World Psychiatry

OFFICIAL JOURNAL OF THE WORLD PSYCHIATRIC ASSOCIATION (WPA)

Volume 20, Number 2 *June 2021*

WILEY

IMPACT FACTOR: 40.595

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 145, spanning 121 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 70 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Research Reports containing unpublished data are welcome for submission to the journal.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: majmario@tin.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 PubMed system (http://www.pubmedcentral.nih.gov/tocrender.fcgi?journal=297&action=archive).

Rediscovering the mental health of populations

The principles of prevention espoused by G. Rose[1] have underpinned many modern successes in health care. In areas such as cardiometabolic diseases, injuries and violence, and substance abuse, shifting the community distribution of risk factors has become the primary strategy. The ensuing reductions in disease burden have been striking.

Psychiatry remains an outlier. Over decades, the quality of clinical care has been improved, greater funding has been attracted, more and better trained mental health professionals have been grown, and the governance of mental health care has been upgraded[2]. However, the emphasis in recent initiatives in high-income countries has been overwhelmingly a further extension of treatment: early clinical intervention has been the dominant initiative taken up in government investments into the mental health of young people[3].

Yet, this continuing expansion of government expenditure, prescribing of antidepressants and availability of psychological services has still not been accompanied by reductions in the prevalence of common mental disorders[3]. While it remains possible that this in part reflects a continuing failure to scale minimally-sufficient treatments, the evidence from other fields of medicine suggests that a more likely explanation is the lack of scalable risk-focused prevention strategies.

This failure to embrace population-based approaches to prevention in psychiatry is understandable. Most clinicians find the endorsement of population perspectives difficult. For them, the individual is the unit of study[1]. For psychiatry, the opacity of pathophysiological processes has supported the tendency to focus on interventions directed at the individual. Recent excitement about progress in genetics and neuroscience has reinforced this tendency, with both major research funding agencies and the pharmaceutical industry emphasizing the individual over the social context.

In this scenario, the paper by Fusar-Poli et al in this issue of the journal[4] raises questions around the optimal strategies for prevention in psychiatry. The overwhelming emphasis to date across common mental disorders, psychosis and bipolar disorder has been on individuals at high risk by virtue of early clinical symptoms or genetic predisposition. These selective and indicated approaches to prevention have targeted subjects at the tail of the distribution, with an aim of reducing the likelihood of transition to clinical caseness. However, this emphasis on individuals has been accompanied by a failure to address structural and social determinants.

E. Durkheim's work, well over a century ago, drew the conclusion that suicide rates are stable and distinctive characteristics of populations. He viewed suicide as a collective phenomenon in which personal factors are less important than the social context. Similarly, strategies focused on the social, economic and regulatory context that bring a reduction in average alcohol consumption have been far more successful in reducing levels of alcohol use disorders than individually targeted interventions[5]. This principle that actions to reduce modest risks in a large group will generate greater benefits than targeting conspicuous risks in a small number should guide the prevention of mental disorders.

One challenge is that most risks for mental disorders lie outside the direct influence of the health sector. For young people, social determinants of mental health derive from inequitable gender norms, shifts in family structure and function, culture and religion, economic development and its consequences, digital technology, urbanization and planetary change. These social and structural determinants shape peer, family and community relationships, accessibility of service systems, the likelihood of experiencing major external events, as well as risks related to lifestyle and individual behaviour. For mental disorders, as for the physiological processes underpinning physical health, there are also sensitive periods in which risks are more likely to become embedded and when prevention will be more effective.

The COVID-19 pandemic illustrates the influence of social and structural factors on the mental health of all age groups, but particularly the young. It also illustrates areas where psychiatry should be acting. The effects of lifestyle risk factors for mental disorders, including physical inactivity, screen time, irregular sleep and poor diets, have been enhanced. Even more profound have been the shifts in relationships, with disruption to friendships and peer interactions, heightened worries about and sometimes conflict with family members, confinement to home and loss of the social milieu of schools, including extracurricular activities.

In taking prevention in psychiatry forward, there are further lessons to be drawn from other areas of medicine[1]. Epidemiology remains the underpinning discipline of public health, and, for psychiatry, epidemiology should adopt both life-course and population perspectives. However, psychiatric epidemiology remains in a parlous state, particularly for children and young people. Global coverage for even basic estimates of prevalence lies under seven percent, with rates in low- and middle-income countries substantially lower, and 124 countries having absolutely no data[6]. Coverage of risk factors is even weaker.

As noted by Fusar-Poli et al, a life-course perspective on mental health is essential[4]. Yet, a life-course perspective would ideally extend across generations, given that familial clustering is the clearest of all risk factors. Beyond genetics, there are malleable intergenerational risk factors for mental disorders, ranging from the biological (e.g., epigenetic) through to the structural (e.g., inequitable gender norms), including those risks that become embedded prior to conception[7]. Longer-term perspectives derived from prospective life-course studies have the potential to guide prevention research and policy, particularly when combined with powerful new analytic tools for causal inference.

Recent intervention trials provide grounds for optimism. Schools will be one important context for prevention. Children and young people spend close to half their waking hours in school and education. Policy-makers increasingly understand that poor student mental health affects learning and academic

achievement. There are now examples from both high- and low-resource settings that interventions promoting a positive school social climate and reducing bullying can substantially reduce symptoms of common mental disorder[8]. Other promising platforms include those based in local communities (e.g., girls clubs) and the new social environments created by digital media.

Interventions well beyond those traditionally regarded as the focus for prevention of mental disorders will also be important. Cash transfers have been widely adopted by governments in other areas of health and social policy, and seem to bring reductions in symptoms of mental disorder and promotion of well-being in low-resource settings where psychological interventions based on cognitive behaviour therapy have little or no effect[9]. Such findings suggest the value of inclusion of mental health into trials of non-mental health interventions.

The dramatic deterioration in community mental health during the COVID-19 pandemic heightens the imperative for psychiatry to shift beyond its comfort zone of the individual patient, and engage with the social, structural and political determinants of mental health.

George C. Patton[1,2], Monika Raniti[1,2], Nicola Reavley[3]
[1]Centre for Adolescent Health, Murdoch Children's Research Institute, Parkville, VIC, Australia; [2]Department of Paediatrics, Melbourne Medical School, University of Melbourne, Melbourne, VIC, Australia; [3]Melbourne School of Population and Global Health, University of Melbourne, Melbourne, VIC, Australia

1. Rose G. The strategy of preventive medicine. Oxford: Oxford University Press, 1992.
2. Priebe S, Burns T, Craig TK. Br J Psychiatry 2013;202:319-20.
3. Jorm AF, Patten SB, Brugha TS et al. World Psychiatry 2017;16:90-9.
4. Fusar-Poli P, Correll CU, Arango C et al. World Psychiatry 2021;20:200-21.
5. Rose G. In: Williams P, Wilkinson G, Rawnsley K (eds). The scope of epidemiological psychiatry. London: Routledge, 1989:77-85.
6. Erskine HE, Baxter AJ, Patton G et al. Epidemiol Psychiatr Sci 2017;26:395-402.
7. Patton G, Olsson C, Skirbekk V et al. Nature 2018;554:458-66.
8. Shinde S, Weiss HA, Varghese B et al. Lancet 2018;392:2465-77.
9. McGuire J, Kaiser C, Bach-Mortensen A. https://doi.org/10.31235/osf.io/ydr54.

DOI:10.1002/wps.20842

Enabling a youth- and mental health-sensitive greener post-pandemic recovery

International bodies such as the United Nations (UN), the World Health Organization (WHO), the International Labour Organization (ILO) and the Organization for Economic Co-operation and Development (OECD) have warned that the COVID-19 pandemic has made the world a yet more difficult place to be young.

The ILO report *Youth & COVID-19: Impacts on Jobs, Education, Rights and Mental Well-Being*[1] found that nearly three-quarters of people aged 18-29 years reported pandemic-related educational disruptions, one-half described themselves as depressed, and one-in-six of those who were employed before the outbreak had stopped working. The effects have been worst among youth in low- and middle-income countries (LMICs) and among young women everywhere, exacerbating pre-existing inequalities.

Perversely, pandemic-related hardship has pushed some young people prematurely into work, particularly in Asia and the Pacific region. In India and Indonesia, for instance, the UN Children's Fund (UNICEF), the Asian Development Bank and the ILO have jointly reported that poor households are increasingly likely to take underage children out of school to work in the home or away in cities, or to marry them off early to boost family income.

In this issue of the journal, Fusar-Poli et al[2] emphasize that "universal public health approaches targeting the social determinants of mental disorders hold the greatest potential for reducing the risk profile of the whole population". We can extend the focus on inequalities in the socioeconomic environment to incorporate the role that physical environments, built and natural, play in shaping youth mental health, and what can be done in this respect.

By May 2020, governments globally had invested over 10 trillion USD in responses to the pandemic, mostly for crisis initiatives such as furlough schemes, financial support for businesses, and the acquisition of medical supplies. The world is now talking about recovery. Scientists and major international bodies – e.g., the International Monetary Fund, the ILO, the International Energy Agency (IEA), the European Union, and the InterAcademy Partnership – have proposed a green approach to rebuilding economies.

Statista's survey of 28,000 individuals from fifteen nations, *Global Green Economic Recovery Support After COVID-19 2020*[3], found that two-thirds want a green recovery, especially young people. The ILO has hosted a meeting of ministers from thirty countries to discuss how to "build back better", and the UN Secretary-General went so far as to suggest that a green recovery approach in LMICs could help post-pandemic economic development switch from "grey to green". The message is clear: post-pandemic rebuilding cannot continue the over-exploitation of the resources of the planet and its peoples – especially young people – without regard for the costs to either.

Substantial steps have been made in the right direction. The IEA's *Global Energy Review 2020* found that COVID-19 restrictions on travel reduced global carbon emissions by 8%, the kind of fall needed to keep the world within the so-called 1.5°C guardrail beyond which global warming becomes dangerous. However, emissions have started to rise again with the relaxation of restrictions. A commitment to a green recovery, which could avoid 0.3°C warming by 2050[4], is urgently needed.

Leading economists have identified five recovery strategies with particularly strong potential for retaining and even accelerating the emission reductions that the pandemic achieved[5].

The strategies embrace building clean physical infrastructure, retrofitting buildings, and investing in education, training, clean research and development, and natural capital. These are consistent with the WHO's six "prescriptions" for simultaneously promoting planetary and human health outlined in their *Manifesto for a Healthy Recovery from COVID-19*[6]: protecting and preserving nature; investing in essential services for health (e.g., clean water, health care facilities); moving quickly to green energy; healthy and sustainable food systems; stopping subsidizing polluters; and building healthy cities. The UN and the World Bank note that cities are an important focus for a green recovery; the latest UN-HABITAT report has estimated that 60% of the world's population will live in cities by 2030, and 60% of these will be children.

All of these prescriptions and strategies could support universal approaches to promoting young people's future health and prosperity, but it may seem hard to sell some of these ideas politically. However, as the *WHO Manifesto* points out[6], the pandemic has shown that people can accept difficult policies where these are evidently necessary. Further, though politicians may not always listen to scientists and health experts, they listen to public opinion. The large majority of the world's adults wants action on climate change and, as the School Strike for Climate led by Greta Thunberg has shown, those under voting age can be influential.

Clinicians, researchers and their representative bodies have a role to play in persuading opinion leaders of the mental health benefits of a green recovery, especially for young people. This is challenging because its greatest benefits are not immediately obvious. Climate change and mental health are both complex phenomena and their relationship is complicated. It begins high up the causal chain, where climate change aggravates the root causes of mental illness, and ultimately involves multiple reciprocal direct and indirect linkages between a host of proximal, intervening and distal factors that lie on interacting paths of influence[7].

Taking a systems approach to elucidating these relationships can help simplify the complexity meaningfully and shift thinking from the narrow perspective of treating illness to the bigger picture that also incorporates promoting well-being and preventing illness. Systems thinking in this case involves mapping the factors linking climate change to mental health outcomes, from direct, proximate causes to distal root causes, and specifying their interactions. For example, one effect of climate change is to increase the frequency, intensity, unpredictability and duration of extreme events, such as the wildfires that ravaged South-Eastern Australia and California in 2020. Destruction on this scale inevitably has mental health implications that go beyond the immediately obvious, incorporating risks as diverse as significant injury or death, and losses to education and employment, cultural practices, outdoor recreation, access to fresh foods and Internet connectivity. Every one of these cascading factors, separately and interactively, is a potential threat to mental health[7].

Young people can be highly motivated to help in health crises and can mobilize whole communities when needed. Indeed, the ILO report[1] found that, by August 2020, nearly one-third of young people globally was engaged in pandemic-related volunteering. They are also leading a research initiative established by the UN Educational, Scientific and Cultural Organization (UNESCO), *Youth As Researchers*[8], investigating how the pandemic has affected young people.

Developing a youth- and mental health-sensitive approach to COVID-19 recovery would harness the interest, optimism, confidence and energy of young people. It would also address their yearning for a greener future. The Tony Blair Institute for Global Change's report, *Listening to Covid-19's "Lost Generation": Insights From Our Global Youth Survey*[9], has pointed out that young people should help design pandemic recovery pathways.

Members of older generations may feel uneasy about a climate crisis that is their collective bequest to younger cohorts, and may want to help. One thing they can do is to come together more effectively to apply the resources, capabilities and wisdom they have acquired in life to helping young people contribute to the pandemic recovery. Young people are ready to meet the challenge – *their* way, a green way.

Helen L. Berry
Australian Institute of Health Innovation, Faculty of Medicine, Health and Human Sciences, Macquarie University, Sydney, NSW, Australia

1. International Labour Organization. Youth & COVID-19: impacts on jobs, education, rights and mental well-being. www.ilo.org.
2. Fusar-Poli P, Correll CU, Arango C et al. World Psychiatry 2021;20:200-21.
3. Statista. Global green economic recovery support after COVID-19 2020. www.statista.com.
4. Forster PM, Forster HI, Evans MJ et al. Nat Clim Chang 2020;10:913-9.
5. Hepburn C, O'Callaghan B, Stern N et al. Oxford Review of Economic Policy (in press).
6. World Health Organization. Manifesto for a healthy recovery from COVID-19. www.who.int.
7. Berry HL, Waite TD, Dear KBG et al. Nat Clim Chang 2018;8:282-90.
8. UNESCO. Youth As Researchers: exploring the impact of COVID-19. en.unesco.org.
9. Tony Blair Institute for Global Change. Listening to Covid-19's "lost generation": insights from our global youth survey. institute.global.

DOI:10.1002/wps.20843

The promise of machine learning in predicting treatment outcomes in psychiatry

Adam M. Chekroud[1,2], Julia Bondar[2], Jaime Delgadillo[3], Gavin Doherty[4], Akash Wasil[5], Marjolein Fokkema[6], Zachary Cohen[7], Danielle Belgrave[8], Robert DeRubeis[5], Raquel Iniesta[9], Dominic Dwyer[10], Karmel Choi[11,12]

[1]Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA; [2]Spring Health, New York City, NY, USA; [3]Clinical Psychology Unit, Department of Psychology, University of Sheffield, Sheffield, UK; [4]School of Computer Science and Statistics, Trinity College Dublin, Dublin, Ireland; [5]Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA; [6]Department of Methods and Statistics, Institute of Psychology, Leiden University, Leiden, The Netherlands; [7]Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Los Angeles, CA, USA; [8]Microsoft Research, Cambridge, UK; [9]Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neurosciences, King's College London, London, UK; [10]Department of Psychiatry and Psychotherapy, Section for Neurodiagnostic Applications, Ludwig-Maximilian University, Munich, Germany; [11]Harvard T.H. Chan School of Public Health, Boston, MA, USA; [12]Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

For many years, psychiatrists have tried to understand factors involved in response to medications or psychotherapies, in order to personalize their treatment choices. There is now a broad and growing interest in the idea that we can develop models to personalize treatment decisions using new statistical approaches from the field of machine learning and applying them to larger volumes of data. In this pursuit, there has been a paradigm shift away from experimental studies to confirm or refute specific hypotheses towards a focus on the overall explanatory power of a predictive model when tested on new, unseen datasets. In this paper, we review key studies using machine learning to predict treatment outcomes in psychiatry, ranging from medications and psychotherapies to digital interventions and neurobiological treatments. Next, we focus on some new sources of data that are being used for the development of predictive models based on machine learning, such as electronic health records, smartphone and social media data, and on the potential utility of data from genetics, electrophysiology, neuroimaging and cognitive testing. Finally, we discuss how far the field has come towards implementing prediction tools in real-world clinical practice. Relatively few retrospective studies to-date include appropriate external validation procedures, and there are even fewer prospective studies testing the clinical feasibility and effectiveness of predictive models. Applications of machine learning in psychiatry face some of the same ethical challenges posed by these techniques in other areas of medicine or computer science, which we discuss here. In short, machine learning is a nascent but important approach to improve the effectiveness of mental health care, and several prospective clinical studies suggest that it may be working already.

Key words: Computational psychiatry, machine learning, treatment outcomes, prediction, external validation, pharmacotherapies, psychotherapies, electronic health records, smartphone data

(World Psychiatry 2021;20:154–170)

Treatment interventions in psychiatry are far from being effective in all cases in which they are indicated. In depression, for example, only 30-50% of individuals achieve remission after whatever initial treatment they receive, even in the context of a well-conducted clinical trial[1]. Eventually, after trying some number or combination of treatments, most patients do attain remission. What if, rather than iterating through the available treatments that a patient *might* benefit from, we could predict the right treatment for each individual from the start?

Researchers have wanted this for decades. Historically, they have tried to understand specific factors involved in treatment response based on theoretical groundings, leading to many studies focusing on single variables such as early childhood stress, suicidality, major life events, or comorbid diagnoses. Since then, the ongoing search for one (or a few) true explanatory variables has included many levels of analysis, including: the patient (clinical characteristics, blood marker levels), his/her brain (structural and functional neuroimaging, cerebral blood flow, scalp electrical recordings), his/her genes (single nucleotide polymorphisms, mutations/rare genetic variants, copy number variations, gene expression), and intervention characteristics (the medication or psychotherapy selected, the way it was delivered, the provider, the therapeutic alliance). If one variable alone could accurately predict treatment response, our field would probably have found it by now. Instead, most characteristics identified so far have shown small explanatory power over treatment outcomes, and researchers' attention naturally turned towards multivariable models that can incorporate many smaller effects.

Machine learning is a collection of statistical tools and approaches that are extremely well suited to this goal of detecting and aggregating small effects in order to predict an outcome of interest[2]. It allows researchers to go from evaluating a small number (~10) of predictor variables to many hundreds or thousands of variables or variable combinations. There are many potential pitfalls when applying these techniques, but, when implemented well, they afford many opportunities for psychiatric research[3,4]. They allow us to examine many variables, even correlated ones, simultaneously. They move away from exclusively additive models and allow us to identify more complex non-linear patterns in data. They more naturally bridge disparate data types, potentially incorporating clinical assessments, geospatial information, and biological findings into a single analysis. By unlocking powerful hypothesis-free approaches, they enable us to discover factors that are less intuitive but nonetheless predictive of outcomes.

The introduction of machine learning in psychiatry is more than just adding an analysis tool for combining and exploring bigger data sets – it marks a paradigm shift[5]. For years, we used classical statistical approaches to confirm or refute specific hypotheses. Now, machine learning studies shift the focus toward the overall predictive power of a model, particularly how accurately it predicts

the desired outcome in a new, unseen dataset. Studies in this field are evaluated primarily by their potential clinical impact: what our model can reliably tell us about the prognosis of new patients in the future, and what we can do with that information to improve clinical practice.

With this in mind, this paper explores the promise of machine learning in predicting treatment outcomes in psychiatry. There are many things that we do not focus on. This is not a primer on machine learning[6], an explanation of how it works[2], or a debate about what counts as machine learning versus traditional statistics or "non-machine-learning". We do not explain how to build predictive models[7] or how to validate them. We are not formally comparing different algorithmic approaches, how each one works, or circumstances where one may be more appropriate than another. We also avoid a distinction between moderators versus mediators of treatment outcomes, or whether a model predicts outcomes specifically for a treatment versus others or predicts outcomes more generically for multiple treatments[8]. Finally, we do not aim to review the many sociodemographic and clinical variables that have been or can be used for prediction of treatment response in psychiatry, which generally have the most predictive power and are cheapest to collect[9,10].

We begin by discussing machine learning methods, how they compare to traditional statistical approaches, and to what extent is machine learning specifically adding value. Next, we provide an overview of the interventions for which researchers have tried to use machine learning methods to predict outcomes, ranging from medications and psychotherapy to digital interventions and neurobiological treatments. In doing so, we highlight characteristics that made them gold standard examples, and discuss the different goals that can be achieved in each context. Next, we focus on the potential utility of electronic health records, smartphone and social media data, and of data from genetics, electrophysiology, neuroimaging and cognitive testing for the development of predictive models based on machine learning. Finally, we help the reader understand the broader context: how close have we come to implementing these prediction tools in real-world clinical practice; and what are the ethical challenges that these tools carry. The intent of this paper is to review studies throughout psychiatry; any emphasis on depression is not intentional, but it does reflect the fact that the majority of research in this field has been conducted in people with that mental disorder.

IS MACHINE LEARNING ADDING VALUE OVER TRADITIONAL STATISTICS?

Machine learning studies generally differ from traditional research in two ways. The first is a focus on prediction (explanatory power of the model) rather than inference (hypothesis testing). The second is a shift towards model flexibility, with the ability to handle large numbers of predictors simultaneously.

Prediction can be performed without machine learning algorithms, and many studies still use traditional statistical techniques such as logistic regression. In fact, when assumptions and sample size requirements are reasonably met, the number of predictors is small (≤25), and non-linear effects are relatively weak, traditional parametric models will likely predict well. Several studies found no benefit of machine learning over traditional logistic regression, for example in predicting treatment resistance in major depression[11], brain injury outcomes[12], or major chronic diseases[13].

One recent systematic review of clinical prediction models found no difference in performance between machine learning and logistic regression[14], although the authors considered in the category of logistic regression some advanced frameworks that could be included within machine learning, such as penalization (e.g., lasso, ridge or elastic net) and splines (which capture non-linearities). In areas of medicine such as diabetes and heart failure, simple logistic models have performed well and have been externally validated more than machine learning models[15,16].

The added value of machine learning approaches emerges when the number of potential predictors is large and/or their effects are non-linear. Many machine learning algorithms are capable of handling large numbers of predictors, even in cases where there are more predictor variables than observations, due to built-in overfitting control. For example, ridge, lasso and elastic net regression[17] include penalization, which forces the regression coefficients to be closer to zero than in the traditional linear or logistic regression models. Machine learning approaches are also good at capturing complex, interactive, or non-linear effects. For example, tree-based models are able to evaluate many possible variables and variable combinations to identify subgroups that could not be captured by traditional linear models. Another common technique adopted by machine learning approaches is "ensembling". Here, several models are fitted on random samples of the original dataset, and then an average is taken amongst the predictions from each model. This approach is a key element of many popular machine learning techniques today, especially gradient boosting machines and random forests[18-20].

