the treatment of anxiety disorders, substance use disorders and psychotic disorders.

When we know that a treatment works in comparison to a control group, and that has been confirmed in multiple randomized trials, we could step away from controlled trials altogether and focus on other clinical questions that are also relevant. New treatments can be compared to established treatments in randomized trials without a control group to examine if they have comparable effects. Fractional factorial designs do not need a control condition either and allow to examine effective components of psychosocial interventions[8]. Stepped wedged designs can be used to examine the effects but also how to implement interventions in routine care. And we can focus on the collection of large, high-quality datasets in routine care, with as many relevant predictors as possible, in order to identify who benefits from which treatment and further develop personalized interventions. Using techniques such as propensity score matching[9], we can estimate the effects of an intervention by accounting for all covariates that may predict its outcome.

In summary, the most frequently used control conditions in trials of psychosocial interventions overestimate the effects (waitlist), are extremely heterogeneous by design and cannot be generalized to other settings or countries (usual care and no treatment), or cannot be delivered as they should (psychological placebos). In randomized controlled trials there is no good alternative for such control groups, and researchers should carefully consider the advantages and disadvantages of each type of control before deciding which one to use. At some point, however, we have enough controlled trials, and new ones do not add very much to what is already known. For several research areas, this moment is there or very close. This allows to step away from conventional randomized trials with control groups and to examine other questions that are relevant to improve outcomes for patients.

Pim Cuijpers
Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; International Institute for Psychotherapy, Babeş-Bolyai University, Cluj-Napoca, Romania

1. Munkholm K, Paludan-Müller AS, Boesen K. BMJ Open 2019;9:e024886.
2. Mohr DC, Spring B, Freedland KE et al. Psychother Psychosom 2009;78:275-84.
3. Freedland KE, King AC, Ambrosius WT et al. J Clin Epidemiol 2019;110:74-81.
4. Furukawa TA, Noma H, Caldwell DM et al. Acta Psychiatr Scand 2014;130:181-92.
5. Cuijpers P, Quero S, Noma H et al. World Psychiatry 2021;20:283-93.
6. Freedland KE, Mohr DC, Davidson KW et al. Psychosom Med 2011;73:323-35.
7. Cuijpers P, Driessen E, Hollon SD et al. Clin Psychol Rev 2012;32:280-91.
8. Collins LM, Murphy SA, Strecher V. Am J Prev Med 2007;32:S112-8.
9. Guo S, Fraser MW. Propensity score analysis. Statistical methods and applications. Thousand Oaks: Sage Publications, 2010.

DOI:10.1002/wps.21359

Bringing future thinking into focus in psychopathology

Mentally representing possible events in our personal future has been theorized as a central organizing feature of human psychology. Even our autobiographical memory can be thought as functioning to provide the building blocks for imagining our future. As Kierkegaard suggested, life must be lived forwards.

Future thinking occurs both voluntarily and involuntarily throughout our daily lives, and plays a fundamental role in mental health and adaptation. It allows us to make *if-then* inferences about consequences, pre-experience possibilities for internal and external conditions, and make decisions and plans about what to do or avoid. It helps us set meaningful goals and devise solutions to potential problems. Moreover, future thinking regulates our emotions and motivation, through mechanisms such as anticipatory rewards, anxiety about negative outcomes or uncertainty, and disgust for experiences that we wish to avoid. It also creates a sense of self-continuity as we imagine how we will change over time. Additionally, it fosters social bonds by enabling us to imagine, anticipate and discuss social plans or shared experiences with others.

However, future thinking can become dysfunctional within the context of psychopathology. While this is partly recognized through symptoms in diagnoses, such as hopelessness in depression, apathy in schizophrenia, or excessive worrying in generalized anxiety disorder, other changes in future thinking cut across various psychopathologies. While best understood within emotional and psychotic disorders, dysfunction is not limited to them.

A fundamental difficulty in generating positive future events and possibilities is observed in conditions such as depression, schizophrenia-spectrum disorders, suicidality, deliberate self-harm, posttraumatic stress, and borderline personality disorder. Hopelessness in depression and suicidality are linked to this impairment, but not reliably to the generation of more negative events[1]. This becomes particularly relevant as individuals struggle to imagine meaningful reasons for living.

In psychopathology, a general finding in future thinking is a ten-

dency to generate fewer specific, individualized future events, and more generalized, categorical types of events. There are also fewer vivid details and less experience of emotionally amplifying mental imagery for positive events, but more for intrusive and negativistic future thoughts[2]. This lack of sensory detail can inhibit the sense of pre-experiencing the future, diminishing the believability of potential rewards or joy, and making it harder to counter negative expectations. In clinical anxiety, future negative events tend to be experienced as particularly detailed and vivid, consistent with heightened alertness and a focus on aversive experiences or potential dangers[3].

Psychopathology is generally associated with future thinking saturated with negative content. Negative future thinking is more easily generated, and accompanied by increased worry about consequences, involuntary and intrusive thoughts, and more extreme catastrophizing. In internalizing conditions, there is less in-the-moment pleasure when anticipating future events, and predictions of stronger negative emotional reactions and consequences for future events prevail. This can diminish motivation for experiences that could otherwise bring emotional, physical, personal or social rewards. In schizophrenia-spectrum disorders, mental imagery is less vivid in its sensory qualities, which is associated with less anticipation of pleasure, less motivation, and less engagement in subsequent activities[4].

A "better safe than sorry" mechanism may help explain this dysfunction. Once negative biases about the self and the world are set up, they create a tendency to generate similarly negative thoughts about the future, along with a bias toward information that is confirmatory, threat-related, more generalized, and lower in sensory-perceptual detail. According to predictive processing theory, this is thought to maintain the predictability of the world and potentially mitigate disappointment. However, this comes at the expense of maintaining negative expectations and less nuanced or positive future thinking that could alter these predictions[5].

Thus, future thinking dysfunction plays a pivotal role in the "vicious cycle" of decreased interest, withdrawal and hopelessness, as seen in conditions such as depression. To illustrate, consider a therapeutic task such as activity scheduling. Challenges arise when individuals attempt to first generate specific possible future activities related to pleasure or mastery. When activities are selected, imagining their details becomes difficult, as does anticipating or experiencing positive emotions or satisfaction. Negative or lackluster outcomes are more likely to be predicted, leading to lower motivation, reduced engagement with plans, and a reinforcement of prior beliefs about a negative future or lack of self-efficacy. This underscores the role of future thinking in how people appraise both how their lives will be and how their lives are.

Intervening in how people project themselves mentally and emotionally into the future may help break this cycle and promote healthier ways of thinking, feeling and functioning. Standalone future thinking interventions tend to focus on bringing prospection under conscious control. This involves repeated practice in simulating specific future events with positive or adaptive outcomes, as well as the promotion of rich detail and mental imagery. Such interventions are typically brief, deliverable alongside other therapy, and generally have moderate-size effects on characteristics of future thinking in clinical samples[6]. Specific examples include Positive Mental Imagery Training, for which a series of controlled trials have shown reductions in anhedonia and depressive symptoms[7]; Future Event Specificity Training, for which controlled trials show changes in transdiagnostic factors, reduced anhedonia, and a higher likelihood of remission from major depression[8]; and Episodic Future Thinking, for which various substance-use related outcomes have been found across controlled trials[9].

Aside from standalone future thinking interventions, many therapeutic tasks in evidence-based treatments already implicitly or explicitly engage future thinking. Examples include generating therapeutic goals, problem-solving, behavioral experiments and exposure techniques, thought challenging, motivational interviewing, and interpersonal or assertiveness skills. The extant literature on future thinking indicates that more concerted targeting of future thinking within these tasks may lead to better outcomes.

A focus on generating specific events located in time and space during therapeutic tasks may increase the likelihood of producing richly detailed future thinking as content for therapeutic work. Promoting mental imagery and sensory details can enhance the sense of pre-experiencing and elicit associated beliefs, concerns and predictions to be leveraged. The personal significance of future events and their outcomes can be emphasized, linking them to clear goals and values related to approach or avoidance behaviors. Importantly, anticipated emotional states – and the felt sense of that anticipation – can help individuals savor positive emotions or habituate to less palatable negative emotions.

Generating divergent possibilities for future events may foster cognitive flexibility and a greater sense of choice. Additionally, spending time elaborating on the process and steps of plans or goals is likely an important factor in improving perceived control. People may be primed for future thinking by recalling concrete past experiences to fuel the generation and anticipation of future events. Tasks such as daily or weekly recording of experiences, including positive evidence logs, and reminders to engage in purposeful future thinking for specific behaviors each day (i.e., remembering to think about the future) can promote this.

Thus, more consideration of the role of future thinking in general, as well as in therapy, is required. Given the known dysfunction of future thinking in psychopathology, this may enhance therapeutic tasks. Bringing future thinking into focus as a central driver of change could prove to be time well spent in the present.

David J. Hallford
School of Psychology, Deakin University, Melbourne, VIC, Australia

1. MacLeod A. Prospection, well-being, and mental health. Oxford: Oxford University Press, 2017.
2. Holmes EA, Blackwell SE, Burnett Heyes S et al. Annu Rev Clin Psychol 2016; 12:249-80.
3. Du JY, Hallford DJ, Grant JB et al. Clin Psychol Rev 2022;95:102162.
4. Pillny M, Hallford DJ, Böge K. Behav Ther 2024;55:885-97.
5. Van den Bergh O, Brosschot J, Critchley H et al. Perspect Psychol Sci 2021;16: 225-46.
6. Xie Y, Huang SN, Cui JF et al. Cogn Ther Res 2024;15:1-8.
7. Blackwell SE, Schönbrodt FD, Woud ML et al. Psychol Med 2023;53:6113-23.
8. Hallford DJ, Rusanov D, Yeow JJ et al. Cogn Ther Res 2023;47:20-37.
9. Collado A, Stokes A. Psychol Addict Behav 2024;38:134-52.

DOI:10.1002/wps.21356

Advancing mental health in university students: future directions in literacy and digital tools

Epidemiological studies in several countries have documented increased prevalence rates of common mental disorders – including anxiety, depression and substance use – both in the general population of young people and in university students[1,2]. Roughly one in four undergraduate students screen positive or meet diagnostic criteria for a common mental disorder[3]. Since the majority of mental disorders have their onset in adolescence and early adulthood, it is not surprising that mental health problems in undergraduate students are often pre-existing or emerge soon after the transition to university and persist[3].

Untreated mental health disorders in young people have very clear associations with academic problems, impaired psychosocial functioning, and self-harm[4]. Given the well-established importance of well-being to academic success and healthy psychological and social coping during this critical period in development, university leadership and stakeholders have increasingly had to grapple with the pressing issue of how to address the growing need for student mental health support[4].

Authoritative papers and national guidelines recommend a blended strategy of mental health and well-being promotion that would benefit all students (universal), offered together with rationalized targeted early intervention (stepped care) for students at higher risk or more symptomatic, and facilitated pathways to specialized mental health services for those with more severe mental illness[4].

There is acknowledgment of the importance of ensuring that programs, initiatives and resources are coordinated, developed in partnership with students and their families, and that these programs are evidence-based and continually evaluated to ensure relevance, ongoing improvement, and responsivity[4]. Moreover, there is recognition of the need for rigorous large-scale longitudinal data collection, using validated methods, to accurately track student mental health needs and inform resource planning, identify support barriers and gaps, and advance understanding of what resources and services work, why they work, for whom do they work, and who gets left behind.

Over the past several years, dedicated research funding has supported large multi-national cross-sectional and longitudinal research documenting trends in student mental health outcomes over the COVID-19 pandemic, investigating modifiable risk and protective factors, and identifying at-risk groups to inform prevention and early intervention efforts.

Studies have provided evidence that students from minoritized groups report a higher risk of common mental disorders and yet the lowest rates of treatment[5], and that increases in the rates of anxiety, trauma, depression and self-harm are much higher in females compared to males[3]. Social connectedness and a sense of belonging seem to be protective against developing anxiety and depressive symptoms, while substance misuse (e.g., binge drinking and regular cannabis use) increase the likelihood of screen positive status for anxiety and depression over the transition to university[6]. Substance abuse is a major risk factor predicting not only reduced academic outcomes, but also increased risk of chronic refractory psychiatric illness. Reduction in the quality of sleep is common over the course of the academic year, and insomnia in students is associated with increased screen time, reduced recreation and exercise, and increased cannabis use, all of which predict worsened well-being, mental health, and academic performance[7].

These findings underscore the importance of effective mental health literacy tailored to be relevant and engaging for university students. This literacy does not only improve understanding of the determinants of well-being, but also accelerates the translation of knowledge into action through improved pro-health behaviors. While some work has been done to show the benefits of mental health literacy in secondary school students, relatively little work has focused on university students. Mental health literacy could be a promising approach to improve student emotional self-awareness and lifestyle choices around sleep hygiene, study-life balance, and substance use, whilst reducing stigma and other common attitudinal and practical barriers to appropriate help-seeking in this population[8].

Preliminary data suggest that mental health literacy embedded in the curriculum can be an effective and acceptable way to deliver this information to undergraduates across disciplines, with positive effects on psychosocial coping, making healthy lifestyle choices, and knowing how to seek help when and if needed[8].

University years are a critical developmental period during which the brain is plastic and exquisitely poised for learning, but at the same time vulnerable to the toxic effects of binge drinking, recreational drugs, and poor sleep. Developing healthy socio-emotional and behavioral coping resources is very much a work in progress[4].

Given the scope of student well-being and mental health needs, scalable and sustainable solutions are of paramount importance. One approach gaining traction is the use of digital tools to enhance well-being efforts and signpost students to the indicated type and level of support. Student mental health portals are being launched across campuses to serve as centralized hubs for care, providing students with easy access to information, self-help materials, and contact points for clinical escalation.

Portals that arrange services along a stepped care continuum can guide students through different levels of care based on their needs, from self-management applications for minor challenges to direct referrals to professional help for more critical care. Those portals that are positioned as "digital front doors" can streamline the intake process, ensuring that students are directed to the most appropriate resources quickly and efficiently. By integrating these digital tools into student mental health programming, universities can coordinate information sharing and enhance efficiency, making it possible to support a larger number of students with diverse needs. These tools can provide immediate and accessible resources, helping to bridge gaps in traditional service delivery and

increase the likelihood that students receive timely and appropriate care.

A major challenge for this work is engaging students. University students are faced with navigating the demands of higher education over condensed academic terms. They must juggle new responsibilities and take on more autonomy in managing their life and learning commitments. While students report wanting ready access to digital tools to enhance and assist with their well-being and mental health, they are often difficult to engage in proactive management and characteristically show low persistence rates in the use of digital supports[9]. Rather, students tend to reach out when in crisis or manifesting clinically significant symptoms.

In feedback sessions, students have expressed the importance of a single point of online access featuring intuitive navigation, interface customization, and the incorporation of feedback based on their own data entries and pertaining to their own well-being and/or care plans[9].

Investing in thoughtful, proactive and effective mental health promotion and early intervention for university students should be a priority for governments, university leaders, and the mental health community. Entry to university coincides with a critical period of biological and psychosocial development, with unique as well as common challenges faced by young people. While most mental health problems fall in the mild to moderate range, if unsupported they can persist and interfere with academic and personal growth and development. In addition, this is a period in which, for a minority of students, severe mental illness emerges, and early recognition and fast-tracking to specialized psychiatric care could have major prognostic implications.

Digital resources thoughtfully embedded in evidence-based services can improve access, inform triage, and increase efficiency in care – all vital to developing sustainable and responsive student mental health support.

Anne Duffy[1,2], Quynh Pham[3,4]
[1]Department of Psychiatry, Queen's University, Kingston, ON, Canada; [2]Department of Psychiatry, University of Oxford, Oxford, UK; [3]Centre for Digital Therapeutics, University Health Network, Toronto, ON, Canada; [4]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

1. Ten Have M, Tuithof M, van Dorsselaer S et al. World Psychiatry 2023;22:275-85.
2. Thorley C. Not by degrees: improving student mental health in the UK's universities. London: Institute for Public Policy Research, 2017.
3. Auerbach RP, Mortier P, Bruffaerts R et al. J Abnorm Psychol 2018;127:623-38.
4. Duffy A, Saunders KEA, Malhi GS et al. Lancet Psychiatry 2019;6:885-7.
5. Lipson SK, Zhou S, Abelson S et al. J Affect Disord 2022;306:138-47.
6. Adams KL, Saunders KE, Keown-Stoneman CDG et al. BMJ Open 2021;11:e047393.
7. King N, Pickett W, Keown-Stoneman CDG et al. BJPsych Open 2023;9:e210.
8. King N, Linden B, Cunningham S et al. BMC Psychiatry 2022;22:515.
9. Pankow K, King N, Li M et al. Early Interv Psychiatry 2024;18:226-36.

DOI:10.1002/wps.21357

A new WHO roadmap for mental health policy reform

Mental health systems across the world remain in urgent need of reform. Despite decades of advocacy and a growing body of evidence highlighting the harmful effects of institutionalization, this persists – not only in large psychiatric hospitals but also in smaller community services that retain an institutional character[1]. High-quality community-based services are scarce, waiting lists are unacceptably long, and rights-based, person-centered, recovery-oriented interventions – though evidence-based – have yet to be meaningfully integrated into policy or practice[2]. Mental health services too often remain focused on diagnosis and medication[2,3]. Coercive practices – including seclusion, restraint, and involuntary admission and treatment – continue to harm people, discourage help-seeking, and violate basic human rights[4,5]. Meanwhile, poverty, discrimination, conflicts, the climate crisis, and social exclusion deeply shape mental health, and tackling these root causes must be central to both policy and service delivery[6].

Nonetheless, resources remain inadequate. Public spending on mental health is critically low, with a global median of just 2% of government health budgets – much of it still directed toward large institutions associated with human rights violations[7]. However, increased funding alone will not be enough. It is equally important to change the paradigm of care towards a holistic, rights-based approach that treats each person as an individual with unique needs, aspirations and strengths. Only by pairing substantial new resources with policy reforms that emphasize dignity, autonomy, and social inclusion can we realize the full potential of mental health services that truly serve people and communities.

Reforming mental health systems cannot be the responsibility of the health sector alone. A comprehensive, cross-sectoral response is essential, one that integrates across housing, education, employment, justice, social protection and beyond. However, many countries still lack robust national policies and strategic action plans reflecting this interconnected reality. Although governments have pledged progress through United Nations (UN) resolutions, international conventions, and global and regional World Health Organization (WHO) action plans, national frameworks remain fragmented and insufficient[7].

Against this backdrop, the WHO has released a Guidance on Mental Health Policy and Strategic Action Plans[8]. Aligned with the UN Convention on the Rights of Persons with Disabilities[9], the Guidance provides a comprehensive, modular framework to support the development or revision of national mental health policies in line with the latest evidence and international human rights standards.

The release of this Guidance is both timely and essential. Countries across the globe are facing a dramatic rise in mental health needs, fueled by economic instability, the climate crisis, armed conflicts, global health threats such as the COVID-19 pandemic, and other emergencies. These converging challenges are exacerbating poverty, food insecurity, and inequality – particularly in low- and middle-income countries – and placing additional strain on the mental health of populations already at risk.

The WHO Guidance is organized into five modules, each addressing a critical component of modern mental health policy. The first module introduces the Guidance by outlining the urgent policy challenges that must be addressed, including chronic underinvestment, overreliance on institutional care, limited stakeholder engagement, widespread human rights violations, and failure to address the social and structural determinants of mental health.

The second module identifies five key policy areas requiring urgent reform: leadership, governance, and enabling conditions; service organization and development; workforce and human resource development; person-centered, recovery-oriented, rights-based interventions; and the mental health sector's role in addressing the social and structural determinants of mental health and broader societal issues. For each area, the Guidance offers a flexible menu of directives, strategies and actions, that can be tailored to suit national priorities, resource levels, and system structures.

The third module outlines a process-oriented roadmap for selecting and implementing policy actions, detailing nine adaptable steps to support the development, implementation and evaluation of mental health policies and strategic action plans. This module also emphasizes the importance of participatory processes that meaningfully involve all key stakeholders – especially people with lived experience. It includes two checklists to help countries both develop and evaluate the content and the process of their policies and strategic action plans.

The fourth module presents three illustrative country case scenarios, demonstrating how countries with varying income levels and systems capacities can tailor the Guidance to their unique contexts and available resources.

Finally, the fifth module compiles a consolidated directory of all recommended policy options and strategies, designed as a practical reference tool to support policy dialogue and decision-making.

This Guidance is vital for a broad spectrum of stakeholders: policy makers and planners, health and mental health professionals, researchers, civil society organizations, community leaders and, notably, people with lived experience of mental health conditions. Different groups can apply the Guidance in multiple ways: policy makers and planners can use it to systematically review policies, identify gaps, align strategies with evidence and rights obligations, and tailor policy reforms to local contexts; service providers can draw on its person-centered, recovery-oriented directives to design and continuously improve quality of care; researchers can lead the development of new evidence and evaluation frameworks, using the Guidance to inform research priorities and shape methodologies; civil society and advocacy groups can employ its tools to foster collaborative dialogue and drive policy change; and people with lived experience can leverage it to build knowledge, support advocacy efforts, and ensure their meaningful participation in decision-making processes.

In light of many entrenched shortcomings in mental health sys-

tems and practice worldwide, this WHO Guidance bridges the gap between recognizing systemic failures and implementing context-sensitive policy reforms that redirect resources from coercive, institutional models toward holistic, community-driven care. It supports safe and strategic deinstitutionalization, the development of high-quality community-based services, and the elimination of coercion, abuse and neglect through rights-based approaches that uphold dignity and autonomy. It promotes care models that integrate responses to social and structural determinants (such as housing, education and employment), champion recognition of legal capacity and decision-making rights for service users, and empower people with lived experience to co-create policies and lead anti-stigma efforts. By broadening access to physical health, lifestyle, psychological, social and economic supports, the document moves policy beyond an overemphasis on biomedical treatment. Far from a one-size-fits-all blueprint, it encourages countries to select and adapt strategies aligned with their unique priorities and constraints, while progressively working toward more ambitious policy reforms.

We urge governments, mental health providers and their professional bodies, academia, civil society organizations and donor agencies to embed this Guidance in forthcoming mental health strategies, shift funding from institutional beds to community supports, legislate for full legal capacity, and ensure meaningful participation of people with lived experience at every stage of reform. Business as usual is no longer acceptable; with these tools now available, the path toward equitable, rights-based, and effective mental-health systems is clearer than ever.

Michelle Funk[1], Natalie Drew[1], Celline Cole[1], Peter McGovern[2], Maria Francesca Moro[1]
[1]Department of Mental Health, Brain Health and Substance Use, World Health Organization, Geneva, Switzerland; [2]Modum Bad Traumepoliklinikken, Oslo, Norway

The WHO Guidance on Mental Health Policy and Strategic Action Plans is available at https://www.who.int/publications/i/item/9789240106796. The WHO acknowledges financial support from the governments of the Republic of Korea and Portugal towards the development of the Guidance.

1. World Health Organization. World mental health report: transforming mental health for all. Geneva: World Health Organization, 2022.
2. World Health Organization. Guidance on community mental health services: promoting person-centred and rights-based approaches. Geneva: World Health Organization, 2021.
3. Gardner C, Kleinman A. N Engl J Med 2019;381:1697-9.
4. Funk M, Drew N. World Psychiatry 2019;18:43-4.
5. World Health Organization and Office of the United Nations High Commissioner for Human Rights. Mental health, human rights and legislation: guidance and practice. Geneva: World Health Organization, 2023.
6. Kirkbride JB, Anglin DM, Colman I et al. World Psychiatry 2024;23:58-90.
7. World Health Organization. Mental health atlas 2020. Geneva: World Health Organization, 2021.
8. World Health Organization. Guidance on mental health policy and strategic action plans. Geneva: World Health Organization, 2025.
9. United Nations. Convention on the rights of persons with disabilities. New York: United Nations, 2006.

DOI:10.1002/wps.21360

The Psychodynamic Diagnostic Manual, 3rd edition (PDM-3)

The DSM-III and its successors have been intended to increase diagnostic reliability and validity without embedded assumptions about the meanings and etiologies of symptoms, thereby providing researchers and clinicians, irrespective of their theoretical orientation, with present-versus-absent criteria for diagnosing psychiatric disorders. However, this primary focus on the symptomatic manifestations of disorders left out critical aspects of the patients' presentation that are essential for good clinical care. The Psychodynamic Diagnostic Manual (PDM) is intended to compensate for this weakness in descriptive taxonomies.

The effort to highlight individuals' full range of functioning (i.e., the implicit as well as the observable components of their emotional, behavioral, cognitive, interpersonal and social patterns), rather than simply label their disorders, is central to the psychodynamic and humanistic clinical traditions. The PDM outlines a diagnostic framework that is symptom-oriented like the DSM and ICD approaches, but also considers individuals' idiographic characteristics and psychological functioning across different life stages.

The first edition of the PDM[1], spearheaded by S. Greenspan and co-edited by N. McWilliams and R. Wallerstein, was published in 2006. Given its positive reception, and in response to feedback about its strengths and weaknesses, a comprehensively revised second edition was published in 2017[2].

Improving the framework of the previous editions, the PDM-3[3] reorganizes its sections in accordance with developmental chronology. The first three sections are devoted to the diagnostic process in infancy and early childhood, childhood, and adolescence, while sections 4 and 5 concern adulthood and later life. Each section, over and above the DSM/ICD development-based reorganization of disorders, presents the same diagnostic entities in the context of the clinical specificities of each age group. For example, depressive disorders are listed in childhood, adolescence, adulthood and later life sections, because their clinical manifestations, and related subjective experiences, may present crucial variations across the lifespan that need to be acknowledged in a diagnostic formulation. A final section describes assessment using the Psychodiagnostic Chart (PDC), a PDM-derived tool[4], and provides several case descriptions to enhance the clinical utility of the manual.

In each section/age group, the PDM-3 adopts a "prototypic" diagnostic approach, which provides descriptions for each style/disorder that can be considered an "ideal" to which an individual can more or less approximate (rather than distinct categories to which a given person belongs or does not belong)[5]. Within this framework, clinicians can describe individuals' functioning dimensionally in each age range along three axes: Personality (P Axis), Mental Functioning (M Axis), and Symptoms (S Axis) (with

the exception of the "Infancy and Early Childhood" section, which follows a specific multiaxial system). Each Axis provides a nuanced perspective on individual functioning to assist clinicians in creating a multifaceted diagnostic profile to determine the best possible treatment plan.

The P Axis considers both levels of personality organization (i.e., a spectrum of functioning ranging from healthy to psychotic) and personality styles/disorders – i.e., clinically familiar personality configurations (such as narcissistic, obsessive-compulsive, dependent, paranoid) as well as other patterns that have been empirically confirmed (e.g., emotionally dysregulated personalities).

The M Axis considers individual profiles of mental functioning across 13 maturational capacities (e.g., affect regulation and expression, mentalization, bodily experiences and representations, impulse regulation, defensive functioning, adaptation and resilience). Clinicians rate each mental capacity on a 5-point scale, in which higher scores reflect more adaptive levels of functioning.

The S Axis considers symptom patterns. They are mostly labeled according to the DSM-5-TR, but with a specific focus on individual differences in the subjective experience of symptoms and disorders (i.e., the affective states, cognitive processes, somatic experiences, and relational patterns most often associated with each listed condition) and the related emotional experiences of clinicians.

Finally, each section concludes with a chapter focusing on the subjective experience of non-pathological conditions that might warrant specialized intervention – such as individual responses to climate change, the recent pandemic, actual or threatened war, and the experiences of patients (and therapists) who are racially, ethnically, culturally, linguistically, politically, or gender and sexually minoritized.

Studies that have been conducted to establish the reliability and clinical utility of the PDC[4], the PDM-derived tool, have shown not only adequate to good interrater reliability in samples of adults[6], but also good sensitivity in placing children into common diagnoses of developmental vs. behavioral disorders with respect to specific mental functioning patterns, global mental functioning, and levels of personality organization[7].

Recent studies involving different clinical populations have supported the relevance of PDM-related dimensions for planning of personalized clinical interventions and for prediction of therapeutic outcomes. For example, a naturalistic study on a sample of patients with feeding and eating disorders (EDs), evaluated with both the Structured Clinical Interview for DSM-5 - Clinical Version (SCID-5-CV) and the PDC, showed that higher levels of personality organization and lower personality pathology severity predicted lower ED-specific psychopathology at treatment termination, even when controlling for baseline ED symptoms. Moreover, higher levels of overall mental functioning, identity integration, mentalizing capacity, and self-coherence were related to better therapeutic outcomes, whereas DSM-5 categories did not have an impact on symptom change[8]. Further on, single case studies have exemplified the clinical utility of in-depth assessment of individual characteristics (i.e., personality style and level of organization, mental functions, subjective experience of symptoms) in children, adolescents and adults with diverse mental health conditions[9].

Unlike the DSM and ICD frameworks, the PDM diagnostic approach provides information for developing a case formulation that is sufficiently psychologically rich to guide effective treatment planning, especially when psychotherapy is the recommended intervention. Even though it is based mainly on psychodynamic research and clinical experience, the PDM-3 case formulation can be also useful in non-psychodynamically oriented practice settings, given that it provides a careful and jargon-free description of the patient, informed by neuroscience and always in dialogue with cognitive-behavioral perspectives.

The PDM reconciles the diagnostic process with its clinical implications, clearly supporting what practitioners have long realized – that every treatment should be tailored to the individuality of the patient and the patient's unique context. The PDM-3 provides updated clinical implications for treatment focus that may be familiar to clinicians trained in psychodynamic approaches, but are also applicable to those with other therapeutic backgrounds. For each condition, clinical guidelines are offered to enhance relevant dimensions of the therapeutic relationship, including the therapist's emotional responses and the therapeutic alliance, both of which have been shown to relate to treatment outcomes.

In summary, the PDM aims at representing a "taxonomy of people", rather than a "taxonomy of disorders", highlighting the clinical value of considering who one is, rather than what one has. Although the approach may appear more complex and time-consuming than that of the DSM/ICD, and despite the fact that PDM-based empirical research is still in its infancy, we strongly believe that the diagnostic process has no simple, easily applied formula. The PDM aspires to constructively bridge the gaps between descriptive psychiatry, psychodynamic research, clinical experience, and the psychometric/empirical traditions that shape diagnostic reasoning. It adds a much-needed perspective on existing taxonomies, enabling clinicians to describe and evaluate personality patterns, related social and emotional capacities, unique mental profiles, and patients' subjective experiences of symptoms.