Several recent treatment outcome prediction studies in psychiatry demonstrated the added value of machine learning. Random forests and/or elastic net regression[21-24], as well as support vector machines[25], were found to outperform traditional regression methods. Large-scale comparisons on benchmark datasets consistently found machine learning to outperform traditional methods[26-29]. Overall, boosted trees (random forests and gradient boosting machines), regularized regression, support vector machines, and artificial neural networks can all perform well, but no one method will have the best performance across all situations.

While researchers typically aim to maximize predictive performance, practical aspects such as explainability or the cost of including more variables should also be considered. In some cases, simpler models with slightly lower predictive accuracy or higher generalizability might be preferred, because they already capture most of the effects[30,31]. There is no silver bullet in statistics, and all prediction algorithms face the so-called bias-variance tradeoff[2,32,33], where flexibility needs to be balanced with the risk of overfitting. For machine learning methods to capture increasingly complex effects, much larger sample sizes are still

needed. Although these methods can deal with large numbers of potential predictor variables, careful pre-selection of variables likely improves predictive accuracy.

While traditional research approaches focused on p values for specific coefficients in a model, prediction studies focus on the overall explanatory power of the model, often in terms of R^2, balanced accuracy, or area under the receiver operating characteristic curve (AUC). Predictive studies require a keen focus on validation approaches, to examine whether the model is learning patterns that are substantive and consistent from one dataset to another, or whether the model has simply learned idiosyncrasies of the initial training data. Table 1 discusses various kinds of validation that are conducted in predictive studies, from internal approaches that use just one dataset, to external validation approaches that use data from independent sites, studies, trials, countries, or consortia to test model generalizability. Validation frameworks, especially external validation, are critical for developing models that are reliable and useful, and understanding whether the fitted model is likely to generalize to unseen data in the future[34-36].

PREDICTING TREATMENT OUTCOMES IN PSYCHIATRY BY USE OF MACHINE LEARNING

Medications

Predicting treatment outcomes for psychiatric medications is the most active area of research in the field, primarily because they were the easiest place to start. Machine learning studies require large volumes of data to build predictive models, ideally with clearly labelled outcomes, control over the intervention, and relevant data about the patients before treatment. Since this describes most large clinical trials, and most large clinical trials in psychiatry are conducted to evaluate efficacy of a medication, most machine learning efforts began by investigating treatment responses to medications treating depression, schizophrenia or bipolar disorder.

These studies mostly used information from demographic intake forms and clinical symptom scales common in clinical trials, although more recently genetic and neuroimaging data have also been incorporated (discussed later in this paper). Despite being the most active area of research, most resulting models have not yet been validated in external samples. Relatively few prediction tools generated by mental health researchers so far have advanced through implementation studies and into clinical practice[37-39]. Here we focus on examples of studies that were adequately powered, underwent external validation, or are notable for other reasons.

Most treatment prediction studies have focused on antidepressants commonly used in the acute phase of depression. For example, Chekroud et al[40] determined a small group of 25 pretreatment variables that were most predictive of remission with citalopram in the Sequenced Alternative Treatments for Depression (STAR*D) trial. This model achieved an accuracy of 64.6%. The model was then applied to data from another clinical trial to examine whether it can generalize to patients from an entirely independent population. The model was able to predict response to two similar antidepressant regimens (escitalopram plus placebo, and escitalopram plus bupropion, each with an accuracy of around 60%), but the model did not predict remission better than chance for patients who took venlafaxine plus mirtazapine (51%).

The five most important variables identified by the model in predicting remission were baseline depression severity, employ-

Table 1 Common validation approaches used in clinical prediction studies

Generalizability test	Description
None, p value testing	The entire sample is used to predict an outcome, and a p value indicates the probability of obtaining the result in the absence of a true effect. The study cannot make any claims concerning translation or generalizability because they have not been tested.
Leave-one-out cross-validation	One subject is randomly chosen and left out. A model is trained on the remaining subjects and applied to the left-out subject to assess generalizability. This procedure is repeated for every subject in the dataset. This is the simplest form of cross-validation. It produces optimistic biased results.
K-fold cross-validation	The sample is randomly divided into subsamples (called "folds"). One fold is left out and statistical models are trained on the remaining subjects. The models are applied to the subjects in the left-out fold to assess generalizability. This is a common technique to reduce overfitting. However, when the data are from one sample (even if collected at multiple sites), generalizability claims need to be tempered.
Leave-one-site-out cross-validation	Instead of randomly leaving out subjects, sites are now randomly left out. Models are fitted on the remaining sites, and applied to the left-out site. This assesses cross-site generalizability, and the same technique can be extended to any other group definition, such as blocks of time, gender or ethnicity. Generalizability and translational claims still need to be tempered.
External validation	A model is created in one study and applied to a completely separate sample. This approach reflects a high degree of generalizability capacity. Demonstrations can be increasingly close to real-life circumstances, which strengthens the evidence of generalizability and translational potential (but does not guarantee it). The approach may still be subject to poor sociodemographic representation, sampling biases, or study designs that do not reflect clinical reality.
Prospective validation	A previously-created model is evaluated in a prospective study that is ideally randomized and in conditions as close to clinical reality as possible, in order to test whether the tool is safe and effective in practice. Prospective validation studies are still susceptible to the same concerns around external validity as all other clinical trials (e.g., participant compensation and meaningful endpoints), and require large sample sizes, a broad and unbiased recruitment process, and good clinical practices. As with other clinical trials, a phased process may be necessary to first evaluate feasibility and safety in a smaller sample before proceeding to broad evaluation of effectiveness.

ment status, feeling restless during the past seven days (psychomotor agitation), reduced energy level during the past seven days, and Black or African American ethnicity. The study was later replicated by Nie et al[41], who similarly trained a model to predict citalopram treatment outcomes using information easily obtainable at baseline. The team trained and tested the model in the STAR*D dataset and validated it in data from a different open-label citalopram trial, using 22 predictor variables that overlapped between the two trials. Despite minor differences depending on the specific algorithm used, the balanced accuracy of the models was roughly 64-67%.

An earlier study by Perlis[11] showed that eventual treatment resistance might also be predictable from the outset. The author developed a model using STAR*D data that was able to predict at baseline whether an individual would not reach remission after two antidepressant treatment trials, with an AUC of 0.71. Early proofs of concept like the Perlis study did not include external validation, at least partly due to the lack of independent datasets with similar trial designs that could be used for that validation.

The above antidepressant studies selected predictors in a purely data-driven way, including all data that could be extracted at baseline and then using machine learning methods that discard irrelevant information or are amenable to including many variables at once. However, the choice of predictors is not always hypothesis-free, and *a priori* knowledge from scientific literature can also guide the choice of variables and yield useful results. Iniesta et al[42] aimed to predict remission of depression in patients treated with escitalopram or nortriptyline using only variables that had previously been confirmed as individual predictors or moderators of response to treatment. Their models predicted overall response to medication with an AUC of 0.74 and response to escitalopram with an AUC of 0.75, but prediction of nortriptyline outcomes was not statistically significant. In subsequent work incorporating genetic data to the models[43], these authors predicted response to escitalopram and nortriptyline with an AUC of 0.77.

A second use of machine learning to predict medication outcomes is to better define subgroups of patients, symptoms, or symptom trajectories, and then use these subgroups to make more nuanced predictions. Drysdale et al[44] used clustering to identify four "subtypes", or groups, amongst 1,188 depressed patients based on patterns of dysfunctional connectivity in limbic and frontostriatal networks. They developed classifiers for each depressive subtype using support vector machines and later tested these models on an independent dataset, accurately classifying 86.2% of the testing sample. As a next step, the team used the subtypes to predict response to transcranial magnetic stimulation, but did not validate these predictions in any independent sample. Although the biotypes approach is interesting, subsequent methodological research has highlighted concerns and limitations[45].

Chekroud et al[46] used clustering to identify groups of symptoms and mixed-effects regression to determine if they had different response trajectories. Three symptom clusters (core emotional, sleep and atypical) emerged consistently from two independent medication trials – STAR*D and Combining Medications to Enhance Depression Outcomes (COMED) – across two commonly used symptom scales. The authors subsequently used data from STAR*D to train gradient boosting machines (one for each combination of cluster and medication arm), finding modest improvements in the ability of clusters of symptoms to predict total severity outcomes. The same symptom clustering approach was also effective in a study of treatments for adolescents[47].

Other researchers first used techniques like growth mixture modeling[48] or finite mixture modeling[49] to identify trajectories of symptom response such as "fast and stable remitter", "sustained response", or "late relapse". Machine learning models were then developed to try and predict the specific response trajectory a patient will have for a given treatment. This approach is potentially more robust to the noise that is naturally present amongst individual patient trajectories and less affected by the way that outcomes are defined in trials – e.g., whether remission is defined as a score of 5 on the Patient Health Questionnaire-9 (PHQ-9) or a score of 5 or 6 on the Quick Inventory of Depressive Symptomatology (QIDS)[48,49]. However, the approach relies on the availability of repeated measures.

Medication treatment outcomes have been most widely studied in depression, due to the prevalence of the condition and extant available data, but the approach has also been proven in other psychiatric conditions. For schizophrenia, Koutsouleris et al[25] used data from the European First Episode Schizophrenia Trial (EUFEST, N=344) to predict good and bad outcomes based on global functioning scores over time using a support vector machine, and validated the ten most predictive features on an unseen sample of 108 patients with a balanced accuracy of 71.7%. The most valuable predictors identified were largely psychosocial variables, rather than symptom data: unemployment, poor education, functional deficits, and unmet psychosocial needs.

Again in schizophrenia, Leighton et al[50] were not only successful in predicting response to medication treatment in first episode psychosis, but also in validating findings in two independent samples. They first identified predictors that were available across three studies – the Evaluating the Development and Impact of Early Intervention Services (EDEN) study in England, two cohorts recruited from the National Health Service (NHS) in Scotland, and the Danish clinical trial called OPUS. This allowed them to build and test harmonized models across the three studies to predict four outcomes capturing different aspects of recovery: symptom remission, social recovery, vocational recovery and quality of life. Next, they used logistic regression with elastic net regularization to identify the most relevant predictors in the EDEN study (N=1027) – much like Chekroud et al[40] – to determine a smaller subset of variables that could still predict outcomes but require less effort for future data collection and improve clinical applicability. These regularized models trained in the EDEN sample reached internal validation AUCs of 0.70 to 0.74 (depending on the outcome measure). When tested in the second Scottish cohort, the AUC ranged from 0.68 to 0.87. In the OPUS trial, it ranged from 0.57 to 0.68.

Predicting medication response in other mental disorders is still in early stages. Two studies[51,52] used baseline sociode-

mographic, clinical and family history information to predict response to medications commonly used in bipolar disorder: lithium and quetiapine. Although both obtained models with performance above chance, neither was validated in independent samples, and one used 180 variables for prediction[51], which limits its clinical applicability.

Psychotherapies

Historically, efforts to predict treatment outcomes in psychotherapies have focused on theoretically-motivated single variables that might moderate treatment outcomes. Only relatively recently have psychotherapy researchers applied machine learning approaches to predict treatment outcomes[53]. Even amongst these studies, the historical focus on moderators of psychotherapeutic effects has persisted, leading researchers to distinguish between "prognostic" and "prescriptive" models. Prognostic models are those that predict whether a patient will recover with a given treatment. Prescriptive models instead predict which of two (or more) treatments is best suited for a particular patient[54]. Both kinds of model can clearly have clinical utility, even if they answer slightly different questions. The differences continue to blur further with more recent attempts to build prescriptive models by developing multiple prognostic models for different treatments and then comparing their outputs[55].

In an early effort, Lutz et al[53] used nearest neighbor modeling to predict rate of symptom change and session-by-session variability. Models were based on age, gender and baseline symptom scores. Compared to non-machine learning models, the nearest neighbor predictions were more highly correlated with actual values of rate of change, but not session-by-session variability.

Since then, other approaches to prediction in psychotherapy proliferated. DeRubeis et al[56] developed a multivariable modeling method, known as the "personalized advantage index" (PAI), that uses interaction effects between baseline variables and treatment condition, to predict whether a patient will respond better to antidepressants versus cognitive behavioral therapy (CBT). Amongst their small sample of 154 individuals, a clinically meaningful advantage (PAI ≥3), favoring one of the treatments relative to the other, was predicted for 60% of the patients. When these patients were divided into those randomly assigned to their "optimal" treatment versus those assigned to their "non-optimal" treatment, outcomes in the former group were better (d = 0.58, 95% CI: 0.17-1.01). Similar approaches have been developed by other groups[55,58], and more recently improved further by the use of machine learning approaches[59] to generate better predictions and incorporate more variables.

Several studies since then have tried to predict which evidence-based psychotherapy is most likely to benefit a specific patient[55,59], including efforts to identify which of two (or more) psychotherapies may be most effective[60,61], and whether a given patient is predicted to respond better to psychotherapy or medications[56]. A recent scoping review[62] identified a total of 44 studies that developed and tested a machine learning model in psychotherapy, but only seven of them reported on the feasibility of the tool. Since psychotherapy trials are often expensive and rarely have large sample sizes, some have argued that predictive models may need to be developed initially with large observational datasets[63].

PAI-style approaches that calculate treatment by variable interactions quickly lead to high-dimensionality prediction analyses that are prone to overfitting (or require very large sample sizes). Using data from two Dutch randomized trials, van Bronswijk et al[60] examined whether PAI models developed in one clinical trial dataset were able to successfully generalize to an independent dataset. Although the models performed statistically above chance in the trial used to train them, they did not generalize to the other clinical trial when predicting benefit for CBT versus interpersonal therapy (IPT) for depression.

The psychotherapy literature has generated several other prediction models, potentially optimizing significant aspects of patient care. For example, models have been developed[64,65] that would enable mental health providers to select low- or high-intensity treatments for patients on the basis of their expected prognosis. Other studies have tried to deconstruct the content that is traditionally combined to form a course of psychotherapy treatment, in order to predict which treatment components should be delivered within a given intervention, as well as the order in which the components should be implemented[66-68]. Other novel directions include using machine learning to match patients to specific therapists[69], replicating human ratings and judgements[70,71], and using natural language processing techniques to discover patterns of therapist-patient interactions that predict treatment response[72,73].

In general, many machine learning approaches to predict responses to psychotherapies are in the early stages of development[62]. However, a notable exception is found in the well-developed literature on routine outcome monitoring and "progress feedback". This involves tracking a patient's response to treatment in real time by entering his/her self-reported outcome/symptom measures into a computerized system that compares his/her response to predicted trajectories of improvement derived from clinical data using conventional statistical analyses (e.g., longitudinal multilevel/mixed models and growth curve modelling). There are now over 20 randomized controlled trials and several meta-analyses indicating that such clinical prediction models can help to improve treatment outcomes[74].

In addition to models investigating differential response to treatment and treatment optimization, the psychotherapy literature also includes adequately powered studies predicting overall response to treatment based on sociodemographic and clinical variables, much like the literature on response to medication. Buckman et al[75] built nine different models, using depression and anxiety symptoms, social support, alcohol use, and life events to predict depressive symptom response after 3-4 months of treatment in primary care settings. Models were trained on data from three clinical trials (N=1,722) and tested on three independent trials (N=1,136). All models predicted remission better than a null model using only one post-baseline depression

severity measurement. Green et al[76] also predicted depressive symptom response to psychotherapy in 4,393 patients from community health services. They found that a model with only five pre-treatment variables (initial severity of anxiety and depression, ethnicity, deprivation and gender) predicted reduction of anxiety and depression symptoms with an accuracy of 74.9%. The number of sessions attended/missed was also an important factor affecting treatment response.

Digital CBT

In recent years, online delivery of mental health interventions has been seen as a promising approach to reducing barriers to care, with growing evidence for the effectiveness of both guided and unguided delivery[77,78]. Interventions such as internet-based CBT (iCBT) may be particularly amenable to the use of machine learning techniques, due to the possibility of longitudinal standardized collection of outcome data at scale, and the potential to directly incorporate machine learning outputs into online or app-based interventions. For example, in guided treatments, machine learning tools could provide feedback to therapists or alerts regarding risk. They could also be used to drive just-in-time adaptive interventions[79]. Smartphone delivery also opens up the possibility of automated collection of sensor data to derive behavioral markers[80], which would open up many possibilities for tailored interventions, while also raising a number of privacy and ethical concerns.

Machine learning-derived outcome predictions for iCBT may have advantages with regard to ease of deployment, for example by providing integrated decision support for case management. However, most existing work focused on predicting outcomes has been exploratory in nature and based on modest sample sizes. A key distinction is between approaches that use only baseline pre-treatment data, and hence may be applied to direct the choice of treatment, and approaches which use data gathered during the course of treatment, such as regular outcome measures or ecological momentary assessment (EMA).

As an example of the former, Lenhard et al[81] examined how clinical baseline variables can be used to predict post-treatment outcomes for 61 adolescents in a trial of iCBT for obsessive-compulsive disorder. Whereas multivariable logistic regression detected no significant predictors, the four machine learning algorithms investigated were able to predict treatment response with a 75 to 83% accuracy.

In a study which included, in addition to demographic and clinical data, therapy-related predictors of treatment credibility and working alliance, assessed at week 2 of treatment, Flygare et al[82] used a random forest algorithm to predict remission from body dysmorphic disorder after iCBT in a sample of 88 patients, comparing the results to logistic regression. Random forests achieved a prediction accuracy of 78% at post-treatment, with lower accuracy in subsequent follow-ups. The most important predictors were depressive symptoms, treatment credibility, working alliance, and initial severity of the disorder.

van Breda et al[83] added EMA data to models using baseline measures in a study predicting outcomes for patients who were randomized to blended therapy (face-to-face CBT and iCBT) or treatment as usual. This approach did not improve prediction accuracy.

The effectiveness of digital CBT interventions is mediated by patient engagement[84]. Detailed patient engagement data can be gathered automatically in online or app-based interventions; this may include data such as content views, completion of exercises, and interactions with clinical supporters[85]. Engagement data may be used within predictive models, providing interpretable and actionable outputs (e.g., the need for more frequent therapist contact in order to motivate greater engagement). Chien et al[86] analyzed engagement data from 54,604 patients using a supported online intervention for depression and anxiety. A hidden Markov model was used to identify five engagement subtypes, based on patient interactions with sections of the intervention. Interestingly, while in general patients who engaged more achieved better outcomes, the best outcomes were found in those who were more likely to complete content belonging to key components of CBT (i.e., cognitive restructuring and behavioral activation) within the first two weeks on the program, despite not spending the highest amount of time using the intervention. This work demonstrates the feasibility of gathering detailed engagement and outcome data at scale.

Interactions between patient and therapist, and the content of text in patient exercises, may also be analyzed using sentiment analysis techniques[87]. Analysis of patient texts might be embedded in therapist feedback tools for guided interventions, or as features within predictive models. Ewbank et al[73] conducted an analysis of 90,934 session transcripts (specifically, CBT via real-time text messages). Deep learning was used to automatically categorize utterances from the transcripts into feature categories related to CBT competences, and then multivariable logistic regression was applied to assess the association with treatment outcomes. A number of session features, such as "therapeutic praise", were associated with greater odds of improvement.

Chikersal et al[88] analyzed 234,735 messages sent from clinical supporters to clients within an iCBT platform, examining how support strategies correlate with clinical outcomes. They used k-means clustering to identify supporters whose messages were linked with "high", "medium" or "low" improvements in client outcomes, as measured by PHQ-9 and Generalized Anxiety Disorder-7 (GAD-7). The messages of more successful supporters were more positively phrased, more encouraging, more often used first person plural pronouns, were less abstract, and referenced more social behaviors. Association rule mining was then applied to linguistic features in the messages in order to identify contexts in which particular support strategies were more effective. For less engaged patients, longer, more positive and more supportive messages were linked with better outcomes. For more engaged clients, messages with less negative words, less abstraction, and more references to social behaviors were associated to better outcomes. Such results could ultimately be used in the design of supporter training materials.

One could also try to predict whether a patient engages or drops out of care. Wallert et al[89] aimed to predict adherence to

an online intervention targeting symptoms of depression and anxiety in people who had experienced a myocardial infarction. The analysis included linguistic features of the homework texts as well as demographic and clinical characteristics. The strongest predictors of adherence were cardiac-related fear, gender, and the number of words in the first homework assignment.

Neurobiological treatments

Numerous neurobiological options have emerged as potential treatments for severe and treatment-resistant depression, such as transcranial magnetic stimulation (TMS) and electroconvulsive therapy (ECT). Given the potential risks and side effects of these treatments, as well as their higher financial costs, there is an especially strong interest in identifying for whom they are safe and effective[90-92].

Recent reviews have examined predictors of treatment response and relapse among depressed patients receiving TMS[92-94]. TMS studies with more female patients tend to have higher effect sizes, suggesting that gender may be a predictor of TMS outcomes[95]. Although several studies have attempted to examine neurobiological predictors of response to TMS, the findings are currently inconsistent[92]. Small sample size generally means that machine learning study designs are likely to overfit and produce results that will not replicate later.

Efforts to predict treatment outcomes for ECT are still primarily traditional association studies. Some of them identified a few variables that appear to replicate across studies. Better outcomes have been found for older patients, those with psychotic depression, those with high suicidal intent, and those who exhibit early symptom changes[90,96]. However, due to the small sample size in most ECT trials, and the typically non-randomized study designs, this area has not seen much progress. These are also obstacles to the application of machine learning techniques.

THE UTILITY OF ELECTRONIC HEALTH RECORDS, SMARTPHONE AND SOCIAL MEDIA DATA

Electronic health records (EHR) are increasingly widely adopted in health care systems. They comprise data routinely collected and maintained for individual patients over the course of their clinical care. As such, these data may be particularly useful for building predictive models in psychiatry that could be readily integrated into points of care within clinical settings[97]. EHR data can be divided into two major types: coded structured data, including diagnostic codes, procedure codes, laboratory and medication prescription codes; and unstructured data, including clinical notes and other text-based documentation, which can be mined using natural language processing.

Recent studies have tested the potential of EHR data to predict treatment outcomes in psychiatry, with the bulk of efforts to date focused on depression, though examples exist for bipolar disorder[98] and schizophrenia[99]. Machine learning-based efforts using EHR data have sought to identify those individuals who are likely to drop out after initiating antidepressants[100], those who will show a stable treatment response to antidepressants[101], and those who may transition to a bipolar diagnosis after starting antidepressants for depression[102]. Such applications have shown promising, though still modest and not yet clinically actionable, results.

Applying logistic regression and random forest approaches, Pradier et al[102] used demographic and structured EHR data (i.e., diagnostic, medication and procedure codes) available at the time of initial prescription to predict treatment dropout after initiating one of nine most common antidepressants. Although mean AUC was below 0.70, they found that incorporating EHR data significantly improved prediction of treatment dropout compared to demographic information alone, and that predictive performance varied by type of antidepressant (AUC as high as 0.80 for escitalopram) and provider type (higher accuracy among psychiatrist-treated individuals).

Hughes et al[101] applied logistic regression and extremely randomized trees with demographic and structured EHR data to predict general and drug-specific treatment continuity in patients receiving any of 11 antidepressants, observing a mean AUC of 0.63-0.66 and similar performance when evaluated at a separate site.

Where symptom score (e.g., PHQ-9) data have been available for smaller EHR cohorts (e.g., N<2,500)[103], LASSO models incorporating demographic information, structured and unstructured EHR data, and baseline symptom scores have shown modest-to-adequate performance in predicting improvements in depressive symptom severity, for both medication treatment (AUC=0.66) and psychotherapy (AUC=0.75). However, the most important predictor in these models was baseline symptom scores. Only when symptom scores are routinely integrated into EHR treatment workflows will such models be relevant for outcome prediction in large-scale health systems.

When using EHR data for predicting treatment outcomes in psychiatry, a key challenge is how to operationalize the outcome of interest using available clinical information. This usually involves establishing a set of rules around which relevant EHR features are observed, or not observed, in a cohort of patients over a defined period. For example, treatment dropout was defined by Pradier et al[100] as less than 90 days of prescription availability after index antidepressant initiation, with no evidence of alternative psychiatric treatment procedures. Antidepressant treatment stability, on the other hand, has been defined as two or more antidepressant medication prescription codes at least 30 days apart over a period of at least 90 days, with additional rules about the maximum time gap between adjacent prescription codes, and other medication possession indicators[101].

EHR data are also highly dimensional, with tens of thousands of possible diagnostic codes in addition to possible medication and procedure codes. Machine learning methods may be particularly suitable for modeling complex signals across a diverse set of EHR-based predictors, but also for reducing their dimensions prior to modeling. In their study of antidepressant treatment stability, Hughes et al[101] applied supervised topic modeling using latent

Dirichlet allocation to reduce 9,256 coded EHR features into 10 interpretable empirically derived topics, finding that a classifier for continuous treatment based on this lower-dimensional set of predictors showed comparable performance to a logistic regression based on a higher-dimensional set of features. Simpler methods, such as selecting only diagnostic codes that meet a frequency threshold in the patient population, have also been used[100].

Smartphones can provide various kinds of data that are difficult to acquire through other means. Their first and biggest feature is that they contain many sensors that can passively collect data across a variety of domains. Passive smartphone data include dynamic measures of sleep quality, exercise, heart rate, geospatial locations, language use, and communication patterns[80,104]. Machine learning methods are indispensable for dealing with complex patterns in these sensor data[105]. Currently available studies applying machine learning to predict mental health outcomes using sensor data have generally employed modest samples of 7 to 70 participants, yielding proofs-of-principle more than generalizable results[80,106-108]. Mobile phones also facilitate the collection of EMA data, allowing investigators to perform measurements at frequent intervals (e.g., several times a day). Furthermore, smartphone-based neurocognitive assessments appear to be a promising way to scalably collect cognitive data[109,110].

Few studies have used smartphone data to predict treatment outcomes. These include studies using text data from emails to predict treatment response in patients with social anxiety[111], EMA data to predict changes in self-esteem from an online intervention[112], and EMA data to predict treatment response in patients with depression[83]. In the study predicting depression outcomes, a model including EMA data did not outperform a model using baseline characteristics[83], showing that the former data do not always provide incremental value.

Social media allow investigators to access large amounts of data relating to language use and online activity. However, to our knowledge, these data have not yet been used to predict treatment responses. One of the tradeoffs between incorporating different types of data is the cost and quantity versus quality of data: very often these data present with noise which may hinder the ability to identify meaningful patterns and signals. Novel methods of topological machine learning are robust to noise, and allow to extract descriptors of the shape and structure of data that can augment performance for the analysis of intensive timepoint measurements[113]. Such data with repeated measures may be useful for testing hypotheses, since sample size may compensate for the increased noise of data[114].

THE USE OF DATA FROM GENETICS, ELECTROPHYSIOLOGY, NEUROIMAGING AND COGNITIVE TESTING

Genetics

Machine learning methods are an appealing analytical approach for bridging genetic data with the prediction of treatment response in psychiatry. They put the focus on prediction rather than association, are able to detect interactions between loci, wisely handle correlation, and do not assume a pre-defined statistical model or additivity[115].

Machine learning has been used with the objective to improve prediction of treatment outcomes from genetics alone in many diseases, including cancer[116,117] and hypertension[118].

The question of whether an individual's genetic background could affect how he/she responds to medication treatment has been investigated in pharmacogenomics. An earlier study applying genome-wide complex trait analysis in a sample of roughly 3,000 depressed patients suggested that common genetic variation could explain up to 42% of observed individual differences in antidepressant treatment response[119], suggesting that modeling common genetic variation could be useful for prediction. However, results of pharmacogenomic studies have so far, in general, been underwhelming[120].

Polygenic scores are a common method for quantifying the overall contribution of common genetic variation to particular traits[121]. Polygenic associations with treatment response have been investigated in relatively small patient cohorts (most N<1000) to date, with mixed findings[122-125]. For example, polygenic scores for major depression and schizophrenia did not significantly predict antidepressant efficacy (based on symptom improvement) in classic treatment studies such as Genome-Based Therapeutic Drugs for Depression (GENDEP) and STAR*D[123]. However, these scores were built on earlier genome wide association studies (GWAS) and were likely underpowered. Well-powered GWAS of antidepressant response have produced mixed results, with one study identifying gene sets of relevance for bupropion response[126] and another observing no significant findings for antidepressant resistance[127]. Larger-scale GWAS meta-analysis efforts are needed and ongoing. Even fewer studies have examined common genetic variation associated with responses to other treatment modalities such as psychotherapy[125] or ECT[128].

DNA methylation and gene expression data have been explored in combination with phenotypic datasets of demographic and clinical variables on their ability to predict response to multiple medications. A recent review[129] pointed out genetic prediction of therapeutic outcomes in depression as the most promising[43,130-133], with an overall accuracy of 0.82 (95% CI: 0.77-0.87)[134]. Models combining multiple data types, such as peripheral gene expression data, neuroimaging and clinical variables, achieved significantly higher accuracy[134].

Tree-based approaches were the most popular machine learning methods, followed by penalized regression, support vector machines and deep learning[129]. Studies were quite heterogeneous in design, methods, implementation and validation, limiting our capacity to elucidate the extent to which machine learning integrated with genetics can predict antidepressant drug response.

Evidence for polygenic risk scores versus support vector machines for the prediction of treatment-resistant schizophrenia from GWAS data have been reviewed[135]. Although support vector machines might be more suitable to take into account complex genetic interactions, the traditional polygenic risk score approach

showed higher accuracy for classifying treatment-resistant individuals[115].

Despite many efforts to use many kinds of genetic information in many different ways, results so far have not been sufficiently compelling or accurate to support the use of these approaches to guide clinical care[136,137]. In the future, until novel analytic techniques become available to extract signal from the genome, or a better understanding of the genetic basis for mental illness emerges, the most promising avenue in this context is to integrate genetic information into multivariable analyses to potentially improve broader model performance[133,137].

Electrophysiology and neuroimaging

Tailoring treatment decisions based on brain measures is intuitively appealing and empirically well-justified. Systematic reviews and meta-analyses indicate that therapeutic outcomes are often related to pre-treatment brain differences and that the brain changes as a result of therapy[138-145]. However, in previous research using traditional statistical methods, effect sizes were too low to make the jump from statistical significance to clinical relevance, external validation was rare, sample sizes were small, methodological and site-related variance was high, and in many cases the techniques were not suited to an integration into clinical routine due to their cost-benefit ratio (e.g., positron emission tomography) or reliance on experimental protocols that are unavailable in most clinical settings[138,139,143,145,146]. Machine learning approaches offer hope in overcoming these barriers to clinical implementation. Preliminary reviews comparing accuracies support this optimism by suggesting superiority for treatment prediction with respect to traditional statistical methods[134].

Early studies in this area applied machine learning to detect outcomes such as response to clozapine in psychosis[147] and to selective serotonin reuptake inhibitors (SSRIs) in depression[148-150], but the majority of research has focused on predicting brain stimulation outcomes for depression[148,151-155]. For example, Corlier et al[156] found that alpha spectral correlation could be used to measure EEG connectivity, which then predicted response to repetitive TMS (rTMS), using cross-validated logistic regression, with an accuracy of 77% in a subgroup of depressed individuals. This increased to 81% when adding clinical symptoms of depression. Most studies report predictive accuracies of >80% on the basis of pilot samples consisting of approximately 50 cases or less[155], reflecting the strong likelihood of bias and overfitting that is also seen with magnetic resonance imaging (MRI)[157].

Task-related functional MRI (fMRI) has been used for treatment prediction[158]: for example, by modelling amygdala engagement interactions with early life stress during an experimental task to predict antidepressant outcome[159] or by using fear conditioning responses to predict panic disorder treatment outcome[160,161]. Similar task-related predictive models have been built in a number of studies of CBT[162] or antidepressant responses[162-164]. In task-based fMRI, however, the translational potential is limited due to the use of lengthy and methodologically complicated experimental paradigms. Resting-state fMRI is a popular alternative, because it measures behaviourally-relevant, synchronized brain network activity at rest, and the imaging protocols can be more easily harmonized across scanners[165]. Studies in this field have demonstrated similar accuracies for CBT[166], trauma-focused psychotherapy[167], antidepressant treatment[168], and antipsychotic therapy[169], while also showing predictive accuracy for ECT[165,170].

A challenge of functional imaging is reliability across scanners, especially in non-experimental clinical settings. Structural neuroimaging may provide an opportunity for faster implementation into existing clinical routines. Most studies have involved grey matter measurements, and ECT treatment prediction has been a frequent focus, with studies using whole-brain approaches[171], regional measurements[172], and combinations of neuroimaging modalities[173]. White matter measurements (e.g., with diffusion tensor imaging) have been relatively less commonly considered.

Overall, the lack of multi-site studies and external validation reflects the pilot-study stage of research in this area, where results can be interpreted as promising but highly experimental. Whether the machine learning results will ultimately agree with the low effect sizes found with classical statistical approaches remains an open question[143,145].

Cognitive testing

Cognitive testing is a straightforward method to indirectly assess brain functioning that has been historically linked to treatment outcomes. Although such testing can be time-consuming and costly when performed by a trained neuropsychologist, more recent computerized methods can facilitate efficient digital assessments that lend themselves especially well to machine learning, including from passively collecting smartphone measurements as described above[80,114,174].

Etkin et al[175] conducted an early study in this area, as part of the international Study to Predict Optimized Treatment in Depression (iSPOT-D), aimed to predict response to antidepressant treatment using a battery of computerized cognitive tasks assessing attention, processing speed, memory, and executive and emotional functions. In order to obtain accurate predictive estimates, they first classified depressed individuals into a subgroup with particularly poor cognition before training a supervised discriminant function to predict remission. Results demonstrated that remission following escitalopram could be predicted with 72% accuracy, but this was not confirmed with sertraline or venlafaxine.

Subtyping or unsupervised learning approaches have also been helpful to identify response trajectories to cognitive training. A recent study found that self-organizing maps detecting multivariate relationships between cognitive functions associated with working memory task performance could identify individuals who differentially responded to the training[176].

HOW CLOSE WE HAVE COME TO REAL-WORLD IMPLEMENTATION

Not all prediction models will translate readily for use in clinical or other real-world settings. In evaluating the readiness of predictive models for real-world implementation, key criteria include external validation, empirical support from implementation trials, and acceptability to users (e.g., clinicians).

External cross-validation remains the gold standard for evaluating real-world performance, as it quantifies performance loss when a trained model is applied to a completely independent sample. In addition, it guards against increased researcher degrees-of-freedom that may result from the many tuning parameters of more complex machine learning methods. A review focusing on machine learning in psychotherapy research reported that only 3 of 51 studies had performed external validation[62].

Studies without external validation are at high risk of overconfidence, as demonstrated by Van Bronswijk et al[60], who developed and then tested a treatment selection model across two randomized controlled trials comparing CBT and IPT. They found that the estimated effect size for the benefit of receiving the model-recommended treatment (generated through internal cross-validation) shrunk by 77% when the model was tested using the second study's data (external validation).

Some prediction efforts using large naturalistic samples have reported positive results following external validation[65,177,178].

When a model undergoes external validation and successfully predicts outcomes, the next step towards real-world use is an implementation trial. These trials provide the most compelling evidence for the value of a decision support tool. Here, patients are usually allocated to algorithm-guided treatment (generally within a shared decision-making framework) or treatment as usual.

Trial-based efforts to evaluate the efficacy of treatment personalization tools have begun to emerge. One example is a multi-service cluster randomized trial[179], in which patients (N=951) were referred to either high- or low-intensity psychotherapy. In one arm, the choice of intensity was informed by an algorithm previously developed in a naturalistic dataset. In the other arm, most patients started on low-intensity psychotherapy and were later referred to high-intensity treatment in the case of non-response, as per usual stepped care. The study found higher depression remission rates in patients whose initial treatment was recommended by the algorithm compared to usual stepped care (52.3% vs. 45.1%, odds ratio, OR=1.40, p=0.025).

Another recent example comes from Lutz et al[180], who used archival data from an outpatient CBT clinic to develop a predictive decision support system providing therapists with treatment strategy recommendations and psychometric feedback enhanced with clinical problem-solving tools. They randomized therapist-patient dyads (N=538) to treatment as usual or to algorithm-informed treatment. They reported that, overall, outcomes for those who were randomized to the intervention did not differ from those who received usual care. However, there was significant variability in the extent to which therapists in the intervention condition followed the recommendations provided by the decision support tool. When the authors analyzed outcomes for patients whose therapists had followed the recommendations, significant benefits emerged.

Browning et al[181] conducted another trial randomizing depressed patients to either algorithm-informed care or usual care for depression. Their algorithm, called PReDicT, used information from symptom scales and behavioral tests of affective cognition to predict non-response to treatment with citalopram. After eight weeks of treatment, the rate of depressive symptom response in the PReDicT arm was 55.9%, versus 51.8% in the usual care arm (not significant, OR=1.18, p=0.25). Of all instances where the algorithm predicted non-response, only 65% prompted a change in treatment regimen, and most consisted of an increase in dosage only.

In combination, the above findings highlight that accurate algorithms are not enough to ensure the success of a decision support system for precision treatment[39]. When randomizing patients to algorithm-informed care or usual care, clinicians may override algorithm recommendations and choose alternative treatments. Patients may refuse the algorithm-recommended treatment, or have restrictions to its use that were not contemplated by the decision support tool (e.g., prohibitive cost of therapy). In light of this, effect sizes for these interventions will often vary when applied in different settings.

The use of predictive models may be uniquely challenging in psychotherapy research and practice. One challenge is that a given therapist is only trained to provide a limited subset of psychotherapies. Whereas a psychiatrist may be qualified to prescribe a large number of different medications or medication combinations, a psychotherapist is less likely to be able to competently provide many different psychotherapies. Another consideration is that predictions from a model may lead to self-fulfilling prophecies, in which clinicians treat "easy" patients (those with good prognoses) differently than "difficult" patients[182].

For both medications and psychotherapies, in real-world, treatment decisions are rarely going to be made solely based on model recommendations. Rather, these decisions will involve the preferences of patients, the recommendations of clinicians, the availability and costs of treatments, and several other considerations[183]. As such, the development of data-driven decision tools should be informed by extensive consultation and co-production with the intended users, in order to implement models that maximize acceptability and compatibility with other clinical guidelines (i.e., risk management procedures, norms about safe dosage or titration of medications).

Another crucial barrier to implementation is the interpretability of machine learning models. As algorithms become increasingly complex, sometimes called "black box" algorithms, they can become very difficult to interpret, and therefore unlikely to be acceptable to clinical users. Methods for explaining predictions of complex models have therefore been developed[184,185], but there is currently no agreed-upon measure for assessing the quality or accuracy of these explanations. In addition, black-box predictive models combined with (similarly complex) explanatory methods may yield complicated decision pathways that increase the likelihood of human error[186].

In order to ensure that algorithm recommendations are used in trials, additional thought and effort must be devoted to issues of dissemination and implementation, with the goal of making the recommendations simple to generate, easy to understand, trustworthy, ethical, cost-effective, and compelling enough to influence the decision-maker(s)[187].

A recent experiment was conducted with 220 antidepressant-prescribing clinicians to assess the impact of providing machine learning recommendations and accompanying explanations[188]. It was found that recommendations did not improve accurate selection of antidepressants in hypothetical patient scenarios, and that accuracy was even lower when incorrect recommendations were presented than when standard information was available. Prospective field-tests[181,189] are one method for identifying the myriad institutional, cultural and contextual factors that could affect the uptake and sustained use of a precision psychiatry tool, aiming to co-produce acceptable and interpretable decision tools with the intended users.

ETHICAL CHALLENGES

From the development of machine learning tools to their potential deployment into clinical care, we can identify several ethical challenges[190-193].

The first challenge concerns responsibility. With the implementation of machine learning programs into clinical practice, physicians and machine learning-based tools would become "teammates" that collaborate in selecting an optimal treatment[194,195]. In such a scenario, who will hold authority and ethical responsibility over the decision made? We believe that a competent human agent should check and take final responsibility on the machine learning-based suggestions[196], as only he/she is equipped with empathy, a good understanding of the contextual environment and, most uniquely, consciousness.

The second challenge is to avoid dehumanization[197]. Machine learning can incorporate a great variety of psychological, environmental and social variables, and there is some progress towards including subjective patient experience into machine learning models[198]. However, giving a patient the space to articulate his/her concerns is essential to ensure accurate diagnosis, health outcomes, and humane care[199].

Third, making decisions is an intricate part of physicians' activity. The non-expert tends to act as a "technician" and more likely relies on protocols, whilst the expert, after the observation of many cases, is more prone to making decisions based on tacit knowledge[200-202]. The ethical mandate is that practitioners use all of their capabilities, including those based on self-experience and observation, even if this is in discordance with a statistical model. Disagreements between physicians and machine learning-based decisions may lead to consultations with other clinicians[193]. However, in the context of modern health care systems, respecting clinicians' judgement is vital[193,203], and they should not be forced to act against their own criteria (freedom of action)[204].

Practitioners (especially those with less expertise) might be in danger of not developing/losing their own clinical judgement and become dependent on automatically deployed machine learning outcomes[205], particularly for those complex cases that they fear they are not competent enough to solve. This would risk disempowerment of clinicians. On the other hand, it is a physicians' duty to train themselves in the use, understanding and interpretation of machine learning applications, so that they can trust the system and its outputs, and contribute to patients' acceptance[206].

Machine learning tools need to be transparent to the human teammates to facilitate understanding[194,207]. The idea of transparency is opposite to that of "black-box" machine learning algorithms, in which the patterns the algorithm follows to make a decision for a given patient are opaque to the person and even to the developer, making very challenging (if not impossible) for the affected person to understand how the system worked out an output for him/her[190]. This risks not only increasing clinicians' resistance to use the tool, but also disempowering patients and disrespecting their autonomy. Developers should consider simpler algorithms that balance interpretability with accuracy[191].

Furthermore, a central issue in fair machine learning development arises when the training dataset is not a good representation of the phenomenon being studied[192,208]. A model trained in such data will predict erroneous outcomes for groups that were underrepresented[209]. For example, a widely used machine learning algorithm assigned the same level of disease risk to Black and White patients, even if Black patients were sicker than White patients[210]. As a consequence, the system was actively causing harm to Black patients by leading to allocation of fewer resources to them. Potentially discriminatory predictors should be left out of the model, but developers should be aware that surrogate variables correlated with the excluded set might still become relevant for prediction. Objective unbiased applications might help reduce discrimination in machine learning[211,212].

Finally, the risk of misuse of personal and sensitive data exchanged in machine learning is high[213]. For this reason, machine learning tools can be only used when data security and privacy are guaranteed.

CONCLUSIONS

This paper reviews several studies suggesting that it is possible to predict outcomes and personalize psychiatric treatment by using machine learning. Several gold standard prediction studies have shown that we can predict whether a depressed patient will respond to specific antidepressants[40,41], to specific psychotherapeutic techniques[177], and whether patients with first episode psychosis will have good prognosis after one year with certain antipsychotic medications[25,50]. At least three predictive models have even been tested in prospective clinical trials.

Despite this progress, the potential for machine learning in psychiatry has just begun to be explored. Predicting treatment response is just one relatively narrow use case where machine learning can add value and improve mental health care. Predic-

tion can help with so many more clinical decisions and clinical processes. We could predict barriers that prevent an individual from engaging in care initially, or non-adherence or dropout from care after initiation. We could streamline patients to the appropriate level of care, such as self-guided programs vs. outpatient care, or intensive outpatient versus inpatient care, to maximize scarce health care resources. In selecting a specific treatment approach, we could optimize dosing or predict side effect profiles in order to improve symptoms but minimize impact on patient quality of life. Some psychiatric treatments carry high cost (e.g., ketamine, ECT) or unwanted side effects (e.g., metabolic disruption and weight gain for antipsychotics). Doing no harm is arguably more important than improving the probability of recovery, and so precision mental health efforts could be especially important in identifying which treatments are safest and most tolerable.

Machine learning could even help sequence treatments over time, or design specific treatment protocols for an individual. For example, modular psychological interventions can be personalized[66,68], or tailored health behavior change interventions can be customized for an individual. This form of personalization and customization has proven effective in contexts like smoking cessation, breast cancer screening, and physical activity[214,215].

Techniques like natural language processing, often using machine learning algorithms, give us the ability to draw insights from text-based data – e.g., social media posts, peer-support conversations, or conversation transcriptions – that might inform the content that is offered to an individual as part of his/her treatment to maximize future outcomes. In addition, the same analytic techniques may form the basis of interventions, such as chatbots, that could provide scalable support for loneliness, stress, or other subclinical psychological issues when human support is unavailable or not clinically warranted. This personalization of iCBT treatment may be particularly necessary for unguided interventions, where non-adherence is widespread and undermines the potential for symptom relief.

Machine learning is a powerful tool that can help sift through multi-modal predictors and model their complex/non-linear contributions. And it can identify specific subtypes of patients, e.g., through clustering, for more nuanced prediction of treatment outcomes. Machine learning techniques are allowing us to extract more knowledge from bigger datasets in a more efficient way – which is a good and promising thing.