Vittorio Lingiardi[1], Nancy McWilliams[2]
[1]Department of Dynamic and Clinical Psychology, and Health Studies, Faculty of Medicine and Psychology, Sapienza University of Rome, Rome, Italy; [2]Rutgers Graduate School of Applied and Professional Psychology, Piscataway, NJ, USA

The authors thank L. Muzi and G. Lo Buglio for their collaboration.

1. PDM Task Force. Psychodynamic diagnostic manual. Silver Spring: Alliance of Psychoanalytic Organizations, 2006.
2. Lingiardi V, McWilliams N. Psychodynamic diagnostic manual - 2nd ed. (PDM-2). New York: Guilford, 2017.
3. Lingiardi V, McWilliams N. Psychodynamic diagnostic manual - 3rd ed. (PDM-3). New York: Guilford, 2025.
4. Gordon RM, Bornstein RF. Psychoanal Psychol 2018;35:280-8.
5. Westen D, Gabbard GO, Blagov P. In: Krueger RF, Tackett JL (eds). Personality and psychopathology. New York: Guilford, 2006:355-84.
6. Biberdzic M, Grenyer BFS. J Pers Assess 2022;105:436-44.
7. Fortunato A, Tanzilli A, Lingiardi V et al. Res Psychother 2022;25:73-87.
8. Muzi L, Tieghi L, Rugo M et al. Psychoanal Psychol 2021;38:300-10.
9. Malone JC, Piacentini E, Speranza M. Psychoanal Psychol 2018;35:339-45.

DOI:10.1002/wps.21361

Trauma under psychedelics: how psychoactive substances impact trauma processing

The Hamas-led attack in southern Israel on October 7, 2023 was one of the deadliest terror attacks in history, resulting in 1,182 fatalities, over 4,000 wounded, and 251 taken hostage[1]. The Nova festival, an 18-hour rave held in the Gaza Envelope region, suffered the highest casualties in the attack, with over 370 killed.

For the Nova attendees, the highlight of the all-night party was sunrise, and many of them reported taking psychoactive substances timed to take effect at dawn. Less than half an hour after sunrise, the first rockets came in sight. As a result, survivors endured prolonged exposure to acute, life-threatening traumatic events, many while under the influence of psychoactive substances. These tragic circumstances created an unprecedented opportunity to examine the effect of peritraumatic exposure to psychoactive substances on short- and long-term impact of severe, life-threatening trauma.

From an estimated population of 3,710 Nova survivors (66% male; 76% 18-24 years old), we identified 1,239 eligible individuals. Of these, 923 (74.5%) completed the study questionnaire by February 21, 2024, of whom a total of 107 (11.6%) did not meet the DSM-5 Criterion A for post-traumatic stress disorder (PTSD), as they were not directly exposed, and 44 (4.8%) did not complete all questions, leading to their exclusion. This resulted in an analytic cohort of 772 survivors (487 males; mean age: 26.96±6.55 years). The study was approved by the University of Haifa Ethics Committee, and all participants provided informed consent. We systematically collected data on exposure to psychoactive substances and peritraumatic experiences.

Primary outcome measures included the PTSD Checklist for DSM-5 (PCL-5) with cutoff score set at 33, and the Kessler Psychological Distress Scale (K6) with cutoff score set at 13. Secondary outcomes included a 0-100 metric of perceived substance helpfulness, sense of control, and feelings of isolation during trauma exposure. Post-traumatic processing measures included self-perceived social interactions, social support, feelings of guilt, and sleep quality.

Seventy-two percent (N=556) of survivors reported being under the influence of psychoactive substances during the attack, with most (79.1%) consuming them within the three hours prior to the attack. Due to the prevalence of polysubstance use and in order to isolate the effects of individual substances, analyses focused on participants that were under the influence of a single substance: hallucinogens (psilocybin or lysergic acid, LSD; N=84); 3,4-methylenedioxymethamphetamine (MDMA) (N=99); or cannabis and/or alcohol (N=68). No substance was used by 216 participants. Group differences were analyzed using a linear regression model with substance groups as independent variables, and a multivariate model with age, sex and time from event as covariates. Only models with the lowest Bayesian information criterion (BIC) are presented. Additional results and polysubstance analyses are presented in the supplementary information.

Significant between-groups differences were found on the metric of perceived substance helpfulness during the traumatic event ($F_{2,247}$=6.14, p=0.003, R^2=0.05). Specifically, individuals in the MDMA (62.6±21.7, β=12.4, p=0.001) and hallucinogens (61.5±28.3, β=11.3, p=0.004) groups perceived more substance helpfulness during the traumatic event, as compared to the cannabis/alcohol group (50.2±20.6). This finding is unlikely to be explained by differential scope and impact of the traumatic event, as we found those to be similar across groups (all p values >0.15). Anecdotal reports suggest that survivors who were under the influence of MDMA during the trauma experienced reduced sensations of fear and threat as the event unfolded.

PTSD symptom severity scores differed significantly between groups ($F_{3,229}$=4.8, p=0.003, R^2=0.06), with significantly higher PCL-5 scores in the cannabis/alcohol group (48.3±12.8) compared to the no-use group (39.8±14.9, β=8.5, p=0.006). Notably, mean PCL-5 scores across groups were high (41.3±15.3), with all four groups exceeding the clinical cutoff score of 33 (all p values <0.05).

Mental distress scores also differed significantly between groups ($F_{3,250}$=4.3, p=0.006, R^2=0.05), due to lower K6 scores in the MDMA group (10.5±5.1) compared to the no-use group (12.2±5.1, β=–1.6, p=0.049), and higher scores in the cannabis/alcohol group (14.4±4.1) compared to the no-use group (β=2.3, p=0.031). Mean K6 scores in the cannabis/alcohol group were significantly higher than the clinical cutoff of 13, whereas scores in the MDMA and no-use groups were significantly below this threshold (all p values <0.05).

During the peritraumatic period, substance groups differed significantly in the extent of social interactions ($F_{3,463}$=4.5, p=0.004, R^2=0.03), with the MDMA group reporting significantly higher levels (76.5±26.2) compared to the no-use group (66.5±27.1, β=10.0, p=0.003). Sleep quality also varied significantly across groups ($F_{3,463}$=4.6, p=0.004, R^2=0.03). Compared to the no-use group (37.5±30.0), the MDMA group reported better sleep quality (45.4±30.9, β=7.9, p=0.025), while the cannabis/alcohol group reported worse quality (29.4±25.0, β=–8.1, p=0.046).

No significant group differences were found in perceived control ($F_{3,463}$=0.59, p=0.62) or feelings of social isolation ($F_{3,463}$=1.53, p=0.21) during trauma exposure. In the peritraumatic period, feelings of guilt ($F_{3,463}$=0.97, p=0.41) and perceived support from friends and family ($F_{3,463}$=2.5, p=0.056, R^2=0.016) did not differ between groups.

These findings suggest that trauma exposure under the influence of MDMA is associated with reduced psychological distress, higher sociality and improved sleep quality in the post-traumatic period, possibly mediated through MDMA's known effects of reducing negative emotions and elevating prosociality[2,3].

This beneficial effect of MDMA aligns with evidence from MDMA-assisted psychotherapy studies highlighting reduction of negative affect as key to its therapeutic efficacy[4-6]. Clinical protocols for MDMA-assisted psychotherapy suggest that re-experiencing traumatic events in a safe setting, while exposed to MDMA's prosocial and fear-reducing effects, may enhance the benefits of psychotherapy for PTSD[2-4]. The present study extends this idea by demonstrat-

ing that, even outside a structured therapy setting, MDMA may facilitate adaptive post-trauma social behaviors that could support psychological recovery.

Survivors who were under the influence of cannabis and/or alcohol during the attack exhibited worse sleep quality and poorer clinical outcomes, including higher mental distress and post-traumatic symptoms. These findings align with previous research demonstrating the detrimental effects of alcohol on trauma processing, including increased risk of peritraumatic dissociation, anxiety, depression, and acute stress disorder symptoms[7].

Limitations of this study include lack of control over substance choice, dosage, and intake time, as well as potential personality-based selection biases. Exposure to substances was self-reported and is prone to information biases. While the study captures real-world trauma survivors' behavior, its generalizability might be limited. Findings reflect only the initial post-traumatic period, which, though predictive, may not capture long-term clinical outcomes. Additionally, cannabis and alcohol were grouped, due to sample size constraints, limiting substance-specific analyses.

Survivor bias is inherent, and survivors with more severe symptoms may be under-represented in our cohort. However, the mean PCL-5 score in our sample across groups is well above the clinical cutoff, suggesting substantial post-traumatic symptoms. Unmeasured confounders are unavoidable, and causal assumptions should be made with much caution, if at all. Further research should explore the mechanisms linking psychoactive substances to trauma recovery and explain the putative protective role of MDMA and detrimental effect of cannabis and alcohol.

This unprecedented natural experiment offers novel insights into how psychoactive substances influence trauma processing during acute trauma and in the initial post-traumatic period. As part of an ongoing longitudinal study, these findings have important clinical implications for both survivors of this attack and trauma survivors more broadly.

Ophir Netzer[1], Noa Magal[1], Yonatan Stern[1], Tzuk Polinsky[1], Raz Gross[2,3], Roee Admon[1,4], Roy Salomon[1,4-6]

[1]School of Psychological Sciences, University of Haifa, Haifa, Israel; [2]Division of Psychiatry, Sheba Medical Center, Ramat Gan, Israel; [3]Department of Epidemiology and Preventive Medicine, and Department of Psychiatry, School of Public Health and School of Medicine, Faculty of Medical and Health Sciences, Tel Aviv University, Tel Aviv, Israel; [4]Integrated Brain and Behavior Research Center, University of Haifa, Haifa, Israel; [5]Department of Cognitive Science, University of Haifa, Haifa, Israel; [6]SafeHeart, Israel

The authors wish to thank SafeHeart management and early volunteers (R. Plonsker, G. Simon, K. Dessau, T. Zagursky, Y. Grynbaum, N. Tadmor, I. Tartakovsky, I. Hacmun, S. Cohen, S. Maor, D. Halperin and E. Atun) who worked tirelessly to ensure treatment for the survivors. They extend their deepest gratitude to the survivors of the festival who made a tremendous effort to help others in such difficult times. R. Admon and R. Salomon contributed equally to this paper. Supplementary information on this study is available at https://osf.io/qwks7/.

1. All Party Parliamentary Group for UK-Israel. 7 October parliamentary commission report. www.7octparliamentarycommission.co.uk.
2. Mitchell JM, Ot'alora GM, van der Kolk B et al. Nat Med 2023;29:2473-80.
3. Kamilar-Britt P, Bedi G. Neurosci Biobehav Rev 2015;57:433-46.
4. Vermetten E, Yehuda R. Neuropsychopharmacology 2020;45:231-2.
5. Mithoefer MC, Mithoefer AT, Feduccia AA et al. Lancet Psychiatry 2018;5:486-97.
6. Schmid Y, Bershad AK. Biol Psychiatry Cogn Neurosci Neuroimaging 2024;9:490-9.
7. Nacasch N, Malka T, Zohar J et al. World Psychiatry 2024;23:452-4.

DOI:10.1002/wps.21363

Prevalence of clinically significant radiological abnormalities in people with first episode psychosis

Magnetic resonance imaging (MRI) can be used to identify secondary psychoses caused by structural brain abnormalities, which may require different treatment from primary psychoses[1]. However, there is no international consensus as to whether MRI should be routinely offered in first episode psychosis[2-5]. We examined MRI radiology reports in a large sample of people with first episode psychosis, drawn from the South London and Maudsley NHS Foundation Trust Biomedical Research Centre Case Register. We determined the clinical significance of MRI scans by assessing the proportion of patients with a scan that was abnormal, and the proportion of scans that led to a change in the clinical management.

The study population comprised all patients who received a first diagnosis of a psychotic disorder within a 14-year period (from January 1, 2007 to June 30, 2021). In those who underwent MRI in an 18-month window around the index diagnosis, we determined the indication for the scan, the results of the scan, and any subsequent change to clinical management. The project was approved by the Oxfordshire Research Ethics Committee (23/SC/0257).

We categorized indication for MRI as: cognitive impairment (including suspected dementia), head injury, neurological features (for example, focal neurological signs or seizures), headaches, suspected encephalitis, suspected space-occupying lesion (including suspected brain metastases), hyperprolactinaemia, other atypical presentation (such as unusual age of onset or rapid onset), routine screening, and not specified. An MRI was coded as "normal" if this was specified in the radiology report, or if the findings were described as "within normal limits", a "normal variant", "no abnormality detected", "normal for age" or words to this effect. When abnormalities were reported, we specified the finding and grouped them following the classification used in the meta-analysis by Blackman et al[6]. The broad categories comprised atrophy, cyst, pituitary abnormality, tumour, vascular abnormality (excluding white matter), ventricular abnormality, white matter abnormality, or other.

We used logistic regression to examine the association of indication for the scan with having an abnormal result. The same approach was used to examine the relationship between indication and a subsequent change in clinical management. Covariates included in these models were age, ethnicity, gender, and primary

diagnosis. All statistical analyses were carried out using R statistical software version 4.3.0, and the glm function was used for the logistic regression models.

We identified 23,953 patients with a first diagnosis of psychosis, of whom 1,693 (7.1%) were referred for an MRI within the 18-month window around the index diagnosis. Radiological reports were available for 1,486 patients. Among these patients, the most common indication for the scan was routine screening (n=615, 41.4%), followed by other atypical presentation (n=184, 12.4%) and cognitive impairment (n=162, 10.9%). The indication was "not specified" in 279 cases (18.8%) (see also supplementary information).

Patients with an abnormal MRI were 380 (25.6%). Any white matter abnormality was found in 206 patients (13.9% of all scans), any atrophy in 101 (6.8%), any vascular abnormality in 71 (4.8%), any cyst in 31 (2.1%), any ventricular abnormality in 18 (1.2%), any pituitary abnormality in 17 (1.1%), any tumour in 10 (0.7%), and any other abnormality in 28 (1.9%). Among patients whose indication for MRI was routine screening, 91 (14.8%) had an abnormal scan (see also supplementary information).

Compared to routine screening, clinical indications of suspected space-occupying lesion (odds ratio, OR=5.3, 95% CI: 2.2-12.9), suspected encephalitis (OR=3.2, 95% CI: 1.3-7.2), neurological features (OR=2.6, 95% CI: 1.5-4.4), cognitive impairment (OR=2.1, 95% CI: 1.3-3.3) and other atypical presentation (OR=1.9, 95% CI: 1.2-2.9) were associated with an abnormal scan in the adjusted model (see also supplementary information). Of the covariates modelled, only age was associated with an abnormal MRI (OR=1.06, 95% CI: 1.05-1.07).

In total, 137 (9.2%) scans were followed by a change in clinical management, including referral to another specialty (n=60), change in diagnosis (n=36), and further investigations (n=34). In the subgroup of patients who had a change in management following a routine screening MRI (n=28), 13 were referred for further investigation and 13 were referred to another specialty. However, none of these routine screening cases were associated with a change of diagnosis or identification of a secondary cause for psychosis.

In the adjusted model, indications of suspected encephalitis (OR=4.1, 95% CI: 1.3-11.3), neurological features (OR=3.5, 95% CI: 1.7-6.9), suspected space-occupying lesion (OR=3.2, 95% CI: 1.0-8.8), head injury (OR=2.8, 95% CI: 1.0-6.8) and cognitive impairment (OR=2.6, 95% CI: 1.4-5.0) were associated with a change in clinical management (see also supplementary information). Of the covariates modelled, only age had a statistically significant effect, with each increasing year of age being associated with a change in management (OR=1.04, 95% CI: 1.03-1.06).