However, the ultimate goal of psychiatry is to better treat mental illness. The path toward machine learning improving psychiatric care in real-life settings is not only governed by statistical, but also by implementation considerations. Recent seminal findings[180,181] highlight that accurate algorithms alone are not enough to ensure the success of a decision support system for precision treatment. This is because many things change in the transition from a research setting into real patient care[39]. In practice, clinicians may override algorithm recommendations and choose alternative treatments. Patients may refuse the algorithm-recommended treatment, or have restrictions to its use that were not contemplated by the decision support tool. Recommendations may be provided in a poorly-designed user interface, and thus may go unseen or be actively ignored. All of these factors contribute to a general phenomenon of reduced effect sizes when an algorithm is implemented in clinical practice.

In our own personal experience, patient concerns around privacy are a very real problem. Because mental health is particularly sensitive, capturing personal data can be challenging and we need to innovate ways of collecting these data so that we do not have a biased perspective of the landscape due to a poor sampling within certain groups. Data needs to be collected in such a way that participants are aware of how and for what purposes those data will be used[216].

Technology systems must implement careful logging processes to examine concept or data drift, where the underlying distribution of a predictor or an outcome changes over time, and to ensure that the inputs and outputs of the system are auditable. This is a collective exercise of building trust in predictive models and how these will be potentially used to enhance patient outcomes, and can avoid the introduction of harm or biases in decision-making processes.

This paper reviews many kinds of data that have been used to predict treatment outcomes in psychiatry. Ultimately, treatment responses emerge from multiple interacting biological, psychological and social factors. Therefore, in theory, multi-modal approaches using demographic, clinical and brain variables should result in the most accurate predictions[217]. However, to this date, it is clear that certain kinds of data – specifically sociodemographic, self-report, psychosocial and clinical data – consistently offer more meaningful and generalizable predictions. Other types of data that might be more scientifically appealing – such as neuroimaging and genetic data – have not yet shown compelling results in a large external sample, let alone in prospective implementation studies.

Ultimately, data types that can be easily integrated into clinical care in a cost-effective and ethical way, which is appropriate for the prevalence and invasiveness of the therapy, are most likely to show favorable return on investment for ultimate decision makers in health systems and health payers.

ACKNOWLEDGEMENTS

A.M. Chekroud holds equity in Spring Care Inc., Carbon Health Technologies Inc., and UnitedHealthcare. J. Bondar is an employee of Spring Care Inc. G. Doherty is co-founder of SilverCloud Health. R. Iniesta is supported by the Brain & Behavior Research Foundation and the National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre.

REFERENCES

1. Rush AJ, Trivedi MH, Wisniewski SR et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am J Psychiatry 2006;163:1905-17.
2. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction, 2nd ed. New York: Springer, 2009.
3. Chekroud AM. Bigger data, harder questions – opportunities throughout mental health care. JAMA Psychiatry 2017;64:44-50.
4. Iniesta R, Stahl D, McGuffin P. Machine learning, statistical learning and the future of biological research in psychiatry. Psychol Med 2016;46:2455-65.
5. Breiman L. Statistical modeling: the two cultures (with comments and a rejoinder by the author). Stat Sci 2001;16:199-231.

6. Delgadillo J. Machine learning: a primer for psychotherapy researchers. Psychother Res 2021;31:1-4.
7. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. BMC Med Res Methodol 2019;19:64.
8. Cohen ZD, DeRubeis RJ. Treatment selection in depression. Annu Rev Clin Psychol 2018;14:209-36.
9. Kessler RC, van Loo HM, Wardenaar KJ et al. Using patient self-reports to study heterogeneity of treatment effects in major depressive disorder. Epidemiol Psychiatr Sci 2017;26:22-36.
10. 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.
11. Perlis RH. A clinical risk stratification tool for predicting treatment resistance in major depressive disorder. Biol Psychiatry 2013;74:7-14.
12. Gravesteijn BY, Nieboer D, Ercole A et al. Machine learning algorithms performed no better than regression models for prognostication in traumatic brain injury. J Clin Epidemiol 2020;122:95-107.
13. Nusinovici S, Tham YC, Yan MYC et al. Logistic regression was as good as machine learning for predicting major chronic diseases. J Clin Epidemiol 2020;122:56-69.
14. Christodoulou E, Ma J, Collins GS et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. J Clin Epidemiol 2019;110:12-22.
15. Desai RJ, Wang SV, Vaduganathan M et al. Comparison of machine learning methods with traditional models for use of administrative claims with electronic medical records to predict heart failure outcomes. JAMA Netw Open 2020;3:e1918962.
16. Lynam AL, Dennis JM, Owen KR et al. Logistic regression has similar performance to optimised machine learning algorithms in a clinical setting: application to the discrimination between type 1 and type 2 diabetes in young adults. Diagn Progn Res 2020;4:6.
17. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010;33:1-22.
18. Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychol Meth 2009;14:323.
19. Bühlmann P. Bagging, boosting and ensemble methods. In: Gentle J, Härdle W, Mori Y (eds). Berlin: Springer, 2012:985-1022.
20. van der Laan MJ, Polley EC, Hubbard AE et al. Super learner. Stat Appl Genet Mol Biol 2007;6:4765-74.
21. Kessler RC, van Loo HM, Wardenaar KJ et al. Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports. Mol Psychiatry 2016;21:1366-71.
22. Pearson R, Pisner D, Meyer B et al. A machine learning ensemble to predict treatment outcomes following an Internet intervention for depression. Psychol Med 2019;49:2330-41.
23. Webb CA, Trivedi MH, Cohen ZD et al. Personalized prediction of antidepressant v. placebo response: evidence from the EMBARC study. Psychol Med 2019;49:1118-27.
24. Hilbert K, Kunas SL, Lueken U et al. Predicting cognitive behavioral therapy outcome in the outpatient sector based on clinical routine data: a machine learning approach. Behav Res Ther 2020;124:103530.
25. Koutsouleris N, Kahn RS, Chekroud AM et al. Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach. Lancet Psychiatry 2016;3:935-46.
26. Fernández-Delgado M, Cernadas E, Barro S et al. Do we need hundreds of classifiers to solve real world classification problems? J Mach Learn Res 2014;15:3133-81.
27. Gacto MJ, Soto-Hidalgo JM, Alcalá-Fdez J et al. Experimental study on 164 algorithms available in software tools for solving standard non-linear regression problems. IEEE Access 2019;7:108916-39.
28. Wainer J. Comparison of 14 different families of classification algorithms on 115 binary datasets. arXiv 2016;1606.00930.
29. Zhang C, Liu C, Zhang X et al. An up-to-date comparison of state-of-the-art classification algorithms. Expert Syst Appl 2017;82:128-50.
30. Hand DJ. Classifier technology and the illusion of progress. Stat Sci 2006;21:1-14.
31. Holte RC. Very simple classification rules perform well on most commonly used datasets. Mach Learn 1993;11:63-90.
32. Yarkoni T, Westfall J. Choosing prediction over explanation in psychology: lessons from machine learning. Perspect Psychol Sci 2017;12:1100-22.
33. de Rooij M, Pratiwi BC, Fokkema M et al. The early roots of statistical learning in the psychometric literature: a review and two new results. arXiv 2018;1911.11463.
34. Larson SC. The shrinkage of the coefficient of multiple correlation. J Educ Psychol 1931;22:45-55.
35. Mosier CI. I. Problems and designs of cross-validation 1. Educ Psychol Meas 1951;11:5-11.
36. Wainer H. Estimating coefficients in linear models: it don't make no nevermind. Psychol Bull 1976;83:213-7.
37. Delgadillo J, Lutz W. A development pathway towards precision mental health care. JAMA Psychiatry 2020;77:889-90.
38. Browning M, Carter CS, Chatham C et al. Realizing the clinical potential of computational psychiatry: report from the Banbury Center meeting, February 2019. Biol Psychiatry 2020;88:e5-10.
39. Chekroud AM, Koutsouleris N. The perilous path from publication to practice. Mol Psychiatry 2018;23:24-5.
40. 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.
41. Nie Z, Vairavan S, Narayan VA et al. Predictive modeling of treatment resistant depression using data from STAR*D and an independent clinical study. PLoS One 2018;13:e0197268.
42. 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.
43. Iniesta R, Hodgson K, Stahl D et al. Antidepressant drug-specific prediction of depression treatment outcomes from genetic and clinical variables. Sci Rep 2018;8:5530.
44. Drysdale AT, Grosenick L, Downar J et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat Med 2017;23:28-38.
45. Dinga R, Schmaal L, Penninx B et al. Evaluating the evidence for biotypes of depression: methodological replication and extension of. Neuroimage Clin 2019;22:101796.
46. Chekroud AM, Gueorguieva R, Krumholz HM et al. Reevaluating the efficacy and predictability of antidepressant treatments: a symptom clustering approach. JAMA Psychiatry 2017;74:370-8.
47. Bondar J, Caye A, Chekroud AM et al. Symptom clusters in adolescent depression and differential response to treatment: a secondary analysis of the Treatment for Adolescents with Depression Study randomised trial. Lancet Psychiatry 2020;7:337-43.
48. Gueorguieva R, Chekroud AM, Krystal JH. Trajectories of relapse in randomised, placebo-controlled trials of treatment discontinuation in major depressive disorder: an individual patient-level data meta-analysis. Lancet Psychiatry 2017;4:230-7.
49. Paul R, Andlauer Till FM, Czamara D et al. Treatment response classes in major depressive disorder identified by model-based clustering and validated by clinical prediction models. Transl Psychiatry 2019;9:187.
50. Leighton SP, Upthegrove R, Krishnadas R et al. Development and validation of multivariable prediction models of remission, recovery, and quality of life outcomes in people with first episode psychosis: a machine learning approach. Lancet Digit Health 2019;1:e261-70.
51. Nunes A, Ardau R, Berghöfer A et al. Prediction of lithium response using clinical data. Acta Psychiatr Scand 2020;141:131-41.
52. Kim TT, Dufour S, Xu C et al. Predictive modeling for response to lithium and quetiapine in bipolar disorder. Bipolar Disord 2019;21:428-36.
53. Lutz W, Leach C, Barkham M et al. Predicting change for individual psychotherapy clients on the basis of their nearest neighbors. J Consult Clin Psychol 2005;73:904-13.
54. Lambert MJ, Bergin AE. Bergin and Garfield's handbook of psychotherapy and behavior change. Chichester: Wiley, 2021.
55. Deisenhofer AK, Delgadillo J, Rubel JA et al. Individual treatment selection for patients with posttraumatic stress disorder. Depress Anxiety 2018;35:541-50.
56. DeRubeis RJ, Cohen ZD, Forand NR et al. The personalized advantage index: translating research on prediction into individualized treatment recommendations. A demonstration. PLoS One 2014;9:1-8.
57. Cloitre M, Petkova E, Su Z et al. Patient characteristics as a moderator of posttraumatic stress disorder treatment outcome: combining symptom burden and strengths. BJPsych Open 2016;2:101-6.
58. Wallace ML, Frank E, Kraemer HC. A novel approach for developing and interpreting treatment moderator profiles in randomized clinical trials. JAMA Psychiatry 2013;70:1241-7.
59. Keefe JR, Wiltsey Stirman S, Cohen ZD et al. In rape trauma PTSD, patient characteristics indicate which trauma-focused treatment they are most likely to complete. Depress Anxiety 2018;35:330-8.

60. van Bronswijk SC, Bruijniks SJE, Lorenzo-Luaces L et al. Cross-trial prediction in psychotherapy: external validation of the Personalized Advantage Index using machine learning in two Dutch randomized trials comparing CBT versus IPT for depression. Psychother Res 2021;31:78-91.
61. Cohen ZD, Kim TT, Van HL et al. A demonstration of a multi-method variable selection approach for treatment selection: recommending cognitive-behavioral versus psychodynamic therapy for mild to moderate adult depression. Psychother Res 2020;30:137-50.
62. Aafjes-van Doorn K, Kamsteeg C, Bate J et al. A scoping review of machine learning in psychotherapy research. Psychother Res 2021;31:92-116.
63. Luedtke A, Sadikova E, Kessler RC. Sample size requirements for multivariate models to predict between-patient differences in best treatments of major depressive disorder. Clin Psychol Sci 2019;7:445-61.
64. Lorenzo-Luaces L, DeRubeis RJ, van Straten A et al. A prognostic index (PI) as a moderator of outcomes in the treatment of depression: a proof of concept combining multiple variables to inform risk-stratified stepped care models. J Affect Disord 2017;213:78-85.
65. Delgadillo J, Huey D, Bennett H et al. Case complexity as a guide for psychological treatment selection. J Consult Clin Psychol 2017;85:835-53.
66. Fisher AJ, Bosley HG, Fernandez KC et al. Open trial of a personalized modular treatment for mood and anxiety. Behav Res Ther 2019;116:69-79.
67. Lutz W, Rubel JA, Schwartz B et al. Towards integrating personalized feedback research into clinical practice: development of the Trier Treatment Navigator (TTN). Behav Res Ther 2019;120:103438.
68. Rubel JA, Fisher AJ, Husen K et al. Translating person-specific network models into personalized treatments: development and demonstration of the dynamic assessment treatment algorithm for individual networks (DATA-IN). Psychother Psychosom 2018;87:249-51.
69. Delgadillo J, Rubel J, Barkham M. Towards personalized allocation of patients to therapists. J Consult Clin Psychol 2020;88:799-808.
70. Atkins DC, Steyvers M, Imel ZE et al. Scaling up the evaluation of psychotherapy: evaluating motivational interviewing fidelity via statistical text classification. Implement Sci 2014;9:49.
71. Goldberg SB, Tanana M, Imel ZE et al. Can a computer detect interpersonal skills? Using machine learning to scale up the Facilitative Interpersonal Skills task. Psychother Res 2021;31:281-8.
72. Ewbank MP, Cummins R, Tablan V et al. Quantifying the association between psychotherapy content and clinical outcomes using deep learning. JAMA Psychiatry 2020;77:35-43.
73. Ewbank MP, Cummins R, Tablan V et al. Understanding the relationship between patient language and outcomes in internet-enabled cognitive behavioural therapy: a deep learning approach to automatic coding of session transcripts. Psychother Res 2020;31:326-38.
74. de Jong K, Conijn JM, Gallagher RAV et al. Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: a multilevel meta-analysis. Clin Psychol Rev 2021;85:102002.
75. Buckman JEJ, Cohen ZD, O'Driscoll C et al. Predicting prognosis for adults with depression using individual symptom data: a comparison of modelling approaches. Open Sci Framework (in press).
76. Green SA, Honeybourne E, Chalkley SR et al. A retrospective observational analysis to identify patient and treatment-related predictors of outcomes in a community mental health programme. BMJ Open 2015;5:e006103.
77. Andersson G, Titov N, Dear BF et al. Internet-delivered psychological treatments: from innovation to implementation. World Psychiatry 2019;18:20-8.
78. Andrews G, Basu A, Cuijpers P et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. J Anxiety Disord 2018;55:70-8.
79. Nahum-Shani I, Smith SN, Spring BJ et al. Just-in-Time Adaptive Interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018;52:446-62.
80. Mohr DC, Zhang M, Schueller SM. Personal sensing: understanding mental health using ubiquitous sensors and machine learning. Annu Rev Clin Psychol 2017;13:23-47.
81. Lenhard F, Sauer S, Andersson E et al. Prediction of outcome in internet-delivered cognitive behaviour therapy for paediatric obsessive-compulsive disorder: a machine learning approach. Int J Methods Psychiatr Res 2018;27:1-11.
82. Flygare O, Enander J, Andersson E et al. Predictors of remission from body dysmorphic disorder after internet-delivered cognitive behavior therapy: a machine learning approach. BMC Psychiatry 2020;20:1-9.
83. van Breda W, Bremer V, Becker D et al. Predicting therapy success for treatment as usual and blended treatment in the domain of depression. Internet Interv 2018;12:100-4.
84. Yardley L, Spring BJ, Riper H et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016;51:833-42.
85. Morrison C, Doherty G. Analyzing engagement in a web-based intervention platform through visualizing log-data. J Med Internet Res 2014;16:e252.
86. Chien I, Enrique A, Palacios J et al. A machine learning approach to understanding patterns of engagement with internet-delivered mental health interventions. JAMA Netw Open 2020;3:e2010791.
87. Provoost S, Ruwaard J, van Breda W et al. Validating automated sentiment analysis of online cognitive behavioral therapy patient texts: an exploratory study. Front Psychol 2019;10:1-12.
88. Chikersal P, Belgrave D, Doherty G et al. Understanding client support strategies to improve clinical outcomes in an online mental health intervention. Presented at the ACM Conference on Human Factors in Computing Systems, Honolulu, April 2020.
89. Wallert J, Gustafson E, Held C et al. Predicting adherence to internet-delivered psychotherapy for symptoms of depression and anxiety after myocardial infarction: machine learning insights from the U-CARE heart randomized controlled trial. J Med Internet Res 2018;20:e10754.
90. Pinna M, Manchia M, Oppo R et al. Clinical and biological predictors of response to electroconvulsive therapy (ECT): a review. Neurosci Lett 2018;669:32-42.
91. Haq AU, Sitzmann AF, Goldman ML et al. Response of depression to electroconvulsive therapy: a meta-analysis of clinical predictors. J Clin Psychiatry 2015;76:1374-84.
92. Kar SK. Predictors of response to repetitive transcranial magnetic stimulation in depression: a review of recent updates. Clin Psychopharmacol Neurosci 2019;17:25-33.
93. Miljevic A, Bailey NW, Herring SE et al. Potential predictors of depressive relapse following repetitive transcranial magnetic stimulation: a systematic review. J Affect Disord 2019;256:317-23.
94. Nord CL. Predicting response to brain stimulation in depression: a roadmap for biomarker discovery. Curr Behav Neurosci Rep 2021;8:11-9.
95. Kedzior KK, Azorina V, Reitz S. More female patients and fewer stimuli per session are associated with the short-term antidepressant properties of repetitive transcranial magnetic stimulation (rTMS): a meta-analysis of 54 sham-controlled studies published between 1997-2013. Neuropsychiatr Dis Treat 2014; 10:727-56.
96. Yao Z, McCall WV, Essali N, et al. Precision ECT for major depressive disorder: a review of clinical factors, laboratory, and physiologic biomarkers as predictors of response and remission. Pers Med Psychiatry 2019;17-18:23-31.
97. Smoller JW. The use of electronic health records for psychiatric phenotyping and genomics. Am J Med Genet 2018;177:601-12.
98. Hayes JF, Marston L, Walters K et al. Lithium vs. valproate vs. olanzapine vs. quetiapine as maintenance monotherapy for bipolar disorder: a population-based UK cohort study using electronic health records. World Psychiatry 2016;15:53-8.
99. Wu C-S, Luedtke AR, Sadikova E et al. Development and validation of a machine learning individualized treatment rule in first-episode schizophrenia. JAMA Netw Open 2020;3:e1921660.
100. Pradier MF, McCoy TH Jr, Hughes M et al. Predicting treatment dropout after antidepressant initiation. Transl Psychiatry 2020;10:60.
101. Hughes MC, Pradier MF, Ross AS et al. Assessment of a prediction model for antidepressant treatment stability using supervised topic models. JAMA Netw Open 2020;3:e205308.
102. Pradier MF, Hughes MC, McCoy TH et al. Predicting change in diagnosis from major depression to bipolar disorder after antidepressant initiation. Neuropsychopharmacology 2021;46:455-61.
103. Huang SH, LePendu P, Iyer SV et al. Toward personalizing treatment for depression: Predicting diagnosis and severity. J Am Med Inform Assoc 2014;21:1069-75.
104. Hallgren KA, Bauer AM, Atkins DC. Digital technology and clinical decision making in depression treatment: current findings and future opportunities. Depress Anxiety 2017;34:494-501.
105. Willetts M, Hollowell S, Aslett L et al. Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. Sci Rep 2018;8:7961.
106. Gravenhorst F, Muaremi A, Bardram J et al. Mobile phones as medical devices in mental disorder treatment: an overview. Pers Ubiquitous Comput 2015;19:335-53.
107. Saeb S, Lattie EG, Schueller SM et al. The relationship between mobile phone location sensor data and depressive symptom severity. PeerJ 2016;4:e2537.

108. Thakur SS, Roy RB. Predicting mental health using smart-phone usage and sensor data. J Ambient Intell Humaniz Comput (in press)..
109. Germine L, Reinecke K, Chaytor NS. Digital neuropsychology: challenges and opportunities at the intersection of science and software. Clin Neuropsychol 2019;33:271-86.
110. Passell E, Dillon GT, Baker JT et al. Digital cognitive assessment: results from the TestMyBrain NIMH Research Domain Criteria (RDoC) Field Test Battery Report. PsyArXiv 2019;10.31234.
111. Hoogendoorn M, Berger T, Schulz A et al. Predicting social anxiety treatment outcome based on therapeutic email conversations. IEEE J Biomed Health Inform 2017;21:1449-59.
112. Bremer V, Funk B, Riper H. Heterogeneity matters: predicting self-esteem in online interventions based on ecological momentary assessment data. Depress Res Treat 2019;2019:3481624.
113. Riihimäki H, Chachólski W, Theorell J et al. A topological data analysis based classification method for multiple measurements. BMC Bioinformatics 2020;21:336.
114. Gillan CM, Rutledge RB. Smartphones and the neuroscience of mental health. Annu Rev Neurosci (in press).
115. Vivian-Griffiths T, Baker E, Schmidt KM et al. Predictive modeling of schizophrenia from genomic data: comparison of polygenic risk score with kernel support vector machines approach. Am J Med Genet B Neuropsychiatr Genet 2019;180:80-5.
116. Kourou K, Exarchos TP, Exarchos KP et al. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 2015;13:8-17.
117. Chabon JJ, Hamilton EG, Kurtz DM et al. Integrating genomic features for non-invasive early lung cancer detection. Nature 2020;580:245-51.
118. Iniesta R, Campbell D, Venturini C et al. Gene variants at loci related to blood pressure account for variation in response to antihypertensive drugs between black and white individuals: genomic precision medicine may dispense with ethnicity. Hypertension 2019;74:614-22.
119. Tansey KE, Guipponi M, Hu X et al. Contribution of common genetic variants to antidepressant response. Biol Psychiatry 2013;73:679-82.
120. Lawrie SM. Clinical risk prediction in schizophrenia. Lancet Psychiatry 2014; 1:406-8.
121. Zheutlin AB, Ross DA. Polygenic risk scores: what are they good for? Biol Psychiatry 2018;83:e51-3.
122. Amare AT, Schubert KO, Tekola-Ayele F et al. Association of the polygenic scores for personality traits and response to selective serotonin reuptake inhibitors in patients with major depressive disorder. Front Psychiatry 2018;9:65.
123. García-González J, Tansey KE, Hauser J et al. Pharmacogenetics of antidepressant response: a polygenic approach. Prog Neuropsychopharmacol Biol Psychiatry 2017;75:128-34.
124. Ward J, Graham N, Strawbridge R et al. Polygenic risk scores for major depressive disorder and neuroticism as predictors of antidepressant response: meta-analysis of three treatment cohorts. bioRxiv 2018:295717.
125. Andersson E, Crowley JJ, Lindefors N et al. Genetics of response to cognitive behavior therapy in adults with major depression: a preliminary report. Mol Psychiatry 2019;24:484-90.
126. Li QS, Tian C, Seabrook GR et al. Analysis of 23andMe antidepressant efficacy survey data: implication of circadian rhythm and neuroplasticity in bupropion response. Transl Psychiatry 2016;6:e889.
127. Wigmore EM, Hafferty JD, Hall LS et al. Genome-wide association study of antidepressant treatment resistance in a population-based cohort using health service prescription data and meta-analysis with GENDEP. Pharmacogenomics J 2020;20:329-41.
128. Baune BT, Soda T, Sullivan PF et al. The Genomics of Electroconvulsive Therapy International Consortium (GenECT-ic). Lancet Psychiatry 2019;6:e23.
129. Lin E, Lin C-H, Lane H-Y. Precision psychiatry applications with pharmacogenomics: artificial intelligence and machine learning approaches. Int J Mol Sci 2020;21:969.
130. Kautzky A, Baldinger P, Souery D et al. The combined effect of genetic polymorphisms and clinical parameters on treatment outcome in treatment-resistant depression. Eur Neuropsychopharmacol 2015;25:441-53.
131. Maciukiewicz M, Marshe VS, Hauschild A-C et al. GWAS-based machine learning approach to predict duloxetine response in major depressive disorder. J Psychiatr Res 2018;99:62-8.
132. Athreya AP, Neavin D, Carrillo-Roa T et al. Pharmacogenomics-driven prediction of antidepressant treatment outcomes: a machine-learning approach with multi-trial replication. Clin Pharmacol Ther 2019;106:855-65.
133. Lin E, Kuo P-H, Liu Y-L et al. A deep learning approach for predicting antidepressant response in major depression using clinical and genetic biomarkers. Front Psychiatry 2018;9:290.
134. Lee Y, Ragguett RM, Mansur RB et al. Applications of machine learning algorithms to predict therapeutic outcomes in depression: a meta-analysis and systematic review. J Affect Disord 2018;241:519-32.
135. Pisanu C, Squassina A. Treatment-resistant schizophrenia: insights from genetic studies and machine learning approaches. Front Pharmacol 2019; 10:1-7.
136. Zubenko GS, Sommer BR, Cohen BM. On the marketing and use of pharmacogenetic tests for psychiatric treatment. JAMA Psychiatry 2018;75:769.
137. Zeier Z, Carpenter LL, Kalin NH et al. Clinical implementation of pharmacogenetic decision support tools for antidepressant drug prescribing. Am J Psychiatry 2018;175:873-86.
138. Enneking V, Leehr EJ, Dannlowski U et al. Brain structural effects of treatments for depression and biomarkers of response: a systematic review of neuroimaging studies. Psychol Med 2020;50:187-209.
139. Colvonen PJ, Glassman LH, Crocker LD et al. Pretreatment biomarkers predicting PTSD psychotherapy outcomes: a systematic review. Neurosci Biobehav Rev 2017;75:140-56.
140. Fonseka TM, MacQueen GM, Kennedy SH. Neuroimaging biomarkers as predictors of treatment outcome in major depressive disorder. J Affect Disord 2018;233:21-35.
141. Fu CH, Steiner H, Costafreda SG. Predictive neural biomarkers of clinical response in depression: a meta-analysis of functional and structural neuroimaging studies of pharmacological and psychological therapies. Neurobiol Dis 2013;52:75-83.
142. Lueken U, Zierhut KC, Hahn T et al. Neurobiological markers predicting treatment response in anxiety disorders: a systematic review and implications for clinical application. Neurosci Biobehav Rev 2016;66:143-62.
143. Molent C, Olivo D, Wolf RC et al. Functional neuroimaging in treatment resistant schizophrenia: a systematic review. Neurosci Biobehav Rev 2019; 104:178-90.
144. Tarcijonas G, Sarpal DK. Neuroimaging markers of antipsychotic treatment response in schizophrenia: an overview of magnetic resonance imaging studies. Neurobiol Dis 2019;131:104209.
145. Phillips ML, Chase HW, Sheline YI et al. Identifying predictors, moderators, and mediators of antidepressant response in major depressive disorder: neuroimaging approaches. Am J Psychiatry 2015;172:124-38.
146. Lueken U, Hahn T. Functional neuroimaging of psychotherapeutic processes in anxiety and depression: from mechanisms to predictions. Curr Opin Psychiatry 2016;29:25-31.
147. Khodayari-Rostamabad A, Hasey GM, Maccrimmon DJ et al. A pilot study to determine whether machine learning methodologies using pre-treatment electroencephalography can predict the symptomatic response to clozapine therapy. Clin Neurophysiol 2010;121:1998-2006.
148. Erguzel TT, Ozekes S, Gultekin S et al. Neural network based response prediction of rTMS in major depressive disorder using QEEG cordance. Psychiatry Investig 2015;12:61-5.
149. Mumtaz W, Xia L, Mohd Yasin MA et al. A wavelet-based technique to predict treatment outcome for major depressive disorder. PLoS One 2017; 12:e0171409.
150. Jaworska N, de la Salle S, Ibrahim MH et al. Leveraging machine learning approaches for predicting antidepressant treatment response using electroencephalography (EEG) and clinical data. Front Psychiatry 2018;9:768.
151. Khodayari-Rostamabad A, Reilly JP, Hasey GM et al. Using pre-treatment electroencephalography data to predict response to transcranial magnetic stimulation therapy for major depression. Annu Int Conf IEEE Eng Med Biol Soc 2011;2011:6418-21.
152. Bailey NW, Hoy KE, Rogasch NC et al. Differentiating responders and non-responders to rTMS treatment for depression after one week using resting EEG connectivity measures. J Affect Disord 2019;242:68-79.
153. Hasanzadeh F, Mohebbi M, Rostami R. Prediction of rTMS treatment response in major depressive disorder using machine learning techniques and nonlinear features of EEG signal. J Affect Disord 2019;256:132-42.
154. Zandvakili A, Philip NS, Jones SR et al. Use of machine learning in predicting clinical response to transcranial magnetic stimulation in comorbid posttraumatic stress disorder and major depression: a resting state electroencephalography study. J Affect Disord 2019;252:47-54.
155. Al-Kaysi AM, Al-Ani A, Loo CK et al. Predicting tDCS treatment outcomes of patients with major depressive disorder using automated EEG classification. J Affect Disord 2017;208:597-603.
156. Corlier J, Carpenter LL, Wilson AC et al. The relationship between individual alpha peak frequency and clinical outcome with repetitive transcranial magnetic stimulation (rTMS) treatment of major depressive disorder (MDD). Brain Stimul 2019;12:1572-8.

157. Woo CW, Chang LJ, Lindquist MA et al. Building better biomarkers: brain models in translational neuroimaging. Nat Neurosci 2017;20:365-77.
158. Costafreda SG, Khanna A, Mourao-Miranda J et al. Neural correlates of sad faces predict clinical remission to cognitive behavioural therapy in depression. Neuroreport 2009;20:637-41.
159. Goldstein-Piekarski AN, Korgaonkar MS, Green E et al. Human amygdala engagement moderated by early life stress exposure is a biobehavioral target for predicting recovery on antidepressants. Proc Natl Acad Sci USA 2016;113:11955-60.
160. Hahn T, Kircher T, Straube B et al. Predicting treatment response to cognitive behavioral therapy in panic disorder with agoraphobia by integrating local neural information. JAMA Psychiatry 2015;72:68-74.
161. Ball TM, Stein MB, Ramsawh HJ et al. Single-subject anxiety treatment outcome prediction using functional neuroimaging. Neuropsychopharmacology 2014;39:1254-61.
162. Tolmeijer E, Kumari V, Peters E et al. Using fMRI and machine learning to predict symptom improvement following cognitive behavioural therapy for psychosis. Neuroimage Clin 2018;20:1053-61.
163. Crane NA, Jenkins LM, Bhaumik R et al. Multidimensional prediction of treatment response to antidepressants with cognitive control and functional MRI. Brain 2017;140:472-86.
164. Nguyen KP, Fatt CC, Treacher A et al. Predicting response to the antidepressant bupropion using pretreatment fMRI. Predict Intell Med 2019;11843:53-62.
165. van Waarde JA, Scholte HS, van Oudheusden LJ et al. A functional MRI marker may predict the outcome of electroconvulsive therapy in severe and treatment-resistant depression. Mol Psychiatry 2015;20:609-14.
166. Whitfield-Gabrieli S, Ghosh SS, Nieto-Castanon A et al. Brain connectomics predict response to treatment in social anxiety disorder. Mol Psychiatry 2016;21:680-5.
167. Zhutovsky P, Thomas RM, Olff M et al. Individual prediction of psychotherapy outcome in posttraumatic stress disorder using neuroimaging data. Transl Psychiatry 2019;9:326.
168. Yuan M, Qiu C, Meng Y et al. Pre-treatment resting-state functional MR imaging predicts the long-term clinical outcome after short-term paroxetine treatment in post-traumatic stress disorder. Front Psychiatry 2018;9:532.
169. Cao B, Cho RY, Chen D et al. Treatment response prediction and individualized identification of first-episode drug-naïve schizophrenia using brain functional connectivity. Mol Psychiatry 2020;25:906-13.
170. Leaver AM, Wade B, Vasavada M et al. Fronto-temporal connectivity predicts ECT outcome in major depression. Front Psychiatry 2018;9:92.
171. Redlich R, Opel N, Grotegerd D et al. Prediction of individual response to electroconvulsive therapy via machine learning on structural magnetic resonance imaging data. JAMA Psychiatry 2016;73:557-64.
172. Cao B, Luo Q, Fu Y et al. Predicting individual responses to the electroconvulsive therapy with hippocampal subfield volumes in major depression disorder. Sci Rep 2018;8:5434-4.
173. Gong J, Cui LB, Xi YB et al. Predicting response to electroconvulsive therapy combined with antipsychotics in schizophrenia using multi-parametric magnetic resonance imaging. Schizophr Res 2020;216:262-71.
174. Germine L, Strong RW, Singh S et al. Toward dynamic phenotypes and the scalable measurement of human behavior. Neuropsychopharmacology 2021;46:209-16.
175. Etkin A, Patenaude B, Song YJ et al. A cognitive-emotional biomarker for predicting remission with antidepressant medications: a report from the iSPOT-D trial. Neuropsychopharmacology 2015;40:1332-42.
176. Rennie JP, Zhang M, Hawkins E et al. Mapping differential responses to cognitive training using machine learning. Dev Sci 2020;23:e12868.
177. Delgadillo J, Gonzalez Salas Duhne P. Targeted prescription of cognitive–behavioral therapy versus person-centered counseling for depression using a machine learning approach. J Consult Clin Psychol 2020;88:14-24.
178. Schwartz B, Cohen ZD, Rubel JA et al. Personalized treatment selection in routine care: integrating machine learning and statistical algorithms to recommend cognitive behavioral or psychodynamic therapy. Psychother Res 2021;31:33-51.
179. Delgadillo J, Ali S, Fleck K et al. StratCare: a pragmatic, multi-site, single-blind, cluster randomised controlled trial of stratified care for depression. Unpublished manuscript.
180. Lutz W, Deisenhofer AK, Rubel J et al. Prospective evaluation of a clinical decision support system in psychological therapy. Unpublished manuscript.
181. Browning M, Bilderbeck AC, Dias R et al. The clinical effectiveness of using a predictive algorithm to guide antidepressant treatment in primary care (PReDicT): an open-label , randomised controlled trial. Neuropsychopharmacology (in press).
182. Jussim L, Harber KD. Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies. Personal Soc Psychol Rev 2005;9:131-55.
183. Patel SR, Bakken S, Ruland C. Recent advances in shared decision making for mental health. Curr Opin Psychiatry 2008;21:606-12.
184. Lundberg SM, Lee S-I. A Unified approach to interpreting model predictions. arXiv 2017;1705.07874v2.
185. Chekroud AM. Pragmatic, scalable, computational solutions to reduce the burden of major depression. ProQuest 2018;10907745.
186. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 2019;1:206-15.
187. 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 (in press).
188. Jacobs M, Pradier MF, McCoy TH Jr et al. How machine-learning recommendations influence clinician treatment selections: the example of the antidepressant selection. Transl Psychiatry 2021;11:108.
189. Delgadillo J, Appleby S, Booth S et al. The Leeds Risk Index: field-test of a stratified psychological treatment selection algorithm. Psychother Psychosom 2020;89:189-90.
190. Müller VC. Ethics of artificial intelligence and robotics. https://plato.stanford.edu/entries/ethics-ai/.
191. Lo Piano S. Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward. Humanit Soc Sci Commun 2020;7:9.
192. Mittelstadt BD, Floridi L. The ethics of big data: current and foreseeable issues in biomedical contexts. Sci Eng Ethics 2016;22:303-41.
193. Starke G, De Clercq E, Borgwardt S et al. Computing schizophrenia: ethical challenges for machine learning in psychiatry. Psychol Med (in press).
194. Trujillo AC, Gregory IM, Ackerman KA. Evolving relationship between humans and machines. IFAC-PapersOnLine 2019;51:366-71.
195. Endsley MR. Toward a theory of situation awareness in dynamic systems. Hum Factors 1995;37:32-64.
196. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25:44-56.
197. Haque OS, Waytz A. Dehumanization in medicine: causes, solutions, and functions. Perspect Psychol Sci 2012;7:176-86.
198. Chen N-C, Drouhard M, Kocielnik R et al. Using machine learning to support qualitative coding in social science: shifting the focus to ambiguity. ACM Trans Interact Intell Syst 2018;8:1-20.
199. Stewart M. Towards a global definition of patient centred care: the patient should be the judge of patient centred care. BMJ 2001;322:444-5.
200. Robin R, Polanyi M. Personal knowledge. towards a post-critical philosophy. Philos Phenomenol Res 1960;20:429.
201. Bennett NL. Donald A. Schön, educating the reflective practitioner. San Francisco: Jossey-Bass, 1987.
202. Borrell-Carrió F, Suchman A, Epstein RM. The biopsychosocial model 25 years later: principles, practice, and scientific inquiry. Ann Fam Med 2004;2:576-82.
203. Faden RR, Kass NE, Goodman SN et al. An ethics framework for a learning health care system: a departure from traditional research ethics and clinical ethics. Hastings Cent Rep 2013;43:S16-27.
204. Boff KR. Revolutions and shifting paradigms in human factors & ergonomics. Appl Ergon 2006;37:391-9.
205. Brangier E, Hammes-Adelé S. Beyond the technology acceptance model: elements to validate the human-technology symbiosis model. In: Robertson MM (eds). Ergonomics and health aspects of work with computers. Berlin: Springer, 2011:13-21.
206. Christophe G, Jean-Arthur M-F, Guillaume D. Comment on Starke et al.: 'Computing schizophrenia: ethical challenges for machine learning in psychiatry': from machine learning to student learning: pedagogical challenges for psychiatry. Psychol Med (in press).
207. Gerber A, Derckx P, Döppner DA et al. Conceptualization of the human-machine symbiosis – A literature review. Presented at the 53rd Hawaii International Conference on System Sciences, Maui, January 2020.
208. Floridi L, Taddeo M. What is data ethics? Philos Trans R Soc Math Phys Eng Sci 2016;374:2083.
209. Binns R. Fairness in machine learning: lessons from political philosophy. Proceedings of Machine Learning Research 2018;81:149-59.

210. Obermeyer Z, Powers B, Vogeli C et al. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019;366:447-53.
211. Yeung K, Lodge M (eds). Algorithmic regulation. Oxford: Oxford University Press, 2019.
212. Brownsword R, Scotford E, Yeung K (eds). The Oxford handbook of law, regulation and technology, Vol. 1. Oxford: Oxford University Press, 2016.
213. Jacucci G, Spagnolli A, Freeman J et al. Symbiotic interaction: a critical definition and comparison to other human-computer paradigms. Presented at the International Workshop on Symbiotic Interaction, Helsinki, October 2014.
214. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010;51:214-21.
215. Lustria MLA, Noar SM, Cortese J et al. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18:1039-69.
216. Bird S, Kenthapadi K, Kiciman E et al. Fairness-aware machine learning: practical challenges and lessons learned. Presented at the 12th ACM International Conference on Web Search and Data Mining, Melbourne, February 2019.
217. Koutsouleris N, Dwyer DB, Degenhardt F et al. Multimodal machine learning workflows for prediction of psychosis in patients with clinical high-risk syndromes and recent-onset depression. JAMA Psychiatry 2021;78:195-209.

DOI:10.1002/wps.20882

Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): II. Externalizing superspectrum

Robert F. Krueger[1], Kelsey A. Hobbs[1], Christopher C. Conway[2], Danielle M. Dick[3], Michael N. Dretsch[4], Nicholas R. Eaton[5], Miriam K. Forbes[6], Kelsie T. Forbush[7], Katherine M. Keyes[8], Robert D. Latzman[9], Giorgia Michelini[10], Christopher J. Patrick[11], Martin Sellbom[12], Tim Slade[13], Susan C. South[14], Matthew Sunderland[13], Jennifer Tackett[15], Irwin Waldman[16], Monika A. Waszczuk[17], Aidan G.C. Wright[18], David H. Zald[19], David Watson[20], Roman Kotov[17]; HiTOP Utility Workgroup*

[1]Department of Psychology, University of Minnesota, Minneapolis, MN, USA; [2]Department of Psychology, Fordham University, Bronx, NY, USA; [3]Department of Psychology, Virginia Commonwealth University, Richmond, VA, USA; [4]US Army Medical Research Directorate - West, Walter Reed Army Institute of Research, Joint Base Lewis-McChord, WA, USA; [5]Department of Psychology, Stony Brook University, Stony Brook, NY, USA; [6]Centre for Emotional Health, Department of Psychology, Macquarie University, Sydney, NSW, Australia; [7]Department of Psychology, University of Kansas, Lawrence, KS, USA; [8]Department of Epidemiology, Columbia University, New York, NY, USA; [9]Department of Psychology, Georgia State University, Atlanta, GA, USA; [10]Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, CA, USA; [11]Department of Psychology, Florida State University, Tallahassee, FL, USA; [12]Department of Psychology, University of Otago, Dunedin, New Zealand; [13]Matilda Centre for Research in Mental Health and Substance Use, University of Sydney, Sydney, NSW, Australia; [14]Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA; [15]Department of Psychology, Northwestern University, Evanston, IL, USA; [16]Department of Psychology, Emory University, Atlanta, GA, USA; [17]Department of Psychiatry, Stony Brook University, Stony Brook, NY, USA; [18]Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA; [19]Department of Psychology, Vanderbilt University, Nashville, TN, USA; [20]Department of Psychology, University of Notre Dame, Notre Dame, IN, USA

*Members of HiTOP Utility Workgroup are listed in the Appendix

The Hierarchical Taxonomy of Psychopathology (HiTOP) is an empirical effort to address limitations of traditional mental disorder diagnoses. These include arbitrary boundaries between disorder and normality, disorder co-occurrence in the modal case, heterogeneity of presentation within disorders, and instability of diagnosis within patients. This paper reviews the evidence on the validity and utility of the disinhibited externalizing and antagonistic externalizing spectra of HiTOP, which together constitute a broad externalizing superspectrum. These spectra are composed of elements subsumed within a variety of mental disorders described in recent DSM nosologies, including most notably substance use disorders and "Cluster B" personality disorders. The externalizing superspectrum ranges from normative levels of impulse control and self-assertion, to maladaptive disinhibition and antagonism, to extensive polysubstance involvement and personality psychopathology. A rich literature supports the validity of the externalizing superspectrum, and the disinhibited and antagonistic spectra. This evidence encompasses common genetic influences, environmental risk factors, childhood antecedents, cognitive abnormalities, neural alterations, and treatment response. The structure of these validators mirrors the structure of the phenotypic externalizing superspectrum, with some correlates more specific to disinhibited or antagonistic spectra, and others relevant to the entire externalizing superspectrum, underlining the hierarchical structure of the domain. Compared with traditional diagnostic categories, the externalizing superspectrum conceptualization shows improved utility, reliability, explanatory capacity, and clinical applicability. The externalizing superspectrum is one aspect of the general approach to psychopathology offered by HiTOP and can make diagnostic classification more useful in both research and the clinic.

Key words: HiTOP, externalizing, disinhibition, antagonism, antisocial personality disorder, Cluster B personality disorders, substance use disorders, clinical utility

(World Psychiatry 2021;20:171–193)

The Hierarchical Taxonomy of Psychopathology (HiTOP) consortium aims to integrate research on the empirical organization of psychopathology, with the goal of delineating a comprehensive descriptive system[1-3]. Taxonomies in frequent use (e.g., the DSM) have notable limitations, such as arbitrary boundaries between psychopathology and normality, diagnostic instability, heterogeneity within disorders, disorder co-occurrence in the modal case, and inability to conceptualize subthreshold cases. The HiTOP approach mitigates such problems by: a) defining psychopathology in terms of continua ranging from normative to maladaptive; b) delineating continua based on observed covariation among signs, symptoms and syndromes, and c) arranging continua in a hierarchy, ranging from more narrow and specific (e.g., clusters of symptoms) to more broad and general (e.g., spectra of inter-related diagnostic phenomena).

An approach based on continua or dimensions of human individual differences resolves issues of arbitrary thresholds and diagnostic instability. Thresholds indicating specific clinical options can be described based on evidence, and test-retest reliability of dimensional psychopathology constructs is notably greater than that of arbitrary diagnostic categories[4-7]. No patients are excluded from the system (i.e., individuals with subthreshold or atypical symptoms are all characterized on a set of dimensions), providing a boon to case conceptualization. The HiTOP approach also reduces diagnostic heterogeneity by grouping empirically related symptoms together and arraying them along distinguishable dimensions[8-11]. Comorbidity is rendered understandable, because related conditions form elements in psychologically coherent spectrums.

The working HiTOP system currently includes six broad spectrums: internalizing, somatoform, disinhibited externalizing, antagonistic externalizing, thought disorder, and detachment[1-3]. These spectrums reflect continuous individual differences in a given domain across the entire population. Broad spectrums, in turn, are combined into larger groupings or superspectra: emotional dysfunction (internalizing and somatoform), externalizing (disinhibited and antagonistic), and psychosis (thought disorder and detachment)[12-16]. Above these superspectra, the HiTOP approach also recognizes a general psychopathology factor[17,18].