The proportion of patients with radiological abnormalities was remarkably similar to that found in a meta-analysis of MRI data based on patients with first episode psychosis (26.4%)[6]. However, in the present study, the proportion of scans that led to a change in clinical management was higher (9.2%) than in that meta-analysis (5.9%). This may reflect differences in the patient populations; notably, around half of the studies included in the meta-analysis were based on research cohorts, and all excluded patients with clinical features suggestive of a secondary cause of psychosis[7]. Although previous studies of radiological findings in psychosis have examined whether these led to a change in management, they did not report the changes in detail. By accessing clinical records in a large sample, the present study was able to determine why each scan had been ordered and what actions, if any, were taken in response to the radiology report.

Scanning in the absence of a specific indication ("routine screening") was not associated with any diagnostic changes and did not identify any secondary psychoses. The latter finding is consistent with data from a study of routine MRI screening in first episode psychosis (n=349), which also failed to identify any cases of secondary psychosis[8]. These data suggest that the likelihood of detecting secondary causes of psychosis in the absence of a clinical indication for MRI is small[9]. However, this may depend on the quality of the clinical evaluation: a detailed history and examination may not be feasible in busy services with limited resources, or if the patient is difficult to assess.

Uncertainty about the prevalence of clinically relevant radiological abnormalities in people with first episode psychosis underlies a lack of consensus on the clinical utility of MRI in the clinical assessment of this population. The present study, the largest to investigate this issue to date, indicates that around 9% of scans in these people are followed by a change in clinical management, supporting the use of MRI as part of the initial assessment of people with first episode psychosis.

Thomas J. Reilly[1-3], Matt Butler[2,3], Ahmed Abdelsamie[3], Ishaac Awatli[3], Sneha Barai[3], Luke Baxter[3], Theo Boardman-Pretty[3], Rebecca Phelps[3], Anam Waqar[3], Amelia Jewell[3], Graham Blackman[1,2,4], Dominic Oliver[1,2,4], Ashwin V. Venkataraman[3,5], Matthew J. Kempton[2], Philip McGuire[1,4]

[1]Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK; [2]Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; [3]South London and Maudsley NHS Foundation Trust, London, UK; [4]NIHR Oxford Health Biomedical Centre, Oxford, UK; [5]Centre for Healthy Brain Ageing, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

T.J. Reilly is supported by a UK Medical Research Council (MRC) Clinical Research Training Fellowship (no. MR/W015943/1). M. Butler is a Wellcome Trust Doctoral Clinical Research Fellow. A.V. Venkataraman is funded by the UK National Institute for Health Research (NIHR) and supported by the NIHR Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. This study was funded in part by the NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed here are those of the authors and not necessarily those of the NIHR, the NHS or the UK Department of Health and Social Care. Supplementary information on this study is available at https://osf.io/4q2xs?view_only=7e493f18a7324f24bd31742a13374248.

1. Keshavan MS, Kaneko Y. World Psychiatry 2013;12:4-15.
2. National Institute for Health and Care Excellence. Structural neuroimaging in first-episode psychosis. London: National Institute for Health and Care Excellence, 2008.
3. American Psychiatric Association. Practice guideline for the treatment of patients with schizophrenia, 3rd ed. Washington: American Psychiatric Association, 2021.
4. Galletly C, Castle D, Dark F et al. Aust N Z J Psychiatry 2016;50:410-72.
5. Grover S, Chakrabarti S, Kulhara P et al. Indian J Psychiatry 2017;59(Suppl. 1):S19-33.
6. Blackman G, Neri G, Al-Doori O et al. JAMA Psychiatry 2023;80:1047-54.
7. Patel R, Oduola S, Callard F et al. BMJ Open 2017;7:e013113.
8. Sommer IE, de Kort GAP, Meijering AL et al. Schizophr Bull 2013;39:815-9.
9. Butler M, Scott F, Stanton B et al. BJPsych Bull 2022;46:152-6.

DOI:10.1002/wps.21362

Incidence of gynecological cancers following prolactin-increasing antipsychotic use: a population cohort study

Long-term use of prolactin-increasing antipsychotics has been found to be associated with breast cancer[1]. A recent meta-analysis of observational studies has also reported an increased risk of gynecological cancers by nearly 70% in relation to antipsychotic use[2]. Nevertheless, that meta-analysis was limited by small sample size, heterogeneity across studies, narrow inclusion of specific cancer types, as well as by the case-control design adopted by most included studies. There was also no direct comparison between users of prolactin-increasing and prolactin-sparing antipsychotics on a comparable timeframe[3-6].

We took advantage of the longitudinal anonymized territory-wide electronic health records maintained by the Hospital Authority of Hong Kong to build a large retrospective cohort and estimate the weighted incidence rate ratio (IRR) and rate difference (RD) of gynecological cancers between users of prolactin-increasing and prolactin-sparing antipsychotics. This study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster and the Hospital Authority Central Institutional Review Board (no. UW 20-113, CIRB-2022-015-5).

Women aged 18 to 85 years who initiated antipsychotic treatment between January 2006 and December 2018 in Hong Kong were identified. Patients with previous antipsychotic use, hysterectomy or mastectomy, or cancers (except non-melanoma skin cancer) before the index date were excluded. Patients were followed from antipsychotic initiation until the diagnosis of gynecological cancer, or 365 days after change of antipsychotic (a switch to antipsychotic other than the one initiated on the index date), or 365 days after the discontinuation of antipsychotics, or death, or December 31, 2023, whichever occurred the earliest. To avoid reversed causality, patients developing cancer within 30 days of the index date were censored upon cancer diagnosis.

Study exposure was defined as use of prolactin-increasing antipsychotics (i.e., those associated with moderately to highly elevated prolactin levels), with use of prolactin-sparing antipsychotics (i.e., those associated with low or non-elevated prolactin)[5] as the comparator (see supplementary information for the classification of antipsychotics). The main outcome was gynecological cancer (ICD-9: 179-180.9, 182.0-184.9) and its subtypes, including cervical (180), endometrial (179, 182), ovarian (183) and vaginal cancers (184)[5,7].

Covariates included age, reproductive history (with or without childbearing experience), year of cohort enrollment, baseline comorbidities, and number of days on other medications over the past two years. Comorbidities included diagnoses of schizophrenia, bipolar disorder, paranoid disorder, other non-organic psychoses, depression, anxiety disorder, dementia, substance use disorders, diabetes, hypertension, obesity (body mass index ≥30), disorder of lipid metabolism, autoimmune diseases, chronic kidney disease, ovarian dysfunction, human papillomavirus infection, prolactinoma, and human immunodeficiency virus (HIV) infection, recorded since the inception of the database in 1993. Prior use of other medications included aspirin, beta-blockers, calcium channel blockers, contraceptives, diuretics, female hormones, hypnotics and anxiolytics, non-steroidal anti-inflammatory drugs, opioids, paracetamol, renin-angiotensin system inhibitors, statins and digoxin.

Poisson regression with inverse probability of treatment weighting and robust variance estimation was used to estimate the IRR with 95% confidence intervals (CIs) of gynecological cancers between prolactin-increasing and prolactin-sparing antipsychotic users. Weighed RD was also estimated. Age-stratified analysis was conducted by the threshold of ≥51. In addition, patients with psychotic disorders, bipolar disorder, paranoid disorder, other non-organic psychoses, and none of these conditions were separately analyzed. Head-to-head comparisons were made between users of specific agents.

Several sensitivity analyses were further conducted. First, in a repeated analysis, only patients who switched to the different prolactin-inducing category of antipsychotics were censored. Second, we excluded antipsychotics with moderate prolactin-increasing effects from the prolactin-increasing group. Third, we reduced the period for defining drug discontinuation from 90 to 60 days. Fourth, we followed up patients for an additional 180 days (instead of 365 days) if they discontinued antipsychotics. Fifth, we limited the maximum follow-up time to 10 years. Sixth, we extended the initial censoring window for cancer occurrence from 30 to 60 days after the index date. Seventh, we conducted the analysis without an initial censoring window for cancers after the index date. Eighth, patients on aripiprazole were excluded, because this drug may counteract the prolactin-increasing effects of other antipsychotics. Ninth, we used breast cancer (174) as a positive control outcome, and tuberculous meningitis (013) as a negative control outcome, for a repeated main analysis. Lastly, we repeated the analysis with weighted Cox regression. All the analyses were conducted in the R statistical programming environment (version 4.1.2). Statistical tests were two-tailed, and a p value ≤0.05 was indicative of statistical significance.

In total, 84,061 female new antipsychotic users were included, with a follow-up period up to 15 years (median: 1.18; 25th-95th percentile: 1.02-1.99). Of these, 61,771 (73.5%) were prolactin-increasing antipsychotic users. The three main antipsychotics prescribed were quetiapine, haloperidol and risperidone. During the follow-up period, 126 cases of gynecological cancers were identified, with a crude incidence rate of 77.86 per 100,000 person-years. Of these, 96 cases occurred in the prolactin-increasing group (88.12 per 100,000 person-years), and 30 in the prolactin-sparing group (56.72 per 100,000 person-years).

Poisson regression showed that prolactin-increasing antipsychotic users had a significantly higher rate of gynecological cancers compared to prolactin-sparing antipsychotic users (IRR=

1.99, 95% CI: 1.13-3.50). The weighted RD of gynecological cancers was 42.58 (95% CI: 11.77-69.32) per 100,000 person-years, with the rate notably elevated among patients aged over 51 years (IRR=2.09, 95% CI: 1.02-4.26), and those diagnosed with psychotic disorders (IRR=9.69, 95% CI: 1.32-71.10).

In head-to-head drug comparisons, the increased rate remained statistically significant for the quetiapine-haloperidol and quetiapine-risperidone comparisons (IRR=2.27, 95% CI: 1.21-4.26; and IRR=2.20, 95% CI: 1.10-4.37). Regarding gynecological cancer subcategories, elevated risks of endometrial cancer (IRR=1.94, 95% CI: 1.02-3.69) and ovarian cancer (IRR=3.68, 95% CI: 1.17-11.56) were observed. The increased risk of cervical or vaginal cancer did not reach statistical significance (IRR=1.33, 95% CI: 0.44-4.04 and IRR= 2.14, 95% CI: 0.43-10.73, respectively). Similar results were seen in the sensitivity analyses, including the positive control outcome analysis on breast cancer (IRR=1.63, 95% CI: 1.10-2.43), with the negative control outcome analysis showing a non-significant IRR close to 1 (IRR=0.96, 95% CI: 0.37-2.55).

To the best of our knowledge, this is the first cohort study, and the largest real-world study, to compare prolactin-increasing with prolactin-sparing antipsychotic use in terms of incidence of gynecological cancers. Findings suggest a two-fold increased rate of this incidence among users of prolactin-increasing antipsychotics compared with those using prolactin-sparing antipsychotics, particularly evident for ovarian and endometrial cancers, but with small RDs, i.e., approximately one case per 2,300 person-years. Preclinical studies suggesting biological links[8] and observational studies supporting an empirical association[9] are far from scarce, with preliminary evidence suggesting that prolactin may promote the growth of ovarian and endometrial surface epithelial cells, enhance ovarian cancer cell survival and migration, and inhibit apoptosis[8]. Key limitations of the study include the lack of randomization and potentially limited generalizability beyond Chinese populations.

In conclusion, our study showed that prolactin-increasing antipsychotics are associated with an increased rate of gynecological cancers compared with prolactin-sparing antipsychotics, although the RD is small, entailing little clinical significance.

Francisco Tsz Tsun Lai[1-4], Wenxin Tian[1], Yuqi Hu[1], Cuiling Wei[1], Qi Sun[1], Song Song[1,2], Lingyue Zhou[1], Yue Wei[1,2], Shan Luo[3,5], Amy Pui Pui Ng[3], Joe Kwun Nam Chan[6], Heidi Ka Ying Lo[6], Esther Wai Yin Chan[1,2], Stanley Sau Ching Wong[7], Wing Chung Chang[6,8]

[1]Department of Pharmacology and Pharmacy, Centre for Safe Medication Practice and Research, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong; [2]Laboratory of Data Discovery for Health, Hong Kong Science Park, Sha Tin, Hong Kong; [3]Department of Family Medicine and Primary Care, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong; [4]Advanced Data Analytics for Medical Science, Hong Kong, Hong Kong; [5]School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong; [6]Department of Psychiatry, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong; [7]Department of Anaesthesiology, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong; [8]State Key Laboratory of Brain and Cognitive Sciences, University of Hong Kong, Hong Kong

F.T.T. Lai is partially supported by the Laboratory of Data Discovery for Health, Hong Kong. The authors thank the Hong Kong Hospital Authority for the provision of data, and C. Chui, C.L. Cheung, K. Lee, R. Chu, B. Liu and W. Liu for advice on the study. F.T.T. Lai and W. Tian contributed equally to this paper. Supplementary information on this study is available at https://drive.google.com/file/d/1LGt6LLM5b3fw7G1JzAPyiiltwAYySJHm/view?usp=sharing.

1. Taipale H, Solmi M, Lähteenvuo M et al. Lancet Psychiatry 2021;8:883-91.
2. de Moraes FCA, Sudo RYU, Souza MEC et al. BMC Cancer 2024;24:712.
3. Chen WL, Nithiyanantham S, Mao YC et al. Clin Psychopharmacol Neurosci 2022;20:526-35.
4. Harlow BL, Cramer DW, Baron JA et al. Cancer Epidemiol Biomarkers Prev 1998;7:697-702.
5. Klil-Drori AJ, Yin H, Abenhaim HA et al. J Clin Psychiatry 2017;78:714-9.
6. Yamazawa K, Matsui H, Seki K et al. Oncology 2003;64:116-23.
7. Hospital Authority. Hong Kong Cancer Registry. https://www3.ha.org.hk/cancereg/icd10.html.
8. Levina VV, Nolen B, Su Y et al. Cancer Res 2009;69:5226-33.
9. Clendenen TV, Arslan AA, Lokshin AE et al. Cancer Causes Control 2013;24: 741-8.

DOI:10.1002/wps.21365

Predicting epigenetic aging by the transdiagnostic internalizing spectrum vs. depressive and anxiety syndromes

Depression and anxiety are linked with higher risk for multimorbidity[1] and all-cause mortality[2]. In health research, they are often indexed as binary syndromes or symptom counts, reflective of traditional diagnostic models. However, there has long been evidence supporting dimensional and hierarchical conceptualizations of psychopathology[3], wherein the widespread comorbidity among mood and anxiety disorders is modeled in terms of a broader, transdiagnostic internalizing spectrum[4].

Evidence has accumulated that the broader internalizing factor has superior reliability and predictive validity relative to traditional diagnoses, and that its components have shared genetic diatheses, environmental risk factors, childhood antecedents, and treatment responses[5]. Kim et al[6] found that the internalizing spectrum significantly predicted mortality risk over a 20-year period (hazard ratio, HR=1.12, p<0.01), while disorder-specific variability (i.e., residuals net of their common variance) did not have predictive power (HRs=0.94-1.02). These results suggest that psychopathology-related mortality risk is captured at the level of the broader spectrum.