The working HiTOP system was created by reviewing a con-

siderable body of research, but external validity and utility have been less well documented, because previous reviews of these topics had limited scope. With this in mind, the Utility Workgroup of the HiTOP consortium assembled teams of experts to comprehensively review evidence on the validity and utility of the working HiTOP model. Expert reviews were organized according to the three superspectra. The present paper is the second in this series (the first focused on psychosis[19] and the third will examine emotional dysfunction) and focuses on the externalizing superspectrum.

The externalizing superspectrum encompasses two spectra: disinhibited externalizing and antagonistic externalizing. The disinhibited externalizing spectrum includes tendencies to act on impulse, without consideration for potential consequences. Empirically, disinhibition tends to be accompanied by societally prohibited behaviors that align psychologically with the core of the construct, for example, the use of psychoactive substances to excess[20] and with minimal regard for future consequences. The antagonistic externalizing spectrum includes tendencies to navigate interpersonal situations using antipathy and conflict, and to hurt other people intentionally[21], with little regard for their rights and feelings.

These spectra encompass both maladaptive traits and more time-limited symptoms, with the distinction pertaining to the timescale of the phenomena[22]. For example, a series of specific disinhibited behaviors (e.g., a brief period encompassing impulsive purchases and other decisions that reflect immediate reward more than longer-term consequences) could be driven by a specific life crisis, rather than being generally characteristic of a person. If such behaviors persist across time and circumstances, they become additionally indicative of a disinhibitory trait. Similarly, a specific hostile interaction is an antagonistic phenomenon, while frequent and recurrent hostile interactions are indicative of an antagonistic trait. As described at length throughout this review, disinhibited and antagonistic behaviors tend to co-occur at notably greater than chance levels, illustrating the phenotypic coherence of the broad externalizing superspectrum[23].

The goal of this paper is to review the extensive evidence documenting the structural coherence and content of the externalizing superspectrum and the disinhibited and antagonistic spectra, and the utility and validity of these diagnostic constructs.

STRUCTURAL EVIDENCE

Composition of major dimensions

The externalizing superspectrum has long emerged in research on the structure of mental disorders and of maladaptive personality traits. Indeed, studies have revealed that externalizing psychopathology is separate from other superspectra, including internalizing psychopathology in youth[24-30] and both internalizing and thought disorder/psychosis in adults[31-34]. Across these bodies of research, clinical diagnoses or dimensional symptom counts of antisocial personality disorder (PD), attention-deficit/hyperactivity disorder (ADHD), alcohol, cannabis, nicotine, and other substance use disorders (SUDs), and intermittent explosive disorder in adulthood, as well as conduct disorder (CD) and oppositional defiant disorder (ODD) in childhood, clearly reflect a distinct and overarching externalizing superspectrum, as summarized in Table 1 and pictured in Figure 1.

The extant evidence further supports parsing the externalizing superspectrum down into disinhibited and antagonistic externalizing spectra[1]. This bifurcation is more clearly evident in maladaptive trait research and in the adult rather than the child psychopathology literature, and these major domains can be observed in the psychiatric diagnosis literature as well.

As summarized in Table 1, three main observations are evident from this literature. First, the majority of studies identify antisocial PD as an indicator of both disinhibition and antagonism, which supports this disorder as a non-specific and core indicator of the general externalizing superspectrum. In fact, the criteria for antisocial PD are quite evenly spread across both disinhibited and antagonistic features. Second, alcohol and other SUDs are specific to the disinhibited externalizing spectrum. Third, some DSM PDs (i.e., paranoid, narcissistic and histrionic) appear relatively specific to the antagonistic externalizing spectrum. These findings are also generally consistent with Krueger et al's multifaceted model of the externalizing spectrum[15], which considers general externalizing together with more specific liability factors for callous-aggression (the unique component of antagonism) and substance misuse (the unique component of disinhibited externalizing).

One condition deserving specific consideration is borderline PD, as its relevance to general externalizing, as well as its specificity to antagonism vs. disinhibition, appears dependent on other indicators included in the structural model. In studies in which internalizing psychopathology is also prominently featured, borderline PD tends to load robustly with internalizing and less consistently with externalizing[39,46,48,52]; moreover, when dimensional traits are considered in addition to psychiatric diagnoses, this PD loads distinctly on internalizing[64]. In other words, the preponderance of research evidence indicates that borderline PD does load with the internalizing spectrum, while its association with externalizing (and even specific placement within antagonism vs. disinhibition)[65,66] is less clear. At this point, borderline PD is therefore best considered an indicator of both internalizing and, to a lesser degree, the general externalizing superspectrum, likely with different components of the disorder being related to these two spectra. As such, borderline PD is only provisionally included in the externalizing superspectrum, as noted in Figure 1.

It is further noteworthy that, while clearly representing antagonistic externalizing in the context of the broader externalizing superspectrum, paranoid and histrionic PD have other influences as well, given their multifaceted nature. For instance, paranoid PD may appear more strongly linked to the psychosis superspectrum when disorders of this type are clearly represented in the set of structural indicators[13,32,65]. Histrionic PD also has direct links (in the negative direction) to the detachment spectrum[64], which is also supported in the general personality literature[67,68].

Finally, although the externalizing superspectrum is well rep-

Table 1 Structural evidence on the externalizing superspectrum and the disinhibited and antagonistic spectra

	Sample size	Sample type	ASPD	AUD	Other SUD	IED	CD	ODD	ADHD	NPD	PPD	HPD	BPD
Externalizing superspectrum													
Dunedin Multidisciplinary Health and Development Study (Caspi et al[31], Krueger et al[35])	1,037	Community/longitudinal	+	+	+		+						
Early Developmental Stages of Psychopathology (Beesdo-Baum et al[36], Wittchen et al[37])	3,021	Community/longitudinal	+	+	+								
NESARC waves 1 and/or 2 (Carragher et al[38], Eaton et al[39], Keyes et al[32], Lahey et al[40])	43,093 & 34,653	Community/adults	+	+	+				+		–		–
Tennessee Twin Study (Lahey et al[27], Waldman et al[41])	4,050	Community/children & adolescents					+	+	+				
WMH Surveys (de Jonge et al[42], Kessler et al[43])	21,229	Community/ longitudinal		+	+	+	+	+	+				
Blanco et al[24]	9,244	Community/adolescents		+	+		+	+	+				
Castellanos-Ryan et al[25]	2,232	Community/adolescents		+	+		+	+	+				
Conway et al[44]	25,002	University/adults		+	+				+				
Cox et al[45]	5,877	Community/adults	+	+	+								
Forbush & Watson[46]	16,423	Community/adults	+	+	+	+	+		+				+
Forbush et al[47]	1,434	Community/longitudinal	+	+	+		+						
Gomez et al[26]	2,099	Outpatient/youth					+	+	+				
James & Taylor[48]	1,197	Community/adults	+	+	+								–
Krueger[49]	8,098	Community/adults	+	+	+								
Krueger et al[50]	1,048	Community/adolescents		+	+		+						
Martel et al[28]	2,512	Community/children					+	+	+				
Martel et al[28]	8,012	Community/adults		+	+		+	+	+				
Miller et al[51]	1,325	Veterans/adults	+	+	+								
Miller et al[52]	214	Veterans/adults	+		–		+						+
Olino et al[29]	541	Community/children						+	+				
Tuvblad et al[53]	1,219	Community/children					+	+	+				
Verona et al[33]	4,745	Community/adults	+	+	+								
Verona et al[30]	223	Mixed/youth		–	–		+	+	+				
Young et al[54]	668	Community/adolescents			+		+		+				
Total positive			11/11	16/17	17/19	2/2	15/15	10/10	14/14	0/0	0/1	0/0	2/4

Table 1 Structural evidence on the externalizing superspectrum and the disinhibited and antagonistic spectra (*continued*)

	Sample size	Sample type	ASPD	AUD	Other SUD	IED	CD	ODD	ADHD	NPD	PPD	HPD	BPD
Disinhibited spectrum													
MIDAS (Forbes et al[13], Kotov et al[55])	2,900	Outpatients/adults	+	+	+		+			–	–	–	–
Norwegian Institute of Public Health Twin Panel (Kendler et al[56], Røysamb et al[57])	2,794	Community/adults	+	+	+		+			–	–	–	+
Conway & Brown[58]	4,928	Outpatients/adults		+	+								
Conway et al[59]	815	Community/longitudinal		+	+								
Farmer et al[60]	816	Community/longitudinal	+	+	+		+						
Kim & Eaton[61]	43,093	Community/adults	+	+	+								
Slade & Watson[62]	10,641	Community/adults		+	+								
Vollebergh et al[63]	7,076	Community/adults		+	+								
Wright & Simms[64]	628	Outpatients/adults	+	+	+					–	–	–	–
Wright et al[34]	8,841	Community/adults		+	+								
Total positive			5/5	10/10	10/10		3/3			0/3	0/3	0/3	1/3
Antagonistic spectrum													
MIDAS (Forbes et al[13], Kotov et al[55])	2,900	Outpatients/adults	+	–	–		+			+	+	+	+
Norwegian Institute of Public Health Twin Panel (Kendler et al[56], Røysamb et al[57])	2,794	Community/adults	–	–	–		–			+	+	+	+
Farmer et al[60]	816	Community/longitudinal	–	–	–		–	+	+				
Kim & Eaton[61]	43,093	Community/adults	+	–	–								
Wright & Simms[64]	628	Outpatients/adults	–	–	–					+	+	+	–
Total positive			2/5	0/5	0/5		1/3	1/1	1/1	3/3	3/3	3/3	2/3

+: indicator included in analysis with meaningful loading (.30 or larger), –: indicator included in analysis but did not load meaningfully. ASPD – antisocial personality disorder, AUD – alcohol use disorder, SUD – substance use disorder, IED – intermittent explosive disorder, CD – conduct disorder, ODD – oppositional defiant disorder, ADHD – attention-deficit/hyperactivity disorder, NPD – narcissistic personality disorder, PPD – paranoid personality disorder, HPD – histrionic personality disorder, BPD – borderline personality disorder, NESARC – National Epidemiologic Survey on Alcohol and Related Conditions, WMH – World Mental Health, MIDAS – Methods to Improve Diagnostic Assessment and Services.

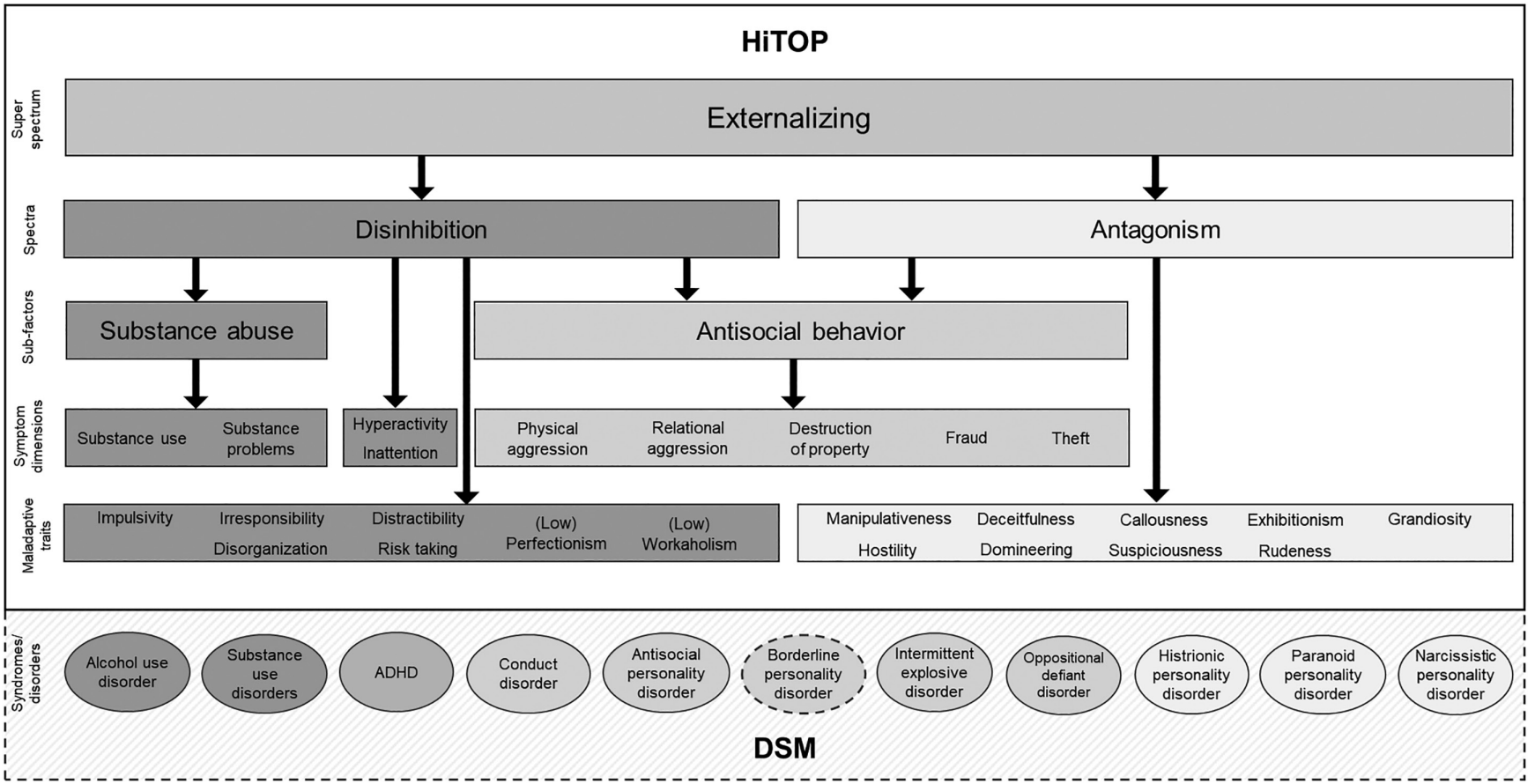

Figure 1 Conceptual model of the externalizing superspectrum. Dashed lines represent provisional inclusion. Specifically, the dashed line surrounding borderline personality disorder represents that this disorder falls under two superspectra (externalizing and internalizing). The dashed line surrounding the diagnoses section indicates that these categorical diagnoses do not belong to the model; they are meant to represent how a dimensional model encapsulates DSM diagnoses. HiTOP – Hierarchical Taxonomy Of Psychopathology, ADHD – attention-deficit/hyperactivity disorder.

resented in the youth psychopathology literature, evidence for bifurcation of disinhibition and antagonism prior to adulthood is less clear (see Table 1), likely owing to the lack of clearly defined indicators for making this separation. In contrast with the adult literature, there are no diagnoses or explicit symptom measures of callous-unemotional traits, narcissism, or paranoia/suspiciousness included in structural modeling studies with children/adolescents, making it virtually impossible for these factors to emerge in the youth literature. Additionally, in young children, substance use is likely to be uncommon. Furthermore, links between personality traits and disorders are less well established in the youth (especially child) literature[69], making an analysis from this perspective less straightforward. Further research is definitely needed to obtain a clearer picture of psychiatric representations of antagonism in youth, especially beyond what are typically referred to as callous-unemotional traits[70].

Role of personality traits

The hierarchical structure of the externalizing superspectrum closely parallels the organization of normal-range personality traits[1,71]. The general externalizing dimension is broadly linked to individual differences in the higher-order trait factor of constraint vs. disconstraint[72], which emerges in three-factor models of normal and abnormal personality[72-75]. When additional factors are extracted, this broad constraint vs. disconstraint dimension divides into two more specific components: agreeableness vs. antagonism, and conscientiousness vs. disinhibition[73,75,76]. These two subdimensions, in turn, form the basis for distinguishing antagonistic from disinhibited forms of externalizing.

Antagonistic externalizing has been linked to a variety of specific maladaptive traits that reflect problematic relations with others. It should be noted that some of these traits also show lesser associations with other forms of psychopathology[1,77-83]. The traits that have been most strongly and consistently associated with the antagonistic externalizing spectrum include manipulativeness (i.e., exploiting and taking advantage of others), deceitfulness (i.e., lying and cheating in pursuit of one's goals), callousness (i.e., being cold-hearted and lacking empathy), exhibitionism (i.e., engaging in attention-seeking behaviors), grandiosity (i.e., being arrogant and feeling entitled to special treatment from others), aggression (i.e., engaging in hostile and even violent behavior), rudeness (i.e., being blunt, tactless, and interpersonally insensitive), domineering (i.e., the proneness to be forceful and controlling in relationships), and suspiciousness (i.e., questioning the honesty, fidelity, and motives of others).

Disinhibited externalizing has also been linked to multiple maladaptive traits reflecting disorganization, poor impulse control, and a lack of concern regarding the consequences of one's behavior[1,77,78,80-83]. The specific traits that have been most strongly and consistently associated with the disinhibited externalizing spectrum include impulsivity (i.e., acting spontaneously on the spur of the moment without concern for consequences), irresponsibility (i.e., being undependable and failing to fulfill obligations), distractibility (i.e., problems in attention and difficulties in focusing on tasks), risk taking (i.e., being reckless and engaging in potentially dangerous activities), (low) perfectionism (i.e., having low standards for the completion of work), and (low) workaholism (i.e., being more interested in having fun than in work-related activities).

These trait correlates, in turn, help to explain the specific types of personality-related pathology that are subsumed within each spectrum, including both adult[1,67,84,85] and youth disorders[86-93]. As can be seen in Figure 1, the antagonistic externalizing spectrum subsumes narcissistic, paranoid and histrionic PDs. Disinhibited externalizing includes ADHD, alcohol use disorder, and SUDs. Disorders such as conduct disorder, antisocial PD, intermittent explosive disorder, ODD, and borderline PD contain trait characteristics related to both spectra (e.g., impulsivity and anger/aggression).

VALIDITY EVIDENCE

Behavior genetic evidence

Evidence for a genetically coherent externalizing superspectrum has emerged most strongly from twin studies of constituent disorders and related personality traits in both youth and late adolescent/adult samples.

Specifically, in youth samples, twin studies have shown high heritabilities (h^2) and moderate levels of non-shared environmental influences, but non-significant shared environmental influences, for ADHD (h^2=~60-80%)[94] and ODD (h^2=~30-70%)[93], as well as for psychopathic traits (such as callous-unemotionality and narcissism)[95]. These studies have also found moderate heritability (h^2=~50%), shared environmental influences, and non-shared environmental influences for CD[96], and moderate heritability for various forms of youth antisocial behavior, including rule breaking and aggression[97], with its various forms such as reactive, proactive and relational aggression[96,98,99].

Most importantly, behavior genetic studies have provided evidence for the coherence of the externalizing superspectrum by showing high levels of genetic overlap across ADHD, ODD and CD[41,100,101], such that the largest contributor to the overlap among these disorders or the covariation among their symptom dimensions is represented by common genetic influences. This is also borne out by studies that have directly estimated the magnitude of genetic influences on an externalizing factor, and have found it to be highly heritable[41,102].

Evidence for the genetic basis of the externalizing superspectrum in youth also includes studies that have demonstrated common genetic influences between these disorders and personality traits such as behavioral disinhibition, neuroticism, and low prosociality[54,103,104].

Twin studies in late adolescent/adult samples provide considerable evidence for the validity of the externalizing superspectrum[54,103,104]. This evidence comes from studies of PDs, SUDs, and their symptom dimensions and related traits (e.g., antisocial behavior).

The "Cluster B" PDs, when examined individually, exhibit moderate to large heritability estimates[105]. The covariance among these disorders can be accounted for by a genetic common factor, with a second genetic factor accounting for variance in antisocial and borderline PDs[106]. Antisocial PD has also been included as an observed indicator in a highly heritable externalizing factor[50,103]. Relatedly, Kendler et al[66] reported evidence for a genetically coherent "Axis I" externalizing factor encompassing antisocial PD as well as CD, alcohol abuse/dependence, and drug abuse/dependence. These authors also found a genetically coherent "Axis II" externalizing factor encompassing dependent, histrionic, narcissistic, obsessive-compulsive, paranoid and borderline PDs, along with eating disorders.

The DSM-5 includes an alternative dimensional model of PDs as opposed to the criteria of the categorical diagnostic model. Most relevant to externalizing are the higher-order domains of antagonism and disinhibition, which are moderately heritable[107,108]. In a joint exploratory factor analysis including the DSM-5 alternative trait model domains, PD symptoms and normal personality domains, three genetic factors emerged: a PD/neuroticism factor, an antagonism/antisocial factor, and a factor reflecting schizoid/detachment[109].

Twin/family studies compellingly demonstrate that SUDs are genetically influenced, with ~50% of the variance in alcohol use disorders[110], 50-60% in problematic cannabis use[111], ~40-80% in cocaine use disorders[105,112,113], 20-50% in opioid dependence[105,112], and ~60% in nicotine dependence[114] being due to genetic influences. Critically, twin studies indicate that genetic influences are largely shared across SUDs[115]. Further, related psychiatric and behavioral manifestations, such as childhood conduct problems, adult antisocial behavior, behavioral undercontrol and impulsivity[116], also load on this shared genetic factor, which is highly heritable (~80%)[50,54]. A general liability towards externalizing explains the majority of genetic influences for alcohol and other SUDs, including 74-80% of the genetic influences on alcohol use disorders and 62-74% of those on other SUDs; it also accounts for 33-37% of the genetic influences on nicotine dependence.

Molecular genetic evidence

Molecular genetic research also supports an appreciable contribution of genes to individual disorders and traits within the externalizing superspectrum.

Candidate gene studies of ADHD have provided some suggestive evidence of association for genes within the dopamine and serotonin neurotransmitter systems, including the dopamine transporter and D4 and D5 receptor genes (*DAT1*, *DRD4* and *DRD5*), the serotonin transporter and receptor 1 genes (*5HTT* and *HTR1B*), and the synaptosomal-associated protein 25 gene (*SNAP-25*)[117].

Genome-wide association studies (GWAS) of various childhood disorders, such as ADHD[118], CD[119], and ODD or CD within the context of ADHD[120], have found evidence for several genome-wide significant associations and polygenic influences, each with a small effect size, that contribute to the risk for these disorders. In addition, moderate genetic correlations have been found between ADHD and other disorders, such as depression and anorexia nervosa; related traits, such as neuroticism and subjective well-being (negative); and important life outcomes, including ever having smoked, the number of cigarettes smoked per day, and intelligence and educational attainment (both negative)[120].

Interestingly, ADHD was not genetically correlated with antisocial behavior in another study, likely due to the relatively small sample size and the heterogeneity of measures of antisocial behavior[121]. In contrast, ODD or CD in the context of ADHD was highly genetically correlated with aggression and antisocial behavior, and its polygenic risk score was more predictive of cognitive functioning, educational outcomes, and having children at a younger age than that for ADHD without ODD or CD[120]. Nonetheless, the maximum variance explained by the polygenic risk score in these outcomes was quite low (0.36%).