Less is known about relations between psychopathology and the biological aging processes underlying morbidity and mortality. Epigenetic alteration is considered one of the "hallmarks of aging", as changes in gene expression can result in the development of many age-related pathologies[7]. One such epigenetic pro-

cess is DNA methylation, or the binding of methyl group molecules to genes in a manner that inhibits or promotes their transcription. Strong associations have been observed between age and methylation in some regions of the genome, leading to the development of epigenetic clocks designed to index biological age as distinct from chronological age. Individuals are said to experience epigenetic age acceleration when their biological (or epigenetic) age exceeds their chronological age. Epigenetic age acceleration as indexed by recent measures (e.g., GrimAge, DunedinPACE) has been validated in novel samples as a replicable predictor of aging-related morbidity and mortality[8].

Epigenetic aging also correlates with a variety of psychosocial and environmental variables, including depression[9]. However, to our knowledge, it has not been studied in relation to the broader internalizing spectrum in adulthood. Given the robust evidence on the structure of psychopathology, modeling a latent internalizing liability might increase predictive power for epigenetic aging and help explain inconsistencies in prior research, as it has with mortality[6]. Consistent with the notion that health-relevant variability within traditional disorders is captured by an overarching liability, we hypothesized that a transdiagnostic internalizing factor would predict future epigenetic age acceleration, and that the variance specific to symptoms of major depressive disorder (MDD), generalized anxiety disorder (GAD), and panic disorder (net of internalizing) would not.

DNA methylation profiling was conducted on 1,309 participants from the National Survey of Midlife Development in the United States (MIDUS). Sociodemographic, psychopathology, and other health factors were assessed during MIDUS survey visits (mean age 51.3; Time 1), and blood was collected for the later Biomarker project (mean age 54.0; Time 2). Whole blood samples were subject to DNA extraction and underwent genome-wide methylation profiling using Illumina EPIC microarrays. Methylation data were scored with existing algorithms to compute four measures of epigenetic age (the Hannum, Horvath, PhenoAge, and GrimAge2 clocks) and the DunedinPACE measure of epigenetic age acceleration.

Exploratory structural equation modeling was then used to examine the covariation among DunedinPACE and the residuals of the four clocks after they were regressed on chronological age. A two-factor model fit the data well. The first factor reflected the earlier Hannum and Horvath measures designed to predict cross-sectional state (age, health state) and the second reflected the later ones developed to predict change in health (GrimAge2 and DunedinPACE). Thomson's factor scores were carried forward and labeled "state-predictive" and "decline-predictive" epigenetic aging. Confirmatory factor analysis was used to model a transdiagnostic internalizing factor with three indicators: continuous symptom counts for MDD, GAD and panic disorder as assessed by the Composite International Diagnostic Interview - Short Form. All three indicators had meaningful and statistically significant factor loadings (λ>0.40, p<0.001).

Structural equation models were then fit that tested the prediction of epigenetic aging (both state and decline factors) by the internalizing factor (Model 1), each of the symptom count variables (Models 2-4), and the residuals for each of the symptom count variables net of the internalizing factor (Models 5-7). In all models, the focal psychopathology variable (Time 1 internalizing or a symptom count) predicted later (Time 2) epigenetic aging. The focal psychopathology variable was regressed on the following six covariates: age, sex, education level, race, body mass index, and smoke pack years. Then the epigenetic aging variable was regressed on the psychopathology variable and the same covariates. All models were fit in Mplus using maximum likelihood estimation with cluster robust (Huber-White) standard errors to account for the nested, within-family structure of the data from siblings in the MIDUS study.

The latent internalizing factor was significantly associated with decline-predictive epigenetic age acceleration (Model 1: β=0.11, p=0.001). Decline-predictive epigenetic age acceleration was also significantly associated with symptoms of MDD (Model 2: β=0.07, p=0.001), but not those of GAD (Model 3: β=0.05, p=0.017) or panic disorder (Model 4: β=0.03, p=0.216). None of the symptom count residuals (net of internalizing) were significantly associated with decline-predictive epigenetic aging (Models 5-7: β=–0.03 to 0.02, p values >0.350). Across models, none of the psychopathology variables were significantly associated with state-predictive epigenetic aging (β=–0.01 to 0.04, p values >0.250). Models accounted for 32-35% of variance in decline-predictive epigenetic aging and 3-4% in state-predictive epigenetic aging (inclusive of covariates) (see also supplementary information).

These analyses evaluated the utility of the internalizing spectrum and symptoms of specific depressive and anxiety syndromes in predicting later epigenetic aging. Accelerated epigenetic aging in decline-predictive measures was significantly associated with greater internalizing and more depressive symptoms. GAD and panic disorder symptoms were nominally positive, but not significant predictors of epigenetic aging. As with mortality[6], the symptom-specific variability, net of internalizing, was not significantly associated with epigenetic aging. These findings suggest that differential methylation of CpG sites (i.e., regions in the DNA sequence where cytosine is followed by guanine in a specific direction) associated with blood biomarkers of morbidity and mortality (e.g., GrimAge) is linked with internalizing and related variables, suggesting a possible mechanism by which psychosocial experiences become biologically embedded.

Since analyses were observational (with a single time-sequenced measurement of focal constructs), little can be concluded about directionality or the role of unmeasured confounders. Additionally, methylation was measured from circulating blood samples, and it is unclear how effects may vary across tissues. Nonetheless, these results suggest that the variability within traditional depressive and anxiety disorders relevant to aging-related health is captured by the overarching internalizing spectrum. Each of the symptom counts had less predictive power than the internalizing spectrum and, critically, little to no predictive power net of internalizing.

Many possible mechanisms may underlie associations between psychopathology and biological aging (e.g., behavioral and genetic factors). However, the relationship may be directly causal, wherein

difficulties with internalizing signal inflammatory and immune-related pathways, potentially involving or leading to epigenetic alteration.

Colin D. Freilich[1], Frank D. Mann[2], Kristian E. Markon[1], Steve W. Cole[3], Robert F. Krueger[1]
[1]Department of Psychology, University of Minnesota, Minneapolis, MN, USA; [2]Department of Medicine, Renaissance School of Medicine at Stony Brook University, Stony Brook, NY, USA; [3]Department of Psychiatry & Biobehavioral Sciences and Medicine, University of California, Los Angeles, CA, USA

The MIDUS study has been funded by the MacArthur Foundation Research Network, and the US National Institute on Aging (grants nos. P01AG020166 and U19AG051426). The investigators were also partly supported by the US National Institute on Aging (grant nos. R01AG053217, R01AG077742, R21AG074705-01, and U19AG51426). Supplementary information on this study is available at https://osf.io/5qa4k.

1. Castro-de-Araujo LFS, Cortes F, de Siqueira Filha NT et al. Front Psychol 2022;13:940978.
2. van Dijk MR, Utens EMWJ, Dulfer K et al. Eur J Prev Cardiol 2016;23:552-8.
3. Ringwald WR, Forbes MK, Wright AGC. Psychol Med 2023;53:533-46.
4. Kotov R, Krueger RF, Watson D et al. J Abnorm Psychol 2017;126:454-77.
5. Watson D, Levin-Aspenson HF, Waszczuk MA et al. World Psychiatry 2022;21:26-54.
6. Kim H, Turiano NA, Forbes MK et al. World Psychiatry 2021;20:276-82.
7. López-Otín C, Blasco MA, Partridge L et al. Cell 2023;186:243-78.
8. McCrory C, Fiorito G, Hernandez B et al. J Gerontol A Biol Sci Med Sci 2021;76:741-9.
9. Oblak L, van der Zaag J, Higgins-Chen AT et al. Ageing Res Rev 2021;69:101348.

DOI:10.1002/wps.21364

Promoting collaboration, harmonization and dissemination in depression research: the ECNP Depression Meta-Network

The European College of Neuropsychopharmacology (ECNP) is a pan-European scientific association in the fields of translational neuroscience and applied brain research. Its core mission is to help ensure that advances in the understanding of brain function and human behavior are translated into better treatments and enhanced public health.

To this purpose, the College makes use of a variety of tools. The best known is its annual congress, attracting over 6,000 participants every year. However, equally vital to ECNP's mission are its networks, established roughly 20 years ago. The ECNP networks are multi-disciplinary collaborative platforms that bring together researchers from across Europe to share ideas, discoveries and best practices in translational neuroscience. Designed to facilitate the collection of essential biological, clinical and therapeutic data in a robust and replicable way, the current 25 existing ECNP networks cover a range of disease-oriented as well as transnosological (e.g., digital health, experimental medicine, nutrition, suicide, resilience) research focus areas.

Until recently, there has been a notable omission from the list of ECNP networks: depression was not included. The reasons are manifold and partially hard to unearth from history. In the beginning, it might have been a simple miss, as the formation of new networks did not follow a structured process. Later on, there was the feeling that a network on depression might be so broad to make almost impossible to select its members, given the enormous breadth of the field, with many ECNP members working on depression from different perspectives.

On the other hand, it was increasingly clear that a network on depression was mandatory, and its lack became noticeable. Depressive disorder is the mental health condition with the largest impact on disease burden, about 6% of the worldwide population being affected in the past year[1]. Its economic and societal impact is huge, not only due to its direct health care costs, but especially to its indirect costs through work absenteeism and impact on caregivers. Unfortunately, existing first-line medication and psychotherapy treatments do not work for all persons, and both treatment gap and treatment inertia for depression are large, leaving many patients inadequately treated, resulting in a sizeable proportion of them suffering from what is called treatment-resistant depression[2,3]. Consequently, within ECNP there is already a strong focus on developing and implementing more appropriate detection, prevention and treatment strategies for depression.

Depression is also a large focus of attention for other ECNP stakeholders, such as patient and family organizations and regulators. Also within industry, there has been recent progress in the drug development field that is relevant for depression (e.g., treatments targeting the glutamate system[4], psychedelic medications). Novel neuromodulation strategies – such as transcranial magnetic stimulation, transcranial direct-current stimulation, and vagus nerve stimulation – are getting more and more attention. Adding to this, depression is a vastly heterogeneous condition, much more syndromal than any other mental disorder. Therefore, "depression" might even be considered a transdiagnostic construct.

To tackle these challenges, we built up the Depression Meta-Network[5], which explicitly connects the various already existing ECNP network activities, by selecting experts working on depression from every relevant network, to build a "network of networks". This Meta-Network was formed at the end of 2023.

The overarching goal of ECNP's Depression Meta-Network is to bring together leading (pre)clinical researchers, industry, regulators and patient representatives from Europe to accelerate, across various disciplines, the understanding of the aetiologies of the depressive syndrome and improve its primary prevention, screening, diagnosis, early intervention, treatment, and thereby outcomes.

Specific aims are the following: a) to serve as a "hub" for connecting different partners (clinicians, researchers, industry, regulators, other stakeholders) in the field of depression, to share expertise and results, discuss clinical/industry/research developments, obtain European grants, and advise on better conduct of clinical trials and selection of appropriate outcome measures; b) to collect

and share information on ongoing, large-scale research projects and infrastructures focusing on depression, in order to stimulate collaboration, integration and replication of research findings, with an emphasis on addressing the heterogeneity within depression and identifying its relevant underlying symptom/behavioral/neurobiological dimensions; c) to showcase the many depression research lines present in ECNP by cross-network presentations, communication and discussions (e.g., at ECNP and other scientific meetings), and training activities.

To ensure a smooth workflow, the Meta-Network is divided into four working groups with a variety of activities:

a) *Dissemination*
- Providing a platform facilitating collaboration between researchers to perform international, multidisciplinary and multi-method research on depression. Cross-network activities that bridge various research methods/lines in the field of depression, e.g. through online meetings, ECNP symposia, symposia at other scientific meetings, and training activities are instrumental in this respect.
- Influencing policy makers, especially at the European level, to increase funding and resources for mental health research in depression, e.g. by providing information on depression prevalence trends as well as prevention and treatment developments and opportunities.
- Educating scientists, clinicians and the general public in innovative research methods and questions in the field of depression, as well as distributing, integrating and/or producing clinical guidance on how to prevent, screen for, and treat depression. As one example, we have set up an Educational Online Course.

b) *Fostering the European Union (EU) landscape in depression research*
- Providing an overview of EU depression studies and large national trials/cohorts, analyzing registry and electronic health record data on depression, pinpointing needs, and describing data collection.
- Maximizing generalizability of findings and study power through sample sharing, and homogenization of research protocols and (depression) concepts across different countries, in line with Findability, Accessibility, Interoperability, and Reuse of digital assets (FAIR) regulations.
- Submitting collaborative international multi-centre study proposals to the EU Commission and/or other national and international funding agencies.

c) *Conceptualization of depression*
- Describing the heterogeneity of depression and relevant personalized medicine approaches, depression's cross-disorder nature and preclinical translation. A Delphi process on some of these questions has recently been initiated. Also, a consensus paper on depression phenotyping has been drafted[6], that follows up recent guidance for clinical characterization of patients with depression[7].

d) *Clinical trial harmonization and stimulation*
- Stimulating platform trials and related interaction with industry and regulators. As one example, Meta-Network members are involved in the Wellcome-funded PEARL consortium, that has been granted to start a platform trial in depression[8].
- Proposing relevant trial outcome measures beyond "traditional" rating scales, to enable EU approval and market access of new antidepressants, also incorporating the voice of people with lived experience.
- Collaborating with the industry involved in pharmacological and non-pharmacological (i.e. wearables, lifestyle/psychotherapy solutions, neurostimulation techniques) interventions for depression.

To achieve these goals, the Meta-Network members meet regularly online as well as face-to-face. While core members are selected from the other ECNP networks and committees, and hence seats are not openly available, we strongly encourage and foster interaction with all relevant stakeholders. Whoever likes to contribute to one of the working groups is invited to reach out to the current network chairs (i.e., the authors of this paper). We especially appreciate concrete activities feeding into one of the working groups.

While the Meta-Network is now active since only one year, it already sparkled many interactions among a large number of researchers, as well as tangible output. This, however, can only be the first step for a long-term endeavour to facilitate and stimulate collaborative depression research in order to improve the lives of millions of affected people.

Andreas Reif[1,2], ECNP Meta-Network on Depression, Brenda W.J.H. Penninx[3,4]
[1]Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, Goethe University Frankfurt, University Hospital, Frankfurt, Germany; [2]Fraunhofer Institute for Translational Medicine and Pharmacology, Frankfurt, Germany; [3]Department of Psychiatry, Amsterdam UMC, Vrije Universiteit, Amsterdam, The Netherlands; [4]Amsterdam Public Health, Mental Health Program, Amsterdam UMC, Vrije Universiteit, Amsterdam, The Netherlands

Members of the ECNP Depression Meta-Network include D.S. Baldwin, M. Bauer, B.T. Baune, E. Binder, P. Brambilla, E. Castrén, A. Cattaneo, J.F. Cryan, G.R. Dawson, L. de Picker, V.G. Frokjaer, C. Harmer, C.A. Hartman, S. Kittel-Schneider, G.M. Knudsen, N. Koutsouleris, A. Lengvenyté, I. Myin-Germeys, C. Otte, C. Pariante, B. Penninx, C.R. Pryce, A. Reif, J. Repple, H.G. Ruhé, A. Serretti, N.J.A. van der Wee, E. Vieta, C. Vinkers, G. Zalsman, J. Brasanac and L. Han.

1. Marx W, Penninx BWJH, Solmi M et al. Nat Rev Dis Primers 2023;9:44.
2. Herrman H, Patel V, Kieling C et al. Lancet 2022;399:957-1022.
3. McIntyre RS, Alsuwaidan M, Baune BT et al. World Psychiatry 2023;22:394-412.
4. Freudenberg F, Reif-Leonhard C, Reif A. Eur Arch Psychiatry Clin Neurosci 2025; doi.org/10.1007/s00406-024-01875-z.
5. European College of Neuropsychopharmacology. Depression Meta-Network. www.ecnp.eu.
6. Repple J, Chevance A, Fried E et al. Submitted for publication.
7. Maj M, Stein DJ, Parker G et al. World Psychiatry 2020;19:269-93.
8. Koenig F, Spiertz C, Millar D et al. EClinicalMedicine 2023;67:102384.

DOI:10.1002/wps.21366

Integral Brain Health: a collaborative approach for psychiatry and neurology

For centuries, philosophers and scientists have debated on the mind-brain dichotomy, generating a division between "functional" and "organic" disorders, and between psychiatry and neurology. However, advancements in genetics and neurobiology continually erode these distinctions, and research findings increasingly support the influence of psychosocial and environmental factors on brain function and development.

The World Health Organization (WHO) defines brain health as the optimal functioning of sensory, motor, cognitive and emotional systems, while mental health is framed in relation to external factors, such as the ability to cope with stress, learn, work, thrive, and contribute to society. However, despite their distinct definitions, brain and mental health share the goal of enabling individuals to fully realize their abilities and potential.

Brain and mental health play a central role in overall health, well-being and productivity. Nevertheless, the excessive specialization and fragmentation of neurology and psychiatry have led to gaps and overlaps, with missed opportunities for collaboration and integration. In the digital era, where a knowledge-driven society and an aging population demand enhanced cognitive abilities and strong mental and social resilience, a comprehensive definition of Integral Brain Health is essential. This definition should encompass cerebral, mental and social components, all supported by a safe and healthy environment[1]. Additionally, there is a need for an index that can measure Integral Brain Health independently of language, literacy, and cultural differences.

According to the Brain Health Atlas, brain-related disorders – including mental disorders, neurological diseases, and cerebrovascular conditions – are among the leading causes of disability worldwide[2]. In 2021, these conditions contributed to over 18% of global health loss, amounting to 522 million disability-adjusted life years (DALYs). This is twice the burden of cancer (260 million DALYs) and exceeds that of cardiovascular disease (402 million DALYs). The economic impact is enormous, with an estimated $1.2 trillion in lost income, primarily driven by migraine, depression and anxiety. Additionally, global health care expenditures on brain conditions reach $1.1 trillion, with dementia, stroke, depression and anxiety accounting for a significant portion of these costs.

The prevalence of brain-related conditions is rising at an alarming rate. Since 1990, depression cases have increased by 89%, strokes by 102%, Alzheimer's disease and other dementias by 161%, and Parkinson's disease by 274%. This trend is primarily driven by rising life expectancy and declining fertility rates, leading to a growing aging population. The WHO estimates that global life expectancy increased by more than six years from 2000 (66.8 years) to 2019 (73.1 years). In contrast, healthy life expectancy rose by only 5.3 years[3].

There is a high comorbidity between psychiatric disorders and neurological conditions. For instance, the prevalence of major depressive disorder is around 20% or above among patients with epilepsy, multiple sclerosis, and Parkinson's disease, and during the acute post-stroke period[4]. The association between psychiatric and neurological disorders appears to be bidirectional; indeed, clinically significant depression and anxiety were shown to increase the risk of developing all-cause dementia at a later time[5].

The comorbidity between psychiatric and neurological disorders is partially due to shared genetic and neurobiological mechanisms, including dysfunctions in the serotonergic system and hypothalamic-pituitary-adrenal axis, and inflammatory processes[6]. Common psychosocial risk factors, such as lifestyle choices, social isolation, poverty and discrimination, also contribute to their overlap.

The diagnosis of psychiatric disorders in neurological patients is challenging. For instance, depressive symptoms can be misinterpreted as cognitive or motor impairments associated with neurological conditions or other age-related issues. Comorbidity complicates disease progression, exacerbating disability, reducing quality of life, limiting treatment response, and increasing the risk for overall mortality and suicide[6]. Antidepressants are the primary treatment for depression in neurological patients, but their efficacy remains uncertain, due to a lack of clinical trials in this population, safety concerns related to drug interactions, and tolerability issues. Cognitive and behavioral interventions have demonstrated small-to-moderate improvements in depression and anxiety symptoms among adults with neurological disorders[7]. Other psychological approaches, such as mindfulness-based and expressive-based interventions, have been shown to enhance well-being and alleviate depressive symptoms[8].

However, several barriers hinder the implementation of psychological interventions in patients with neurological disorders. A lack of training and resources can limit accessibility. The lack of flexibility of some interventions, as well as the time and effort required for participation, may not fit patients with severe physical and cognitive disabilities, advanced disease progression, or individual needs and preferences[9].

Considering shared neurobiological mechanisms and comorbidity between disorders, psychiatric and neurologic disciplines should increase collaboration by focusing on six major common goals[10]:

- *Adopting a life-course approach and strengthening prevention*. A life-course approach to Integral Brain Health (cerebral, mental and social health) is essential for improving outcomes, reducing disease burden, and fostering resilience in affected individuals. This entails the promotion of healthy lifestyles and the implementation of early preventive interventions at key life stages, such as pregnancy, childhood/adolescence, career initiation and retirement, where individuals may be more receptive to adopting brain-healthy habits[11,12].
- *Promoting Integral Brain Health in the workplace and the com-*

munity. The growing aging population, coupled with a declining birth rate, is leading to a shrinking workforce and increasing pressure on employers to improve workplace mental health conditions. Fostering Integral Brain Health among employees becomes vital to equip them for the high demands of modern economy, characterized by automatic production processes, the use of artificial intelligence, and the need to be efficient within limited work time. Workplace and community-based healthy lifestyle interventions promoting Integral Brain Health should go hand in hand.

- *Defying stigma and discrimination, and increasing awareness*. People with psychiatric and neurological disorders often experience stigma and discrimination due to persistent misconceptions and negative societal attitudes. Addressing this stigma is fundamental to success of preventive interventions and to grant universal access to treatment and care. This requires increasing knowledge and awareness at all levels of society. It is crucial to understand that our future depends on healthy brains, and that everyone has a role in developing, maintaining and enhancing Integral Brain Health for all.
- *Fostering interdisciplinary research and joint education curricula*. Timely diagnosis and care for psychiatric and neurological disorders and their comorbidities require improved epidemiological studies and evidence-based treatments, as well as a more comprehensive and personalized approach to patient care. There is a pressing need to deepen our understanding of the neurobiological mechanisms of these conditions, study biomarkers for early diagnosis and targeted treatments, and develop evidence-based and practical guidelines for managing comorbidities. While specialized research and care will always be necessary, strengthening interdisciplinary collaborations and integrating neurological, mental and social health programs will reduce redundancy and enhance impact.
- *Closing the treatment gap*. A significant proportion of individuals with psychiatric and neurological disorders lack access to timely and effective treatment. Integrating research in real-world implementation can enhance treatment accessibility and adherence. In low-resource settings, building capacity within non-specialized health care services and training general health care workers to provide essential neurological and psychiatric services is critical. Additionally, digital health solutions may expand access to care, particularly in underserved regions.
- *Advocating for Integral Brain Health policies*. Collaboration between institutions and organizations, such as the WPA and the World Federation of Neurology, is crucial to secure funding for research and public health initiatives and influence health care policies that support integrated care models.

The integration of psychiatry and neurology is dictated by the advancement of neuroscience and is necessary to ensure that Integral Brain Health is accessible, equitable, and universally supported[13]. Through joint efforts in education, research and policy-making, we can build a future where mental and neurological well-being are prioritized across all stages of life, ensuring lifelong brain health for all.

Danuta Wasserman[1,2], Miriam Iosue[2], Abolfazl Avan[3], Vladimir Hachinski[3]

[1]President, World Psychiatric Association; [2]National Centre for Suicide Research and Prevention of Mental Ill-Health, Karolinska Institutet, Stockholm, Sweden; [3]Robarts Research Institute, Department of Clinical Neurological Sciences, and Department of Epidemiology and Biostatistics, Schulich School of Medicine and Dentistry, University of Western Ontario, London, ON, Canada

1. Hachinski V. Alzheimer's Dement 2023;19:3226-30.
2. Institute for Health Metrics and Evaluation. Brain health atlas. https://brainhealthatlas.org.
3. World Health Organization. Life expectancy and healthy life expectancy. www.who.int.
4. Gold SM, Köhler-Forsberg O, Moss-Morris R et al. Nat Rev Dis Prim 2020;6:69.
5. Kuring JK, Mathias JL, Ward L. J Affect Disord 2020;274:247-61.
6. Ibanez A, Zimmer ER. Nat Ment Health 2023;1:441-3.
7. Gandy M, Woldhuis T, Wu W et al. Psychol Med 2024;54:3237-50.
8. Lai ST, Lim KS, Low WY et al. Clin Neuropsychol 2019;33:490-518.
9. Pinto C, Geraghty AWA, McLoughlin C et al. Heal Psychol Rev 2023;17:416-38.
10. Wasserman D. World Psychiatry 2024;23:165-6.
11. Wasserman D. World Psychiatry 2023;22:343-4.
12. Wasserman D, Arango C, Fiorillo A et al. World Psychiatry 2023;22:488-9.
13. Wasserman D. World Psychiatry 2024;23:302-3.

DOI:10.1002/wps.21367

Advancing global mental health through education

In 2024, the WPA launched *Education and Psychiatry*, an official e-journal representing the organization's efforts to strengthen psychiatric education and promote global knowledge exchange. This initiative aligns with the strategic priorities outlined in the WPA Action Plan 2023-2026[1-3], and is further supported by the WPA's Blueprint for Advancing Psychiatric Education and Scientific Publications[4]. The launch of *Education and Psychiatry* responds to pressing international needs: rising global mental health demands, critical shortages in the psychiatric workforce, and the call for accessible, culturally relevant, educational platforms.

Led by the WPA Committee on Education and Scientific Publications, *Education and Psychiatry* is a peer-reviewed e-journal dedicated to innovations in psychiatric education and international collaboration. Since the publication of its first issue in June 2024, the journal has attracted submissions from all continents, showcasing the diversity of psychiatric training systems, pedagogical and adult learning approaches, and professional challenges across the globe.

The journal is grounded in the belief that educational transformation is essential to addressing the global mental health workforce crisis – a priority highlighted in the WPA's strategic planning documents[1,4], and the WPA Global Study on Psychiatric Training[5]. It promotes scalable and equitable models of psychiatric education, especially encouraging contributions from under-represented

colleagues, including those in low- and middle-income countries, where publishing remains more challenging[4].

The scope of *Education and Psychiatry* encompasses diverse content types, including country reports, special issue interviews with global leaders, research updates, digital education innovations, and policy-oriented perspectives. The journal highlights contributions across psychiatry subspecialties and educational methodologies, with thematic foci, such as artificial intelligence in psychiatric training and postgraduate curriculum reform.

Thematic issues have honoured prominent figures, such as Profs. N. Sartorius and A. Okasha, both of whom offered important historical reflections, insights into present-day events, and visions for the future of psychiatry and mental health. These interviews were published in celebration of their 90th birthdays. The articles, released ahead of print, have already generated thousands of reactions and feedbacks on social media, reflecting strong global engagement with the journal's content.

Looking ahead, the editorial team is actively collaborating with the WPA Education Portal to enhance the journal's multimedia capacity and expand access to educational content globally[4]. The WPA Education Portal, a well-established online platform, offers a wide range of freely accessible learning resources – including modular courses, webinars, clinical toolkits, and e-handbooks – designed to support psychiatric education and professional development across diverse settings. It is currently undergoing comprehensive content and information technology upgrades to improve functionality, expand multilingual access, and deliver a more user-friendly and efficient learning experience.

Education and Psychiatry is more than an e-journal; it is a collaborative project that embodies the WPA's commitment to equity, innovation, and educational excellence in global mental health. As a collective platform for sharing expertise and fostering systemic change, it plays a vital role in shaping the future of psychiatric education worldwide.

We invite all WPA Member Societies, Scientific Sections, and professionals to contribute to and engage with this evolving community of learning.

Norbert Skokauskas[1,2], Gary Chaimowitz[3,4], Dina Aly El-Gabry[4-6], Andrea Fiorillo[4,7], Anusha Lachman[4,8], Angeles Lopez Geist[4,9], Paul Robertson[4,10], Hee Jeong Yoo[4,11], Bennett L. Leventhal[4,12]
[1]WPA Secretary for Education and Scientific Publications; [2]Norwegian University of Science and Technology, Trondheim, Norway; [3]McMaster University, Hamilton, ON, Canada; [4]WPA Committee on Education and Scientific Publications; [5]Ain Shams University, Cairo, Egypt; [6]Department of Psychiatry, College of Medicine and Health Sciences, United Arab Emirates University, Abu Dhabi, United Arab Emirates; [7]University of Campania "Luigi Vanvitelli", Naples, Italy; [8]Stellenbosch University, Stellenbosch, South Africa; [9]Association of Argentinian Psychiatrists; [10]University of Melbourne, Melbourne, VIC, Australia; [11]Seoul National University College of Medicine, Seoul, South Korea; [12]University of Chicago, Chicago, IL, USA

1. Wasserman D. World Psychiatry 2023;22:343-4.
2. Wasserman D, Arango C, Fiorillo A et al. World Psychiatry 2023;22:488-9.
3. Wasserman D. World Psychiatry 2024;23:165-6.
4. Skokauskas N, Chaimowitz G, El-Gabry DA et al. World Psychiatry 2024;23:304-6.
5. Skokauskas N, Chaimowitz G, El-Gabry DA et al. World Psychiatry 2025;24:147-8.

DOI:10.1002/wps.21368

Recent activities of the WPA Scientific Sections

The first WPA Scientific Sections were established in 1961. They are now 66, encompassing virtually all the various sub-specialties of psychiatry. Their activities are regulated by the WPA Statutes, By-Laws and Manual of Procedures, according to which their mandate includes: organization of scientific meetings, including symposia at WPA congresses and co-sponsored meetings, on topics within their expertise; development of educational programs, guidelines and related scientific publications; development of proposals for adoption as WPA consensus and position statements; promotion, conduct and facilitation of international collaborative research activities; development of programs in consultation with other Scientific Sections and promotion of intersectional activities.

On the WPA website, there is a separate page for WPA Scientific Sections, where all interested people can follow their activities, considering to join one of them. Previous WPA Secretaries for Sections have done a great work in this direction[1], but further effort is needed in order to have the multiple activities of the Sections regularly described in that page.

In addition, WPA Scientific Sections are welcome to publish their own journals or have links with international ones. Currently, journals which to various extent are linked to WPA Scientific Sections include *Journal of Affective Disorders, Psychopathology, Academic Psychiatry, History of Psychiatry, Personality and Mental Health, Journal of Mental Health Policy and Economics, International Journal of Mental Health, Journal of Intellectual Disability Research, Activitas Nervosa Superior, Psychiatry in General Practice, Transcultural Psychiatry*, and *Archives of Women's Mental Health*.

The activities of Sections are regularly discussed in meetings of the WPA Standing Committee on Scientific Sections. Below a non-exhaustive summary of some recent activities is provided.

Two European Union Horizon research projects are featuring the participation of several WPA Scientific Sections[2]: the PSY-PGx Consortium, focusing on the implementation of pharmacogenetics in psychiatry, and the Psych-STRATA network, aimed at the identification of biological and clinical markers predicting resistance to pharmacological treatment.