In adolescent and adult samples, GWAS of externalizing PDs are still in their infancy, with only borderline and antisocial PDs being investigated to date, using relatively small samples. One molecular genetic study indicated that borderline PD is heritable[122], but did not test for its genetic association with any other form of externalizing psychopathology. Current GWAS evidence indicates that antisocial behavior is heritable and significantly genetically correlated with CD and neuroticism, but not with schizophrenia, bipolar disorder or ADHD[121]. Furthermore, a study found high genetic correlations of antisocial behavior with lifetime cannabis use and cigarette smoking, but not with alcohol consumption[123], while another study did not find an association between polygenic risk scores for antisocial PD and either tobacco or alcohol use[124]. A GWAS of antisocial PD[125] reported the most associated gene (*ABCB1*) to be one involved in immune function and associated with various forms of substance abuse. These studies have also found that many common genetic variants, each with a small effect size, contribute to risk for antisocial behavior. Finally, a large GWAS of normal personality traits did not find that agreeableness has genetic correlations with any externalizing disorders or other forms of psychopathology[126].

The majority of GWAS on substance use have focused on alcohol-related phenotypes, including alcohol dependence[127], alcohol use disorder[128], number of alcoholic drinks per week[129], and maximum alcohol intake. Studies of these phenotypes have employed moderately to extremely large sample sizes, thus being well-powered. One finding which robustly emerged from these GWAS is that genetic influences on alcohol consumption are only moderately correlated with those on alcohol use disorders[130]. Cannabis related GWAS are beginning to reach adequate power[131-133], but still require even larger samples. GWAS on cocaine dependence[134,135] and opioid dependence[136-138] are currently underpowered. It is important to note that, even in large cohorts, polygenic risk scores continue to predict only small proportions of the variance in independent samples (e.g., the polygenic risk score from a GWAS involving ~1 million participants explained only about 2.5% of the variance in alcohol consumption).

Newer multivariate methods such as genomic structural equation modeling (genomic SEM)[139,140] can be used to model the underlying factor structure of genetic correlations from a set of phenotypes of interest using GWAS summary statistics. These methods enable researchers to move beyond a single disorder or behavior in gene identification efforts, and instead focus on identifying genes contributing to the underlying latent factor(s). Genomic SEM is currently being applied in the international Externalizing Consortium, which analyzed genome-wide data on seven phenotypes related to the externalizing superspectrum from ~1.5 million people and identified nearly 600 significant genetic loci associated with a general liability to externalizing[141]. A polygenic risk score derived from this dataset predicted up to 10% of the variance in general externalizing scores in independent samples, and emerged as significant in both within-sibling and between-sibling comparisons. These analyses suggest that focusing gene identification efforts on general externalizing liability, rather than on individual externalizing disorders/behaviors, is a fruitful approach to advancing knowledge of genes contributing to this psychopathological domain.

Environmental risk factors

Decades of observational research have identified a wide range of environmental risk factors for externalizing problems, spanning a variety of social domains. Meta-analyses document that abuse, neglect, hostile parenting, neighborhood violence, and affiliation with deviant peers all exhibit significant associations with diverse externalizing phenomena[142-144]. Longitudinal research in the community confirms that these effects can endure through adolescence and beyond[145].

Effects of toxic environments are not only robust, but also diffuse. That is, prominent etiological events appear to engender risk for a variety of externalizing mental health conditions and maladaptive personality traits[146]. Indeed, there are essentially no known unique environmental risk factors for any substance use or behavioral disorder.

This observation prompted research on how environmental pathogens relate to composites of externalizing phenotypes. In an epidemiologic sample, various forms of childhood maltreatment predicted individual differences on a latent externalizing dimension constructed from substance use and antisocial behavior disorders[147]. This effect was replicated in a number of cohort studies[148,149]. Across studies, the severity of social stress predicted variation in the broad externalizing factor, but not unique components of the specific observed externalizing conditions. This pattern is evident in research on other risk factors that focus on externalizing outcomes which transcend traditional disorder boundaries. Peer victimization, discrimination experiences, and other chronically stressful conditions such as romantic conflict and unemployment, all predicted standing on a latent externalizing spectrum in separate community samples[150,151].

The connection between externalizing problems and environmental stressors over time is almost certainly bidirectional. Research in community samples shows that variation in a latent externalizing factor predicts future rates of both acute life events (e.g., arrest) and ongoing strains (e.g., marital discord)[152,153]. These stressful conditions, in turn, presumably set the stage for continued externalizing behavior. This type of person-environment fit implies a vicious cycle of stress exposure and worsening externalizing problems, akin to the transactional peer selection and socialization effects on externalizing risk in adolescence[145].

As a whole, longitudinal research has revealed strong connections between a wide range of environmental exposures and the externalizing superspectrum. Much less is known about whether certain environments predispose selectively to disinhibited vs. antagonistic spectra (or any other more homogeneous components) within the superspectrum. The available data at this time suggest that environmental risk is largely non-specific. More research using genetically informative designs is needed to verify the etiologic roles of putative environmental risk factors by controlling for passive gene-environment correlation (e.g., parents creating a home environment that is influenced by their heritable characteristics)[154].

Cognitive and emotional processing abnormalities

Generally speaking, the externalizing superspectrum model helps to organize the literature on cognitive deficits, as reflected in Figure 1. In particular, there is overwhelming evidence that cognitive impairment is prominent in disinhibited forms of externalizing.

Evidence of impaired executive functioning is most substantial for antisocial PD[155-159] and CD[160,161], followed by disinhibitory traits[162-166]. Additionally, deficits in sustained attention, inhibitory control, and sluggish cognitive tempo are associated with ADHD[162,167-172]. There is evidence of cognitive deficits in children with ODD, albeit less abundant[173,174], which might be partly explained by high comorbidity with both ADHD and CD[175,176]. There is even less evidence of cognitive deficits related to intermittent explosive disorder, which is mostly characterized by impairments in social cognition and emotion regulation[177-180]. Impairments in executive functions are extensively reported in individuals with drug and alcohol dependence[181-188].

Under the antagonistic externalizing spectrum, the evidence of cognitive deficits is strong for borderline PD[189,190], whereas findings concerning narcissistic, histrionic and paranoid PDs are mostly derived from symptom, descriptive and trait checklists[191,192].

Antisocial traits are linked with deficits in the ability to regulate emotions and diminished responsiveness to distress in others[193-196]. ODD is associated with deficits in empathy, and impaired emotion regulation has been reported in both ODD and intermittent explosive disorder[180,197,198]. There is evidence for emotion dysregulation impairments also in substance dependent individuals[199-201].

Impaired facial affect recognition and emotional regulation

deficits are observed in individuals with borderline PD[202,203]. The evidence concerning narcissistic and paranoid PDs (respectively, difficulties in emotional empathy and regulation[204], and hypervigilance and stress reactivity[205]) has come from symptom, descriptive and trait checklists, rather than behavioral task performance.

Neurophysiological indicators

The best-established neurophysiological indicator of broad externalizing is reduced amplitude of the visual P300 (P3) event-related potential (ERP)[206], a positive-going ERP that occurs in relation to rare or otherwise salient visual events within an ongoing stimulus series.

Originally thought to be indicative of proneness to alcohol problems[207], subsequent research showed reduced P3 to be related to various other externalizing conditions as well[208]. Ultimately, it became clear that P3 operates as an indicator of the highly heritable liability for externalizing problems in general[209,210]. Like broad externalizing, P3 amplitude is appreciably heritable, and its association with this superspectrum factor reflects additive genetic influences in common between the two[211,212].

Other evidence points to a genetically-based association between broad externalizing and performance on executive control tasks[213], and overlap is evident in the relations of P3 amplitude and executive task performance with broad externalizing[166,214]. The implication is that reduced P3 reflects a weakness in cognitive control capacity that is associated with heritable risk for externalizing problems in general[215,216], highlighting P3 as a marker of the broad externalizing factor at the superspectrum level of HiTOP.

Another less well-established candidate indicator of broad externalizing is reduced amplitude of the error-related negativity (ERN), a negative-going ERP that is evident following errors in a speeded reaction time task, and is theorized to reflect performance monitoring and error detection processes. Reduced ERN was initially reported for individuals high in impulsive traits[217,218], and later for individuals high in broad externalizing[219]. Further research is needed, though, to evaluate the specificity of the relationship of reduced ERN to broad externalizing, and the neural systems basis of this relationship. In addition, research is needed on the etiologic basis of the association between ERN and externalizing problems, given the limited work of this kind to date[220].

Studies that have specifically assessed antagonistic externalizing tendencies along with broad externalizing have shown reduced P3 and ERN in relation to the latter, but not to antagonism-specific variance[221,222]. By contrast, high antagonistic externalizing is reliably associated with reduced brain reactivity to fearful face stimuli. Multiple studies have reported reduced amygdala activation to fearful faces in children/adolescents exhibiting antagonistic externalizing tendencies (sometimes termed "callous unemotionality")[223,224] along with conduct problems, compared to children lacking in antagonistic externalizing. Importantly, this effect has been found to be specific to antagonistic externalizing (callous-unemotionality) by contrasting groups of children matched for externalizing problems but differing in levels of callous-unemotionality[210]. Consistent with this, two studies[225,226] reported reduced early-ERP responses to fearful faces in adults scoring high on a measure of antagonistic externalizing (termed "callousness"); broad externalizing was also assessed in these studies, and effects were shown to be attributable to callousness-specific variance. This impaired responsiveness to fearful faces may reflect general emotional insensitivity among those high on antagonistic externalizing, or perhaps a more specific deficit in the capacity for empathy or affiliative capacity among these individuals[227].

Interpretation of the research literature on neurophysiological indicators of problems situated specifically within the disinhibited externalizing spectrum of HiTOP – in particular, substance use problems – is hampered by a failure to differentiate between specific factors versus broad externalizing liability[228], neglect of the distinction between liability indicators and symptom or "scar" indicators[229], and the substance-specific nature of particular indicators[230]. For example, while there is considerable evidence for a distinct role of reward system dysfunction in substance addictions, it remains unclear at this time whether addiction proneness entails heightened or diminished sensitivity to naturally occurring rewards[231-233], due to limitations of existing research. To address these limitations, longitudinal studies are needed that differentiate between neural measures of premorbid liability to externalizing problems in general, as opposed to measures indicative of addiction liability more specifically, or active symptoms or persisting consequences of substance addiction[229].

Neuroimaging

As with other psychiatric domains, the neuroimaging literature on externalizing has been dominated by case-control studies of individual disorders, but these are now complemented by growing research taking the transdiagnostic dimensional approach. This work is identifying alterations in a number of circuits involved in social-emotional processing, aversive learning, emotional regulation, and cognitive control, with varying levels of specificity between antagonism and disinhibition domains, as well as narrower lower-order constructs that contribute to these domains. We highlight some of the key circuits as a demonstration of the compatibility of neuroimaging data with the HiTOP model of externalizing.

Among the most frequent findings is the observation of reduced amygdala volume, which has been seen in case-control studies or disorder-specific symptom measures of psychopathy and antisocial personality[234], conduct and oppositional problems[174], borderline personality[235,236], aggression and violence[237], risk for substance use problems[238,239], and ADHD[240]. While amygdala volume reductions correlate with broad measures of externalizing traits[241,242], they appear most pronounced for callous-unemotional and antagonistic traits[174,243] as opposed to disinhibition features.

Given the importance of the amygdala in social-emotional

processing, emotional responses to aversive stimuli, and aversive learning[244], such findings fit with psychological and psychophysiological models emphasizing social-emotional and fear learning deficits as core features in the etiology of antagonistic spectrum problems[245-247]. This having been said, reduced amygdala volume has also been reported for other diagnostic constructs (e.g., post-traumatic stress disorder)[245-247], such that the specificity of this association would benefit from further study.

Reductions in amygdala volume are paralleled by changes in task-related activity in disorders with high antagonism characteristics, as repeatedly demonstrated in functional magnetic resonance studies of individuals with antisocial-psychopathic and borderline personality traits[248,249]. Again, the associations appear to most robustly reflect antagonism/callous-unemotionality rather than disinhibition. For instance, lower task-relevant activations are seen in the bilateral amygdala among individuals with ODD/CD as compared to ADHD in a number of tasks[174], and studies using dimensional measures of symptom severity have repeatedly observed reductions in the amygdala response to social-emotional stimuli in relation to callous-unemotional traits[210,223,250].

The amygdala is just one part of a limbic/paralimbic network that has been implicated in different aspects of externalizing[251,252]. Neuroimaging studies of psychopathy especially emphasize the orbitofrontal/ventromedial prefrontal cortical (OFC/VMPFC) region[253], that shares strong structural and functional connectivity with the amygdala[254]. Such an involvement is consistent with the key role of this region in social cognition, including empathy and moral reasoning[255,256], and has helped form the basis of one of the most prominent neural models of psychopathy[253]. Critically, portions of this region have long been associated with the ability to inhibit behavior, with lesions often causing both antisocial behavior and problems with impulsivity and disinhibition[257]. It is thus notable that phenotypic associations with structural and functional features in these circuits extend beyond antagonism or callous-unemotional traits. Both human and animal models demonstrate the importance of the OFC/VMPFC region to both substance use history and the risk for developing substance use[258-260] as well as behavioral addictions[261].

Despite indications of overlap that point to involvement beyond antagonism or disinhibition domains, important differences emerge between ventromedial and ventrolateral prefrontal regions, which appear generally consistent with the core cognitive and emotional functions of these regions. Problems with social antagonistic factors are more prominently reflected in ventromedial regions, while alterations in ventrolateral regions (lateral orbital/inferior frontal) are more related to cognitive control (including response inhibition) and executive functions[262]. For instance, deficits in cognitive control show significant associations with task-related inferior frontal gyrus engagement in both substance dependence and ADHD[263,264].

The dorsal anterior cingulate cortex has been of particular interest in relation to externalizing due to its role in both attention and error monitoring. Alterations in both structure and function have been reported for this area in relation to various externalizing conditions, including ADHD[263,265,266], psychopathy and violent behavior[252], disruptive behavior[267], substance dependence[260,268,269], and behavioral addictions[261]. These findings are of particular interest given the importance of this region in the generation of the ERN[219], providing convergent evidence for a core role of this area in cognitive control deficits in externalizing problems as a whole (i.e., at the superspectrum level of the HiTOP system).

In considering the involvement of cortical areas in externalizing psychopathology, it should be noted that some neural correlates may extend quite broadly, even if particular areas play more focal roles in the expression of specific forms of externalizing. For example, the largest meta-analysis to date of findings for ADHD[265] reported evidence not only for lower surface area in frontal, cingulate and temporal cortical regions, but also lower average effect across the whole cortical area, with the severity of this overall deficit declining from childhood to adolescence and eliminated by adulthood. It will be increasingly important to consider how phenotypic expressions of externalizing are related to, and change with, processes of brain maturation[270,271].

The basal ganglia have been a further focus of interest in the externalizing literature. In particular, dysfunction in mesolimbic and nigrostriatal systems has been repeatedly implicated in reward-motivational processes relevant to risk for and development of addictions[272-274], and also ADHD[265,275,276]. Differences in the functioning of these systems have been linked to altered processes of reward valuation, discounting behavior and impulsivity that characterize externalizing problems[264,277-279]. Even with respect to antagonistic behaviors, individual differences in the functioning of mesolimbic circuits may dramatically affect the manner in which antagonistic actions are expressed – for example, in the sort of impulsive-antisocial actions that emerge in these conditions.

In one of the few studies to examine neuroimaging activation in relation to an externalizing factor, while controlling for scores on a general psychopathology factor, fronto-parietal network hypoactivation during a working memory task was related to increased scores on a "behavioral disturbance" factor, primarily comprising ADHD and CD symptoms[280]. These findings are complemented by recent work reporting relations for the same behavioral disturbance factor with enhanced connectivity within the fronto-parietal control network, but decreased connectivity within the default mode network[281]. Other dimensional measures of externalizing have similarly been associated with network dysfunction in many of the same regions identified in the foregoing summary of findings[282,283]. Consideration of neural networks and their features, as opposed to individual brain regions, almost certainly will prove essential to characterizing the role of neural systems and processes in externalizing problems.

Other biomarkers

Aberrant patterns of DNA methylation have been linked to externalizing psychopathology, including addiction[284,285] and antisocial behaviors[286-288]. Epigenetic findings also indicate common

downstream biological processes in ODD and ADHD, including dysregulation of long-term neuronal synaptic plasticity[289]. DNA methylation is thought to represent a molecular pathway through which environmental exposures become translated into phenotypic variation, conferring increased susceptibility to externalizing disorders[290,291]. Accordingly, one study identified an epigenetic risk score to broad (tobacco, cannabis and alcohol) substance abuse liability, which mediated the prospective association between prenatal maternal tobacco smoking and adolescent substance use[292].

An inflammation-related epigenetic risk score at birth was associated with higher externalizing problems across childhood and adolescence[293]. Elevated levels of pro-inflammatory markers (e.g., cytokines, C-reactive protein) in peripheral tissues such as blood have also been reported in externalizing psychopathology[294-296], including ADHD[297-299], antisocial PD[300], and substance abuse[300-303], although the overall evidence in this respect is mixed.

Meta-analytic evidence supports lower cortisol levels in patients with ADHD[304]. In general, reduced cortisol is also associated with persistent aggression and other antisocial and disinhibited behaviors in children and adults[305-307]. Moreover, blunted cortisol response to stress has been associated with relapse in patients with addiction[308]. Thus, lower cortisol may reflect an impairment in the ability to regulate stress responses that underpins chronic externalizing psychopathology, as well as other forms of psychopathology more broadly[309].

Low platelet monoamine oxidase B (MAO-B) enzyme activity, which is a proxy of low central serotonergic functions, has been consistently shown to correlate with impulsive, aggressive and antisocial personality traits and behaviors, including ADHD[304], alcohol-related problems, and smoking[310]. The role of MAO-B in externalizing disorders is thought to be independent of the effects of tobacco smoking on the enzyme[311]. Moreover, there is evidence for low cerebrospinal fluid serotonin metabolite 5-hydroxyindoleacetic acid (5-HIAA) levels characterizing alcohol abuse and antisocial behavior, including disinhibited forms of aggression[312,313], although this effect remains debated[314]. Thus, serotonin hypofunction may be a shared biological mechanism underlying disinhibited and antagonistic psychopathology.

Overall, research evidence suggests that conditions within the disinhibited and antagonistic externalizing spectra share common biological signatures. However, conclusions have been constrained by methodological limitations of the existing studies, including small sample sizes, focus on a single disorder, and paucity of longitudinal designs, which are particularly relevant for disentangling biological markers of risk vs. consequences of substance use and/or medication.

Temperamental antecedents

Continuity in the traits that underlie the externalizing superspectrum, beginning in early childhood through adolescence and adulthood, has been documented by research[74,315-319].

For example, disinhibition is captured by low effortful control in early childhood[315,316], which has been shown to be a robust predictor of subsequent externalizing behaviors[320,321]. This literature is paralleled by evidence that low agreeableness and conscientiousness (captured together in low effortful control) together predict externalizing behaviors later in childhood and adolescence[315,320]. Negative affectivity has also been found to consistently predict externalizing[320,321], but with low specificity, as it tends to act as a broadband risk for subsequent psychopathology[31,322].

A similar pattern of low effortful control and high negative affectivity has been found to prospectively predict antisocial behavior indicators, including CD, ADHD, ODD and antisocial PD[13,321,323-327]. By contrast, limited evidence exists for intermittent explosive disorder[328]. A large prospective study (N=4,983) in Australia found that high negative affectivity, low effortful control, and high surgency (extraversion) at age 4-5 each uniquely predicted the development of ADHD and CD symptoms to age 12-13[329]. Similarly, a study of two birth cohorts from Norway (N=797) found that high negative affectivity and high surgency predicted increases in ODD symptoms from age 4 to 6[326]. Although not included in traditional models of temperament, callous-unemotional traits in childhood and adolescence (i.e., low empathy, lack of remorse, and insensitivity to distress of others) also robustly and prospectively predict risk for severe antisocial and related behaviors[224,330].

There is little evidence regarding the childhood antecedents of adult PDs included in the antagonistic spectrum of the HiTOP model (e.g., histrionic, narcissistic and paranoid PD)[317], while some research has found that negative affectivity[316,331] and low effortful control[331,332] predict borderline PD, mirroring the findings for other externalizing disorders. Finally, SUDs, reflecting disinhibited externalizing in the HiTOP model, are consistently related to low effortful control[333,334], as well as high negative affectivity[334], with some evidence also pointing to an association with surgency/extraversion (e.g., for cannabis use)[335].

Overall, the combination of negative affectivity with low effortful control represents a consistent constellation of temperamental traits that acts as an antecedent to the externalizing superspectrum. Disinhibited and antagonistic spectra do not tend to show differential associations with childhood temperament, although there is some evidence that callous unemotionality represents an additional risk factor for severe antisocial behavior.

Illness course

Several authors have described a trajectory of externalizing behaviors that begins with hyperactivity and impulsivity in preschool-age children, followed by delinquency in middle school, and SUDs and antisocial personality in late adolescence and emerging adulthood[336-338]. This pattern of progression of externalizing behaviors suggests a shared etiology, and has led to the suggestion that the so-called "co-occurrence" among individual DSM externalizing disorders is largely artifactual, stemming from the split of a unitary construct into multiple diagnoses.

The validity of the externalizing superspectrum is also supported by the high stability over time of externalizing behaviors[339,340], from middle childhood through late adolescence[340]. Olson et al[341] measured externalizing outcomes throughout the school-age period and at age 17 using a multi-informant approach. They found that children at risk for externalizing problems later in childhood and at age 17 were perceived as "difficult" and resistant to control as toddlers. Parental perceptions about child behaviors predicted externalizing behavior as early as at 13 months and remained persistent predictors throughout late adolescence.

Antagonistic and disinhibited spectra have not shown substantial evidence of differential patterns of developmental trajectories.

Treatment response

Numerous treatments have proven effective for a broad array of externalizing disorders in children and adolescents, including behavioral/psychosocial[342-344], systems- or school-based[345-347], and psychopharmacological interventions[348-351], while only few treatments have been successfully used for externalizing in adults (for instance, motivational interviewing has long been used to treat SUDs, and treatment effects have been found to last up to two years, with 75% of participants gaining some type of improvement[352]).

A meta-analysis of 36 randomized, between-subjects comparison studies of psychosocial treatment efficacy for externalizing problems in children less than 8 years of age[353] found that general externalizing symptoms showed the largest treatment response, followed by opposition/non-compliance. Impulsivity/hyperactivity showed the weakest response (although the effect size was still within the "medium" range). These findings suggest that a dimensional approach designed to treat specific components of externalizing may have greater clinical utility than applying individual treatments to individual disorders.