Guided by the WPA Action Plan 2023-2026[3-5], several educational initiatives are being implemented in the areas of perinatal psychiatry and infant mental health[6], exercise and sports psychiatry[7], mental health care for migrants and refugees[8], personalized psychiatry[9], addictive disorders[10], anxiety and obsessive-compulsive disorders[11], the relationship between physical and mental health[12], intellectual disabilities[13], early intervention in psychosis[14], climate

change and mental health, advocacy and public engagement, psychiatric epidemiology, quality assurance in psychiatry, psychiatry in primary care, urban mental health, military psychiatry, human rights of older persons, evolutionary psychiatry, genomics in clinical practice, digital mental health, and childhood attention-deficit/hyperactivity disorder (ADHD).

Partnerships are being fostered by WPA Scientific Sections with a range of international organizations, including the World Health Organization, the European Psychiatric Association, the European Union of Medical Specialists (UEMS), and the European College of Neuropsychopharmacology (ECNP)[11,15].

WPA Scientific Sections are actively involved in World Congresses of Psychiatry. In the 24th World Congress of Psychiatry, held in Mexico City in November 2024, fourteen State-of-the-Art Symposia and 29 Regular Symposia originating from the Sections were included in the scientific programme.

Several Scientific Sections are involved in the activities of the WPA Advisory Committee on Responses to Emergencies (ACRE), which was funded in May 2020 to provide help in coping with mental health consequences of natural and human-made disasters[16].

The WPA aims to contribute to the achievement of the third United Nations (UN) Sustainable Development Goal ("Ensure healthy lives and promote well-being for all at all ages"), and the work of the Scientific Sections is being crucial in this respect. This requires a cooperation with a variety of partners, also beyond the field of public health, targeting areas such as climate action, labour, housing and education.

Promotion of healthy lifestyles and suicide prevention at the population level, and enhancement of the mental and physical well-being in psychiatric patients and staff, are among the main priorities highlighted in the WPA Action Plan 2023-2026[3-5], and are being a main component of the activities of many WPA Scientific Sections.

Finally, some Sections still need to be revitalized, and we are trying to create the conditions for this, under the guidance of the WPA Executive Committee and the leaders and members of all the Sections.

Armen Soghoyan
WPA Secretary for Scientific Sections

1. Schulze TG. World Psychiatry 2023;22:245.
2. Soghoyan A. World Psychiatry 2024;23:306-7.
3. Wasserman D. World Psychiatry 2023;22:343-4.
4. Wasserman D, Arango C, Fiorillo A et al. World Psychiatry 2023;22:488-9.
5. Wasserman D. World Psychiatry 2024;23:165-6.
6. Keren M, Apter G. World Psychiatry 2024;23:168-9.
7. Gorczynski P, Baron D. World Psychiatry 2024;23:169-70.
8. Schouler-Ocak M, Kastrup MC, Kuey L et al. World Psychiatry 2024;23:457-8.
9. Perna G, Caldirola D, Schatzberg AF et al. World Psychiatry 2024;23:460-1.
10. Elkholy H, Bhad R, Ekhtiari H et al. World Psychiatry 2024;23:459-60.
11. Fineberg NA, Stein DJ, Domschke K et al. World Psychiatry 2024;23:456-7.
12. Fiorillo A, de Girolamo G, Simunovic IF et al. World Psychiatry 2023;22:169-70.
13. Bertelli MO, Bianco A, Salvador-Carulla L et al. World Psychiatry 2023;22:168-9.
14. Singh SP, Javed A, Thara R et al. World Psychiatry 2023;22:489-90.
15. Sampogna G, Al-Taiar H, Baessler F et al. World Psychiatry 2024;23:461-2.
16. Wasserman D, Herrman H, Javed A et al. World Psychiatry 2023;22:166-7.

DOI:10.1002/wps.21369

The WPA Section on Genetics in Psychiatry: scientific progress and clinical translation

The WPA Section on Genetics in Psychiatry focuses primarily on the role of genetics in the etiology and treatment of psychiatric disorders, with a particular interest in translating these findings to the clinic. The members of the Section are a mixture of active clinicians and researchers, many of whom are also regular members of the International Society of Psychiatric Genetics. Compared to that larger Society, the WPA Section is more focused on potential uses of genetics in clinical settings and on cultivating a global diverse community consistent with the goals of the broader WPA parent organization[1].

The Section meets every month to discuss clinically relevant topics and provide a global link between scientists, people with lived experience, and other mental health practitioners. As part of its mandate to promote the growth of a community of clinically oriented scientists and clinicians, the Section has organized several symposia and courses at WPA congresses across three continents (in Thailand, Malta, India, United Arab Emirates, Poland, Austria and Mexico) to inform local scientists and clinicians about recent advances in psychiatric genetics, with an emphasis on how these can be translated into their clinical practice. In preparation for the 2023 World Congress of Psychiatry in Vienna, the Section published an educational booklet to inform clinicians about psychiatric genetics and the potential value of pharmacogenetic testing[2].

Over the past decade, scientific progress has greatly advanced our understanding of how both common and rare genetic variations contribute to the etiology of major psychiatric disorders, including schizophrenia, bipolar disorder and major depression. Strong evidence now indicates that disease risk is influenced by hundreds, potentially thousands, of common alleles with small effect sizes that likely make up the bulk of the genetic risk. Although these small effect sizes pose challenges for study, they can be aggregated into polygenic risk scores, which are increasingly explored in clinical research as potential predictors of psychopathology and treatment response. Genome-wide association studies (GWAS) of almost all major psychiatric disorders have now identified tens to hundreds of significant loci[3]. The main challenge facing the field is how to progress from statistical associations to a biological understanding of how risk gives rise to the pathophysiology experienced by our patients.

With the advent of high-throughput and increasingly cost-effective means of performing whole exome and whole genome se-

quencing, large-scale sequencing studies are now increasingly being performed, with a main focus on novel variant discovery[4,5]. Rare variants offer a distinct advantage by potentially pinpointing specific genes (rather than broader loci identified by GWAS) and revealing the direction of effect. While these studies are just beginning to uncover these variants, efforts are underway worldwide to increase sample size and diversity in order to enhance efforts for equitable discovery.

One specific type of rare variants already used in clinical practice includes rare duplications or deletions known as copy-number variants (CNVs). These variants are often pleiotropic (i.e., associated with a range of phenotypes), but are particularly relevant in the study of intellectual disability and severe autism spectrum disorders. The clinical relevance of CNV testing for neurodevelopmental disorders in children has been recently reviewed[6], and surveys are being conducted to evaluate the knowledge of best practices among professionals.

Another major focus of the Section is the potential of pharmacogenetics to improve clinical care. There is now robust evidence that variations in drug-metabolizing enzymes (in particular *CYP2C19* and *CYP2D6*) are associated with levels of several antidepressant and antipsychotic medications. However, the role of such variation in clinical response and overall tolerability needs to be explored in larger, more diverse studies across the world. Our Section is therefore proud to be participating in the PSY-PGx project[7], a global non-industry-sponsored study funded by the European Union Horizon 2020 initiative that will test the effectiveness of pharmacogenetic testing in a broad range of psychiatric disorders.

A key priority of the Section is advancing diversity and inclusion in genetic research by fostering the inclusion of a broader range of populations. Ethically, it is essential to address health disparities by ensuring that the various populations are adequately represented. Scientifically, incorporating more diverse populations in GWAS can lead to a more comprehensive understanding of the genetic architecture of psychiatric disorders, and can improve the accuracy of polygenic risk scores. In pharmacogenomics, research has shown that actionable alleles vary significantly across populations, highlighting the need for more trans-ancestry studies. So, this approach is crucial not only for ethical reasons of equality, inclusion and representation, but also for advancing scientific discovery and enhancing clinical applications.

Finally, we aim to ensure that research and implementation in psychiatric genetics also address critical topics such as social justice, stigma reduction, autonomous decision-making, the right to know or not know, and data protection. The WPA provides an ideal platform to evaluate, discuss and monitor these issues globally, bringing together diverse perspectives that matter to those living with psychiatric disorders. We enthusiastically invite scientists, clinicians and patients who share these goals to join us in this endeavor.

Urs Heilbronner[1,2], Marcella Rietschel[1,3], Thomas G. Schulze[1,2,4-6], Roos van Westrhenen[1,7-9], Fernando S. Goes[1,5], Smita N. Deshpande[1,9]
[1]WPA Section on Genetics in Psychiatry; [2]Institute of Psychiatric Phenomics and Genomics (IPPG), LMU University Hospital, Ludwig Maximilian University of Munich, Munich, Germany; [3]Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany; [4]Department of Psychiatry and Behavioral Sciences, SUNY Upstate Medical University, Syracuse, NY, USA; [5]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [6]German Center for Mental Health (DZPG), partner site Munich/Augsburg, Munich, Germany; [7]Department of Psychiatry, Parnassia Groep, Amsterdam, The Netherlands; [8]Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; [9]St. John's National Academy of Health Sciences, Bangalore, India

F.S. Goes and S.N. Deshpande contributed equally to this paper.

1. Wasserman D. World Psychiatry 2024;23:165-6.
2. van Westrhenen R, Goes F, Deshpande S et al. Highlights of the recent activities of the WPA Section Genetics in Psychiatry. www.wpanet.org.
3. Andreassen OA, Hindley GFL, Frei O et al. World Psychiatry 2023;22:4-24.
4. Liu D, Meyer D, Fennessy B et al. Nat Genet 2023;55:369-76.
5. Palmer DS, Howrigan DP, Chapman SB et al. Nat Genet 2022;54:541-7.
6. Shimelis H, Oetjens MT, Walsh LK et al. Am J Psychiatry 2023;180:65-72.
7. van Westrhenen R, Young AH, Heilbronner U et al. World Psychiatry 2025;24:141-2.

DOI:10.1002/wps.21371

Evolutionary psychiatry and the activities of the relevant WPA Section

Evolutionary psychiatry has made unique contributions to key areas of psychiatry at the conceptual level, with the aim to foster understanding not only of *what* happens, but also of *why* problems arise. For example, Wakefield's concept of harmful dysfunction[1], the application of Tinbergen's four causal domains (also known as Tinbergen's four questions) and Nesse's "pathways" for the persistence of disease and disorder[2] highlight novel ways of thinking about psychiatric conditions. Without evolutionary science, such "why" questions are not conceivable. Mental disorders can just be carefully described. The evolutionary approach enables psychiatry to move from the descriptive to an explanatory phase, an important step in the development of any scientific discipline.

Evolutionary psychiatry has been concisely defined as the subfield of evolutionary medicine that uses the basic science of evolutionary biology to better understand and treat mental disorder[3]. However, it is a fundamentally cross-disciplinary field with vital contributions from evolutionary anthropology, evolutionary psychology, and behavioral genetics, as well as many other disciplines.

The evolutionary approach teaches us that selection shapes vulnerability to disease and disorder and not disorders themselves. It also helps us recognize that selection (on average) shapes functional systems and not dysfunctional ones. But functional systems, whether biologically evolved or man-made, can malfunction under certain conditions. Also, biological systems frequently

have multiple, overlapping functions, and this applies particularly to neurobiological systems, whose functions often remain poorly understood. Several pathways may be implicated concurrently or sequentially in the causation of a given mental disorder.

Selection is unable to eliminate all harmful mutations and can be too slow to respond to rapidly changing environments, which generates states of evolutionary mismatch. The concept of "mismatch" is arguably one of the most important insights in evolutionary medicine, and is crucial for understanding and explaining the existence of "diseases of civilization", such as obesity, metabolic syndrome, type 2 diabetes, eating disorders[4], attention-deficit/hyperactivity disorder (ADHD), postpartum depression, and many others. Evolutionary mismatch occurs when the environment changes too rapidly for selection to be able to track it, resulting in residual traits that are no longer suited to the new environmental conditions. The insights gained through recognition of evolutionary mismatch often point to the need for a public health response, rather than blaming the individual.

It is important to recognize that selection inevitably involves trade-offs. Thus, increasing the potency of one trait is often at the expense of worsening performance of another. For example, reducing the threshold for environmental risk avoidance can result in a greater risk of anxiety disorders, whereas raising the threshold can lead to hypophobia and dangerous risk-taking. Additionally, over-activation of useful emotional defences (e.g., mood states and anxiety) can result in harmful outcomes, leading to what have been termed defence activation disorders (e.g., anxiety and depressive disorders).

Evolutionary psychiatry can also help make sense of the genetics of mental disorder. Taking an evolutionary perspective, one comes to recognize that viewing the human genome as a static "blueprint" for the human phenotype is both erroneous and misleading. At the population level and over multiple generations, the frequency of genes is in a state of continuous change, with some genes increasing in frequency while others decreasing or being eliminated completely as a result of positive and negative selection pressures (natural, sexual, social) as well as drift (random, chance events).

This raises the question of why apparently harmful or disease-causing genes exist and persist in the human gene pool. We suggest that this question can only be addressed by taking an evolutionary perspective. Non-evolutionary approaches simply note the existence of such genes and study their effects. Evolutionists recognize that the human genome is a historical record of past selection pressures. Variation is the rule rather than the exception and, most importantly, there is no such thing as a single normative human genome. Variation truly is the spice of life and supports the current patient-led movements to recognize that, for instance, high-functioning autism spectrum disorder and ADHD may be conceptualized as normative variations and therefore viewed as manifestations of neurodiversity rather than pathological conditions.

With the exception of *de novo* harmful genes which arise from mutations in the parental germ-line, all other disease-causing genes have been subject to selection pressures of some form. Some harmful genes that are compatible with survival and reproduction are subject to purifying selection over many generations, while others may be eliminated more quickly. However, those that persist within the population over numerous generations may be subject to a process known as balancing selection. This takes place where there is a trade-off between the positive and the negative effects of a given genetic variant, thus maintaining that variant at a more or less steady level in the population. Examples include genes that are advantageous in the heterozygous state but harmful in the homozygous state (e.g., sickle cell anaemia); genes that are subject to frequency dependent selection (i.e., potentially advantageous when low in frequency in the population, but disadvantageous at higher frequencies – e.g. psychopathy); and genes with pleiotropic effects (i.e., multiple effects, some advantageous and others not at different life stages).

Many have considered the biopsychosocial model originally proposed by Engel[5] to be in need of review and updating. One way of achieving this could be the incorporation of an evolutionary dimension through subjecting Engel's three levels (the bio, psycho and social) to Tinbergen's four causal domains (mechanism, development, phylogeny and function). We argue that this "evo-biopsychosocial model" may provide a more coherent, scientifically complete and philosophically sound model of mental disorder (and of disease and disorder generally)[6].

The WPA Section on Evolutionary Psychiatry is being active in a number of different ways. Over the last five years, we have set up a series of webinars which are recorded with a live audience and then placed on our YouTube channel. We have participated with three symposia in the Malta WPA Conference in 2022 (one jointly with the Section on Public Policy and Psychiatry) and with two symposia in the Vienna World Congress of Psychiatry in 2023 (jointly with the Sections on Perinatal Psychiatry and Infant Mental Health, and on Public Policy and Psychiatry). Our Section also collaborates closely with the two extant special interest groups on evolutionary psychiatry in the UK and Ireland. In addition, our members have published a number of books that serve as standard texts in the field. These include *Good Reasons for Bad Feelings*[2], *Evolutionary Psychiatry: Current Perspectives on Evolution and Mental Health*[7], and *The Evolutionary Roots of Human Brain Diseases*[8].

Riadh Abed[1,2], Annie Swanepoel[1,3], Henry O'Connell[1,4,5]
[1]WPA Section on Evolutionary Psychiatry; [2]Evolutionary Psychiatry Special Interest Group, Royal College of Psychiatrists, UK; [3]North East London NHS Foundation Trust, Chelmsford, UK; [4]School of Medicine, University of Limerick, Limerick, Ireland; [5]Evolution and Psychiatry Special Interest Group, College of Psychiatrists of Ireland

1. Wakefield JC. World Psychiatry 2007;6:149-56.
2. Nesse RM. Good reasons for bad feelings. London: Allen Lane, 2019.
3. Nesse RM. World Psychiatry 2023;22:177-202.
4. Abed R. Br J Med Psychol 1998;71:525-47.
5. Engel G. Am J Psychiatry 1980;137:535-44.
6. Abed R, Hunt A, St. John-Smith P. Br J Psychiatry 2024;225:424-6.
7. Abed R, St. John-Smith P (eds). Evolutionary psychiatry: current perspectives on evolution and mental health. Cambridge: Cambridge University Press, 2022.
8. Diederich NJ, Brüne M, Amuntset K et al (eds). The evolutionary roots of human brain diseases. New York: Oxford University Press, 2024.

DOI:10.1002/wps.21370

Report from the WPA Section on Military Psychiatry

Ongoing armed conflicts highlight the profound psychological and behavioral effects of war and mass violence on military personnel, impacting their well-being and long-term mental health. With tens of millions active service members and veterans worldwide, addressing military mental health is a global priority.

Military personnel not only experience combat stressors, but also civilian atrocities, military sexual trauma, family separation, reintegration difficulties; and heightened risks of post-traumatic stress disorder (PTSD), depression, substance abuse, and suicide. The evolving nature of warfare – e.g., autonomous weapons, cyber threats, and remote operations – further compounds mental health challenges, often leaving remote operators and cyber personnel with unacknowledged psychological burdens. Understanding these risks across cultures and over time is essential for developing effective policies, programs and interventions to safeguard military and veteran mental health, ultimately strengthening global health security.

Besides military-related risk factors for negative mental health consequences of exposure to trauma, there are unique protective factors related to military service that can mitigate the effects of exposure to war stressors. Loyalty and unit cohesion among members may provide emotional support and a sense of belonging. An immediate access to support services, including counselling and peer support groups, also helps in preventing worsening of mental health. Proper training in stress management, resilience and coping skills, and leadership support may lessen the risk for negative mental health problems in military personnel[1].

Combat and operational stress reactions, the military equivalent of acute stress reaction, often viewed as "normal" reactions to an "abnormal" experience in high-threat situations, can impair well-being and functioning. Promising practices to address these reactions, such as buddy systems (i.e., arrangements in which two individuals are paired for mutual safety in a hazardous situation), have been adapted globally. The armed forces implement combat and operational stress control programs to help service members manage stress, emphasizing proximity, immediacy, expectancy and simplicity to provide support and normalize reactions[2]. Teleconferencing has emerged as a promising solution for managing combat and operational stress reactions, offering accessible and continuous mental health support.

PTSD is a major mental health issue among military personnel and veterans. Its prevalence varies based on deployment experiences and the specific population studied. First-line interventions for PTSD in military personnel have had limited success in preventing the disorder. Trauma-focused psychotherapy, particularly exposure therapy, is the recommended initial treatment, and teleconferencing is suggested when in-person therapy is not feasible[3]. Regardless of treatment setting, the approach should be evidence-based, patient-centered, and culturally appropriate. Leaders must foster an environment that supports help-seeking behaviors and maintains trauma-informed care. A significant subset of individuals with PTSD may develop complex PTSD[4], which requires longer treatment and more diverse interventions to address chronic trauma and disturbances in self-organization[5].

Depression is a significant issue among military personnel and veterans, often co-occurring with PTSD. Among veterans with depression, 36 to 51% also have PTSD[6]. Depression is linked to physical health problems, substance use, and increased suicide risk. Treating depression in military personnel requires a multi-faceted approach, including evidence-based therapies such as cognitive-behavioral and interpersonal psychotherapy, and antidepressants. Teletherapy, peer support, and wellness programs also offer effective, accessible solutions[6].

Alcohol use has been a prominent part of military culture, often used to manage stress during active duty and post-conflict periods. Substance use disorders are common among military personnel, especially veterans dealing with chronic pain, trauma or homelessness, and are linked to increased suicide risk[7]. Despite efforts to address this issue, substance use disorders continue to rise. Treatment should be tailored to military life's unique challenges, including leadership involvement, peer support, and integration with military health care systems for effective recovery.

Military personnel are at high risk for suicide. A meta-analysis shows a prevalence of suicidal ideation and attempts of 11%. Veterans experience higher rates (14%) compared to active-duty personnel (10%), and women are more likely to experience suicidal thoughts[8]. Firearms are commonly used for suicide in the military. Educational programs on mental health and suicide prevention are vital in raising awareness and fostering positive attitudes. A comprehensive suicide prevention approach should address pre-event, event and post-event factors, including risk management and interventions. Policies such as the Brandon Act, which allows service members to request a mental health evaluation, and tailored interventions can help reduce suicide rates.

Mental health stigma is a major barrier to treatment for military personnel and veterans, with 60% not seeking help despite needing it[9]. This stigma can lead to isolation, worsening of mental health, and increased suicide risk. The military's ethos of "service before self" often views mental health struggles as weakness, deterring service members from seeking help. Addressing stigma requires education, training, peer support, and leadership that fosters a help-seeking culture, ultimately improving care utilization and readiness.

After a successful contribution in 2019 to the book *Advances in Psychiatry*, where a group of the WPA Section on Military Psychiatry experts participated with the chapter "Mental health consequences of war conflicts"[2], the Section gave new momentum to professional and scientific work last year by bringing together a working group of experts from around the world to draft the Position Statement on Military Personnel and Veterans' Mental Health[10]. After several months of work, the group submitted the final document for review to the WPA Executive Committee, which approved it in February 2025. This statement provides key updates on global conflicts, unique mental health risks in military roles,

common mental disorders among service members and veterans, and challenges related to stigma and suicide.

The Position Statement on Military Personnel and Veterans' Mental Health offers recommendations across multiple levels to enhance awareness, improve access to mental health care, strengthen leadership and clinical governance, and develop evidence-based policies for prevention, treatment and reintegration. Considering ongoing global conflicts that make the prospect of lasting peace seem remote, the WPA Section on Military Psychiatry remains steadfast in promoting the critical importance of early detection, ensuring access to evidence-based treatments, and fostering a supportive culture that mitigates stigma related to mental health problems within military and veteran communities.

Marina Letica-Crepulja[1,2], Georgios Pagkalos[3], Joshua C. Morganstein[4], Kristina Adorjan[5], Malcolm Hopwood[6]
[1]Department of Psychiatry and Psychological Medicine, Faculty of Medicine, University of Rijeka, Rijeka, Croatia; [2]Department of Psychiatry, Clinical Hospital Center Rijeka, Referral Center for PTSD, Rijeka, Croatia; [3]424 General Military Training Hospital, Thessaloniki, Greece; [4]Center for the Study of Traumatic Stress, Department of Psychiatry, Uniformed Services University of Health Sciences, Bethesda, MD, USA; [5]University Hospital of Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland; [6]University of Melbourne, Melbourne, Australia

The ideas expressed in this paper are those of the authors and do not necessarily reflect those of any agency or branch of the US government or other institution.

1. Benedek DM, Morganstein JC, Holloway HC et al. In: Sadock BJ, Sadock VA, Ruiz P (eds). Kaplan & Sadock's comprehensive textbook of psychiatry. Baltimore: Lippincott, Williams & Wilkins, 2017: 2655-67.
2. Rosanov V, Frančišković T, Marinić I et al. In: Javed A, Fountoulakis KN (eds). Advances in psychiatry, Vol. 4. Cham: Springer, 2019:281-304.
3. US Department of Veterans Affairs. Clinical practice guideline: management of post-traumatic stress, version 4.0. Washington: US Department of Veterans Affairs, 2023.
4. Brewin CR, Atwoli L, Bisson JI et al. World Psychiatry 2025;24:52-60.
5. Letica-Crepulja M, Stevanović A, Protuđer M et al. Eur J Psychotraumatol 2020; 11:1716593.
6. US Department of Veterans Affairs. Management of major depressive disorder. Washington: US Department of Veterans Affairs, 2022.
7. Osborne AK, Wilson-Menzfeld G, McGill G et al. Occup Med 2022;72:313-23.
8. Moradi Y, Dowran B, Sepandi M. BMC Psychiatry 2021;21:510.
9. Sharp ML, Fear NT, Rona RJ et al. Epidemiol Rev 2015;37:144-62.
10. Letica-Crepulja M, Pagkalos G, Morganstein JC et al. Position statement on military personnel and veterans' mental health. www.wpanet.org.

DOI:10.1002/wps.21373

Patient involvement in undergraduate psychiatric education: an international survey

The concept of patient involvement refers to the active participation of individuals with lived experience in roles such as teaching, assessment, or curriculum development. These individuals bring a unique perspective enriching education with their experiential knowledge[1,2]. Patient involvement can vary widely, ranging from being entirely absent, to existing in a limited or partially operational capacity, to being fully active, integrated and sustainable[1]. Previous research has emphasized its potential to enrich students' learning by adding realism and fostering stronger connections between students and patients[3]. It also empowers patients, making them feel valued through their contribution to teaching[4]. Despite these benefits, the implementation of patient involvement strategies remains inconsistent across medical schools worldwide.

The existing literature highlights the positive impact of patient involvement, demonstrating improvements in students' communication skills, empathy, and understanding of patient-centred care [5-8]. However, much of the available research is outdated, country-specific or focused on general medical education rather than addressing the unique needs within psychiatry[9,10]. The lack of international and standardized approaches further underscores the need to assess and improve patient involvement in psychiatry curricula.

To address these gaps, and in line with the WPA Action Plan 2023-2026[11-13], the WPA Section of Early Career Psychiatrists conducted an international online survey of medical schools, in collaboration with the International Federation of Medical Students' Associations (IFMSA). This study aimed to assess the current state of patient involvement in undergraduate psychiatry education globally. Responses were gathered from medical students in 47 countries across six continents. The sample distribution included 23 medical schools from 13 countries in Europe (25.8%); 21 medical schools from 11 countries in Africa (23.6%); 17 medical schools from 10 countries in Asia (19.1%); 18 medical schools from 9 countries in the Americas (20.2%); one medical school from Oceania (1.1%); and 9 medical schools from three countries in transcontinental regions (10.1%). This diverse sample provided a broad perspective on patient involvement in psychiatry education across different regions and economic contexts.

Our findings revealed that patient involvement in undergraduate psychiatry education remains limited in many medical schools. Over half (53.3%) reported no active participation of patients in developing or revising the psychiatry undergraduate curriculum. Furthermore, 62.2% did not involve patients in curriculum quality assurance or as patient-educators, and 63.3% indicated that patients do not play an active role in curriculum assessment. More than half (52.3%) of the institutions lack independent bodies within the medical school or teaching hospital for patients with lived experience of a mental health condition, and 52.2% do not provide financial support for patient-educators.

This study also identified widespread knowledge gaps. Over a quarter (27.7%) were unaware of whether academic staff engaged with patients involved in the undergraduate psychiatry course. Nearly a quarter (23.4%) did not know if patient-educators received educational support, and over a third (35.6%) were uncer-

tain about the availability of counselling and well-being support services for patients involved in the psychiatry undergraduate course. There was also uncertainty about the provision of financial (32.2%) and non-financial (32%) support, the existence of independent bodies (26.7%), and the presence of dedicated administrative personnel to support patient-educators (41.1%).

Interestingly, some of the highest levels of patient involvement were reported in low-income countries. Notable exceptions included Sweden, where patients played a significant role in curriculum design and assessment, and Austria, which reported the existence of full-time administrative personnel dedicated to supporting patient involvement as well as active staff engagement with patients in the undergraduate psychiatry course. Oman adapted facilities and covered all costs associated with patient-educators, while Slovenia provided ongoing well-being support. Bulgaria also offered ongoing training for patient-educators. The highest levels of patient involvement in psychiatry undergraduate education were reported in Peru and Argentina.

Differences in patient involvement were also found. All Ethiopian medical schools in this study reported conducting regular emotional safety and well-being checks during and after the teaching period. Bulgarian medical schools consistently paid salaries to those involved in the educational process. In Hungary, independent bodies within medical schools or teaching hospitals were generally present, albeit with varying levels of involvement. In Romania, there were stark contrasts between institutions: some medical schools reported lack of patient involvement, while others offered continuous well-being support throughout the course, with high levels of patient involvement in curriculum assessment and design, active staff engagement with patients, and financial support for patient-educators.

These findings highlight the urgent need to standardize patient involvement in psychiatry education globally. Standardization would help establish benchmarks, promote best practices, and ensure that patient involvement is not only present but effectively integrated into educational strategies. Addressing challenges, such as the lack of financial and administrative support, is essential for ensuring sustainability and equity. Collaborative efforts from medical schools, health care institutions, and policy makers are needed to ensure that patient involvement becomes a more integral and effective component of psychiatry undergraduate education worldwide.

This study, which benefited from patients' advice in both its design and the interpretation of findings, provides a foundation for further research. Patient involvement holds immense potential to transform psychiatry education by fostering empathy, communication skills, and patient-centred care among future doctors. However, inconsistent implementation and inadequate support hinder its widespread adoption.

Mariana Pinto da Costa[1-3], Tommaso Squeri[4], Lucía Péréz Gómez[5], Bernardo Flôr-Rodrigues[6]

[1]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; [2]Institute of Biomedical Sciences Abel Salazar, University of Porto, Porto, Portugal; [3]WPA Section of Early Career Psychiatrists; [4]Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK; [5]International Federation of Medical Students' Associations; [6]Faculty of Medicine, University of Coimbra, Coimbra, Portugal

1. Towle A, Bainbridge L, Godolphin W et al. Med Educ 2010;44:64-74.
2. Towle A, Farrell C, Gaines ME et al. Int J Health Gov 2016;21:18-25.
3. Henriksen AH, Ringsted C. Med Educ 2011;45:913-9.
4. Walters K, Buszewicz M, Russell J et al. BMJ 2003;326:740.
5. Jha V, Quinton ND, Bekker HL et al. Med Educ 2009;43:10-20.
6. Jha V, Setna Z, Al-Hity A et al. Med Educ 2010;44:347-57.
7. Morgan A, Jones D. Med Teach 2009;31:82-95.
8. Stacy R, Spencer J. Med Educ 1999;33:688-94.
9. Galderisi S, Appelbaum PS, Gill N et al. World Psychiatry 2024;23:364-86.
10. Royal College of Psychiatrists. Learning from experience. Working in collaboration with people with lived experience to deliver psychiatric education. London: Royal College of Psychiatrists, 2021.
11. Wasserman D. World Psychiatry 2023;22:343-4.
12. Wasserman D, Arango C, Fiorillo A et al. World Psychiatry 2023;22:488-9.
13. Wasserman D. World Psychiatry 2024;23:165-6.

DOI:10.1002/wps.21372

Acknowledgement
This publication has been supported by an unrestricted educational grant from Lundbeck Italia S.p.A. and Otsuka Pharmaceutical Italy S.r.l. (OPIT) which is hereby gratefully acknowledged. They have not had any influence on content, review or editorial control in relation to this issue.


FSC
www.fsc.org
MIX
Board | Supporting
responsible forestry
FSC® C020438