In support of a dimensional approach, Epstein et al[354] carried out a meta-analysis of 28 studies of psychosocial interventions for childhood externalizing problems. Using random effect variances, they found that dimensional externalizing severity scores accounted for significant additional variance in predicting treatment outcomes.

Furthermore, there appears to be utility for assessing the full range of the externalizing superspectrum in randomized clinical trials designed to treat externalizing psychopathology. For example, in the meta-analysis by Battagliese et al[355], the authors stated that they could not examine effects of cognitive-behavior therapy on certain diagnostic subgroups because no studies measured ADHD symptoms in children with a diagnosis of ODD and only two studies included children affected by CD. Given the high rates of diagnostic co-occurrence within the externalizing spectrum, assessing and treating the full range of externalizing problems for an individual client may be a parsimonious and effective approach to designing future interventions.

Summary of validity evidence

The validity evidence reviewed herein is summarized in Table 2. This table shows a substantial coherence within the disinhibited and antagonistic spectra, as well as an overlap between them. This supports the validity of a hierarchical conceptualization, involving an overarching externalizing superspectrum with two distinguishable spectra. As shown in the column "Summary of specificity", most validators (sixteen) are evident for the broad externalizing superspectrum, with some (eight) evident for disinhibition and one for antagonism.

Notably, cells that are blank in the table indicate a lack of evidence, not the absence of an effect. These may therefore be fruitful areas for future inquiry. Generally speaking, large sample designs where all elements of the externalizing superspectrum are well characterized, along with multiple validators, can improve inferences by helping to address questions of generality and specificity.

Many validators considered here may not be specific to externalizing. For example, pro-inflammatory biomarkers were characterized as also related to the psychosis superspectrum of HiTOP in our previous paper in this series[19]. These and other factors (e.g., childhood adversity) are likely broadly relevant to psychopathology risk, and not specific to externalizing risk.

Generally speaking, these validity findings dovetail well with the structural perspective on psychopathology taken in the HiTOP consortium. In contemplating the validity of psychopathological concepts, it is no longer sufficient to focus on putative diagnostic categories in isolation. Rather, broad characterization of psychopathological phenomena, along with assessment of specific validators in large samples, can deepen our understanding by revealing the interplay between the structural organization of psychopathology and multiple putative causes and correlates.

UTILITY EVIDENCE

Reliability

Some of the largest studies on the reliability of the diagnosis of mental disorders have come from field trials of the official classification systems, the DSM and ICD. Results of the DSM-5 field trials documented moderate/good reliability for alcohol use disorder (test-retest kappa coefficient of 0.40) and questionable reliability for antisocial PD (kappa=0.21)[356]. These estimates are lower than those observed in field trials of DSM-IV, largely due to the fact that "usual clinical interview approaches"[356] were utilized in the DSM-5 field trials instead of highly structured diagnostic interviews as in the DSM-IV field trials[357]. Nevertheless, complementary analysis of DSM-5 cross-cutting dimensional measures of externalizing-related constructs (confined to alcohol, tobacco and illicit drug use) demonstrated higher reliability compared to their categorical counterparts[358].

Direct comparisons of continuous and categorical measures of psychopathology are rare. In a comprehensive review, Markon

Table 2 Validators of the disinhibited and antagonistic spectra

	Both spectra						Disinhibited spectrum				Antagonistic spectrum				Summary of specificity
	ASPD	CD	ODD	IED	BPD	Traits	AUD	SUD	ADHD	Traits	NPD	HPD	PPD	Traits	
Genetics															
Family/twin	++	+	+		++	+	+	+++	+	++	++	++		+	B
Polygenic risk		+	+						+						B
GWAS	+	+	+		+		+++	+	+	++					D
Environment															
Neighborhood risk factors						++		++							B
Peer interactions						++		++							B
Childhood maltreatment	+++	+++	+++	+++	+++	+++		+++							B
Cognition/Neurobiology															
Cognitive deficits	+++	+++	+	–	++	+++	+++	+++	+++	+++	–	–	–		B
Emotional processing abnormalities	+++	+++	++	+	+++	+++		++	+++	+	+	+	+		B
Reduced amygdala volume	++	+	+		++	+		++	+					+++	B
Involvement of OFC/VMPFC						++		+++							B
Task-related inferior frontal gyrus engagement								++	++						D
Aberrations in dorsal anterior cingulate				+				++	+++					+	B
Reward system dysfunction								+++							D
Dysfunction in mesolimbic and nigrostriatal systems								+++	+++	++					D
Reduced amygdala activation to fearful faces		++												+++	A
Blunted P300							++	++							B
Blunted error-related negativity										++					D
Biomarkers															
DNA methylation						+				+					B
Elevated pro-inflammatory markers	+							+	+						D
Low cortisol						+		+	+						D
Low MAO-B						+	+		+	+					D
Antecedents/Course															
Low effortful control	++	++	+		++	+++		+++	++	+++				++	B
High negative affectivity	+	+	++		++	+++		+	++						B
Extraversion/surgency		+	+					+	+						B
Treatment															
Response to psychosocial interventions						+				+				+	B

+: some evidence for effect, ++: some replications, +++: repeatedly replicated finding, –: some evidence for reverse effect, – –: some replications for reverse effect, – – –: repeatedly replicated reverse effect, A – linked to antagonism, D – linked to disinhibition, B – linked to both, ASPD – antisocial personality disorder, CD – conduct disorder, ODD – oppositional defiant disorder, IED – intermittent explosive disorder, BPD – borderline personality disorder, AUD – alcohol use disorder, SUD – substance use disorder, ADHD – attention-deficit/hyperactivity disorder, NPD – narcissistic personality disorder, HPD – histrionic personality disorder, PPD – paranoid personality disorder, GWAS – genome-wide association studies, OFC/VMPFC – orbitofrontal/ventromedial prefrontal cortex, MAO-B – monoamine oxidase B.

et al[4] found that continuous measures of psychopathology were generally more reliable than discrete measures across all psychopathology domains, and that the overall meta-analytic reliability estimate for the externalizing domain was 0.77.

A growing body of research has examined the reliability of PDs and personality dimensions that fall within the externalizing spectrum. Using the Personality Inventory for DSM-5 (PID-5)[359], a questionnaire specifically developed to operationalize the DSM-5 dimensional trait model for PDs, a high internal reliability of the disinhibition (McDonald's omega = 0.80) and antagonism (omega = 0.83) domains was documented[360].

In a study of the stability of PID-5 domains over a one-year period, the externalizing domains of the PID-5 were relatively stable across a one-year period in individuals diagnosed with PDs[361]. In a study examining both personality traits and PDs, high levels of stability over a two-week period (referred to as dependability by the authors) were reported in PID-5 domains of antagonism (0.86) and disinhibition (0.86)[362]. In addition, the authors provided evidence of clear advantages of dimensional over categorical ratings for PDs traditionally linked to the externalizing domain (e.g., antisocial PD).

Explanatory and prognostic power

Using data from two waves of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) study, a large general population longitudinal investigation, Kim and Eaton[61] demonstrated that an externalizing dimension at Wave 1 predicted Wave 2 mental disorder diagnoses more strongly than individual diagnoses.

Externalizing dimensions have also outperformed diagnoses when explaining variance in suicidality, psychotic-like experiences and internalizing-type disorders[363]. Furthermore, the externalizing dimension has been shown to mediate the relations of constructs such as childhood maltreatment with diagnosed externalizing-type mental disorders (e.g., SUDs)[147]. Similar general vs. disorder-specific findings are evident when examining constructs such as perceived racial discrimination[151], stress responsivity[59], and transmission of externalizing disorders from parents to offspring.

Collectively, this research points to the superiority of the HiTOP conceptualization of externalizing psychopathology in predicting a wide range of disorder validators.

Clinical utility

The utility of integrating the HiTOP model into clinical practice has been recently addressed[364]. Conway et al[44] demonstrated that the HiTOP structure generalizes well to patterns of comorbidity among diagnoses assigned by health practitioners in everyday practice. They further demonstrated that categorical diagnoses did not offer additional incremental validity when predicting suicidality and self-injury, over and above the identified HiTOP dimensions.

Research on the clinical utility of dimensional versus categorical conceptualizations of externalizing largely comes from the PD field, and draws heavily from studies that examine practitioner ratings of utility. Using case vignettes as well as data obtained from actual patients, these studies evaluate the clinical utility of dimensional and categorical frameworks across various dimensions of utility (e.g., ease of use, utility in communicating with other health professionals, usefulness in formulating a therapeutic intervention, and usefulness in treatment planning). Recently, Bornstein and Natoli[365] summarized this literature and found that dimensional models of PD are rated more positively than categorical models with respect to most areas of clinical utility.

MEASUREMENT

The Externalizing Spectrum Inventory (ESI) is one of the most well-validated instruments to measure individual facets and global levels of the externalizing superspectrum. The ESI was developed using a bottom-up process to target 23 unidimensional facets of externalizing and capture the hierarchical structure of broad externalizing (or disinhibition) along with specific factors associated with callousness/aggression and substance abuse[15].

Independent validation studies have demonstrated that the broad factors of the ESI possess concurrent validity against the Multidimensional Personality Questionnaire (MPQ)[366], measures of integrity, and a range of DSM-IV symptoms of externalizing disorders, personality traits, psychopathy, and symptoms of substance dependence[219,367,368].

Recent efforts have focused on improving the clinical utility of the ESI via the development of data-driven brief forms and adaptive scales. Patrick et al[16] constructed brief forms of the 23 facets with a total of 160 items (down from 415 items), ranging from 3 to 11 items per facet, which maintained high internal consistency, replicated the structure of the full ESI, and demonstrated similar validity in relation to the MPQ. Additional independent validation has confirmed the favorable psychometric properties of the brief form[369]. More recently, Sunderland et al[370] have demonstrated the feasibility of computerized adaptive versions of the ESI, producing similar scores as the full ESI with acceptable levels of reliability using very few items tailored to each respondent.

Omnibus clinical personality inventories are also available to assess the externalizing spectrum. Primary examples include the Minnesota Multiphasic Personality Inventory-2-Restructured form (MMPI-2-RF)[371] and the Personality Assessment Inventory (PAI)[372]. Specifically, the MMPI-2-RF captures behavioral/externalizing dysfunction at the higher order level, which comprises pervasive dysfunction with under-controlled or acting-out behaviors, as well as specific facet measurement (juvenile conduct problems, substance abuse, aggression, and anger proneness), all of which have been shown to directly map onto the same externalizing spectrum model as HiTOP and the ESI[373-376].

The MMPI-2-RF Personality Psychopathology Five (PSY-5-RF) scales also have trait-level measures of higher-order antagonism

(aggressiveness) and disinhibition (disconstraint). Furthermore, factor analytic research with the PAI scales has typically revealed three- or four-factor structures, with factors resembling disinhibition and antagonism usually emerging[373].

There are also several personality measures that map onto, and therefore operationalize, the externalizing superspectrum via dimensional personality traits, including the PID-5[359], the NEO Personality Inventory 3 (NEO-PI-3)[377], and the Comprehensive Assessment of Traits Relevant to Personality Disorder (CAT-PD)[378]. The PID-5 explicitly measures the antagonism and disinhibition trait domains that emerge from a broader externalizing superspectrum[379]. Similarly, a conjoint analysis of several dimensional personality trait inventories – the PID-5, CAT-PD and NEO-PI-3 – has provided evidence of a five-factor solution that bore strong resemblance to the HiTOP model and included factors for antagonism and disinhibition that converged onto a single externalizing dimension in hierarchical analysis[83].

In child and adolescent populations, a number of measures have been used extensively to assess externalizing and disinhibited behaviors, such as the Child Behavior Checklist (CBCL)[380], the Strengths and Difficulties Questionnaire (SDQ)[381], and the Diagnostic Interview Schedule for Children (DISC)[382], with factor analysis consistently identifying strong coherence between these measures and the broader HiTOP structure[383-385].

Finally, there are numerous scales designed to measure specific facets of externalizing and disinhibited behavior, such as substance use, impulsiveness, and aggression[386-388].

IMPLICATIONS

The HiTOP approach aims to advance our understanding of the natural organization of externalizing psychopathology in at least three major ways.

First, externalizing psychopathology encompasses two spectra, disinhibition and antagonism. These spectra show both similarities and differences, consistent with the fundamentally clarifying idea of disinhibitory and antagonistic aspects of a broader and more general externalizing superspectrum. Nevertheless, to characterize a patient fully, a profile across major psychopathology spectra needs to be considered, as detailed in previous HiTOP publications[1,19,364].

Second, the HiTOP approach underscores a growing consensus that clinically significant externalizing problems lie on a continuum with normative functioning and maladaptive traits. Developmentally earlier expressions of disinhibitory and antagonistic traits often precede the onset of serious sequalae (e.g., behaviors that are grounds for arrest). In this way, the HiTOP approach melds dimensional and developmental perspectives on psychopathology, as parts of a more integrated approach to understanding both development and broad population-level variation in socially consequential externalizing tendencies.

Third, the HiTOP approach addresses heterogeneity within externalizing problems by explicating specific trait and symptom dimensions that constitute broader spectra. Figure 1 provides an evidence-based guide to constituent narrow-band elements of externalizing. Nevertheless, continued research on the nature of specific sub-elements of externalizing psychopathology would be welcome, as the field pivots toward basing nosology on evidence, as opposed to diagnosis by tradition and putative authority[15,374].

FUTURE DIRECTIONS

The proposed HiTOP model of the externalizing superspectrum is based on extensive evidence. Nevertheless, intriguing possibilities exist to explore the discrete vs. continuous nature of psychopathology based on data. The HiTOP model is meant to include all empirical psychopathological entities, whether dimensional or categorical in nature. Only dimensions have been established empirically to date. Setting aside the complex political issues implied by this situation (e.g., the way authoritative nosologies tend to recognize committee-derived categories as opposed to empirically-derived dimensions), quantitative techniques can adjudicate between more continuous and more discrete accounts of the structure of psychopathology. Further research along these lines can help to continue to place psychiatric nosology on firmer empirical footing[3,19].

Systematic research can also provide a means for linking psychopathological variation with intervention implications in a principled manner. Rather than imposing arbitrary diagnostic thresholds, diagnostic algorithms can link clinical presentations with optimal intervention strategies. Practically speaking, clinical decisions rarely focus on "to treat or not to treat". Rather, an ordinal set of interventions varying in intensity is typically deployed in response to a corresponding level of clinical need. For example, externalizing problems frequently present as SUDs, because substance dependence creates an acute clinical need. Substance use intervention can range from medically responsible outpatient detoxification, to partial hospitalization, to inpatient services. This rough ordinal scale of intervention is typically tethered to clinical need (e.g., medical complications may require close observation to resolve, and a corresponding inpatient stay). Ultimately, these sorts of treatment options can be tethered to intensity of presentation in a principled way, relying on the types of evidence reviewed herein.

CONCLUSIONS

The HiTOP approach to clinical diagnosis provides an empirically based and hierarchical conceptualization of externalizing disorders that was derived from evidence. The validity evidence reviewed herein is extensive, and the utility of the approach was also reviewed and is readily apparent. Assessment instruments for externalizing constructs already exist, and the HiTOP approach can therefore be readily implemented.

Further research will be beneficial, but the HiTOP model is ready for use by scientists and clinicians interested in basing

their approaches on evidence as opposed to putative authority. The HiTOP approach addresses problems of heterogeneity, comorbidity and low reliability, thereby providing valid and reliable descriptions of patients to drive both discovery and intervention.

APPENDIX

Members of HiTOP Utility Workgroup include, in addition to the authors of this paper, Kamran Afzali, Marina A. Bornovalova, Natacha Carragher, David C. Cicero, Anna R. Docherty, Michael B. First, Eiko I. Fried, Michael N. Hallquist, Katherine Jonas, Kristian E. Markon, Les C. Morey, Stephanie N. Mullins-Sweatt, Kristin Naragon-Gainey, Brady Nelson, Thomas M. Olino, Praveetha Patalay, Aaron L. Pincus, Craig Rodriguez-Seijas, Lauren A. Rutter, Giovanni A. Salum, Alexander J. Shackman, Andrew E. Skodol, Kathryn Tabb, Nicholas Wagner, Ashley L. Watts, Amanda A. Uliaszek, Johannes Zimmermann and Richard E. Zinbarg.

ACKNOWLEDGEMENTS

R.F. Krueger is partially supported by the US National Institutes of Health (grants no. R01AG053217, U19AG051426). K. Hobbs is supported by the National Science Foundation Graduate Research Fellowship Program (grant no. 00074041). K. Forbush is supported by the US Department of Defense, and the University of Kansas, Research Excellence Initiative. D. Dick is supported by the US National Institutes of Health (grants no. K02AA018755, P50AA022537, R01DA050721). Further information on the HiTOP consortium can be found at http://medicine.stonybrookmedicine.edu/HITOP. D. Watson and R. Kotov are joint senior authors of this paper.

REFERENCES

1. Kotov R, Krueger RF, Watson D et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): a dimensional alternative to traditional nosologies. J Abnorm Psychol 2017;126:454-77.
2. Kotov R, Krueger RF, Watson D. A paradigm shift in psychiatric classification: the Hierarchical Taxonomy Of Psychopathology (HiTOP). World Psychiatry 2018;17:24-5.
3. Krueger RF, Kotov R, Watson D et al. Progress in achieving quantitative classification of psychopathology. World Psychiatry 2018;17:282-93.
4. Markon KE, Chmielewski M, Miller CJ. The reliability and validity of discrete and continuous measures of psychopathology: a quantitative review. Psychol Bull 2011;137:856-79.
5. Shankman SA, Funkhouser CJ, Klein DN et al. Reliability and validity of severity dimensions of psychopathology assessed using the Structured Clinical Interview for DSM-5 (SCID). Int J Methods Psychiatr Res 2018;27:e1590.
6. Chmielewski M, Clark LA, Bagby RM et al. Method matters: understanding diagnostic reliability in DSM-IV and DSM-5. J Abnorm Psychol 2015;124:764-9.
7. Watson D. Subtypes, specifiers, epicycles, and eccentrics: toward a more parsimonious taxonomy of psychopathology. Clin Psychol Sci Pract 2003; 10:233-8.
8. Cuesta MJ, Peralta V. Integrating psychopathological dimensions in functional psychoses: a hierarchical approach. Schizophr Res 2001;52:215-29.
9. Kotov R, Foti D, Li K et al. Validating dimensions of psychosis symptomatology: neural correlates and 20-year outcomes. J Abnorm Psychol 2016; 125:1103-19.
10. Reininghaus U, Böhnke JR, Chavez-Baldini U et al. Transdiagnostic dimensions of psychosis in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). World Psychiatry 2019;18:67-76.
11. Ruggero CJ, Kotov R, Watson D et al. Beyond a single index of mania symptoms: structure and validity of subdimensions. J Affect Disord 2014;161: 8-15.
12. Anderson JL, Sellbom M, Ayearst L et al. Associations between DSM-5 section III personality traits and the Minnesota Multiphasic Personality Inventory 2-Restructured Form (MMPI-2-RF) scales in a psychiatric patient sample. Psychol Assess 2015;27:801-15.
13. Forbes MK, Kotov R, Ruggero CJ et al. Delineating the joint hierarchical structure of clinical and personality disorders in an outpatient psychiatric sample. Compr Psychiatry 2017;79:19-30.
14. Krueger RF, Chentsova-Dutton YE, Markon KE et al. A cross-cultural study of the structure of comorbidity among common psychopathological syndromes in the general health care setting. J Abnorm Psychol 2003;112:437-47.
15. Krueger RF, Markon KE, Patrick CJ et al. Linking antisocial behavior, substance use, and personality: an integrative quantitative model of the adult externalizing spectrum. J Abnorm Psychol 2007;116:645-66.
16. Patrick CJ, Kramer MD, Krueger RF et al. Optimizing efficiency of psychopathology assessment through quantitative modeling: development of a brief form of the Externalizing Spectrum Inventory. Psychol Assess 2013;25:1332-48.
17. Caspi A, Moffitt TE. All for one and one for all: mental disorders in one dimension. Am J Psychiatry 2018;175:831-44.
18. Lahey BB, Krueger RF, Rathouz PJ et al. A hierarchical causal taxonomy of psychopathology across the life span. Psychol Bull 2016;143:142-86.
19. Kotov R, Jonas KG, Carpenter WT et al. Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): I. Psychosis superspectrum. World Psychiatry 2020;19:151-72.
20. Mullins-Sweatt SN, DeShong HL, Lengel GJ et al. Disinhibition as a unifying construct in understanding how personality dispositions undergird psychopathology. J Res Personal 2019;80:55-61.
21. Lynam DR, Miller JD. The basic trait of antagonism: an unfortunately underappreciated construct. J Res Personal 2019;81:118-26.
22. DeYoung CG, Chmielewski M, Clark LA et al. The distinction between symptoms and traits in the Hierarchical Taxonomy of Psychopathology (HiTOP). Submitted for publication.
23. Krueger RF, Eaton NR. Transdiagnostic factors of mental disorders. World Psychiatry 2015;14:27-9.
24. Blanco C, Wall MM, He J-P et al. The space of common psychiatric disorders in adolescents: comorbidity structure and individual latent liabilities. J Am Acad Child Adolesc Psychiatry 2015;54:45-52.
25. Castellanos-Ryan N, Struve M, Whelan R et al. Neural and cognitive correlates of the common and specific variance across externalizing problems in young adolescence. Am J Psychiatry 2014;171:1310-9.
26. Gomez R, Stavropoulos V, Vance A et al. Re-evaluation of the latent structure of common childhood disorders: is there a general psychopathology factor (P-factor)? Int J Ment Health Addict 2019;17:258-78.
27. Lahey BB, Rathouz PJ, Van Hulle C et al. Testing structural models of DSM-IV symptoms of common forms of child and adolescent psychopathology. J Abnorm Child Psychol 2008;36:187-206.
28. Martel MM, Pan PM, Hoffmann MS et al. A general psychopathology factor (P factor) in children: structural model analysis and external validation through familial risk and child global executive function. J Abnorm Psychol 2017;126:137-48.
29. Olino TM, Bufferd SJ, Dougherty LR et al. The development of latent dimensions of psychopathology across early childhood: stability of dimensions and moderators of change. J Abnorm Child Psychol 2018;46:1373-83.
30. Verona E, Javdani S, Sprague J. Comparing factor structures of adolescent psychopathology. Psychol Assess 2011;23:545-51.
31. Caspi A, Houts RM, Belsky DW et al. The p factor: one general psychopathology factor in the structure of psychiatric disorders? Clin Psychol Sci 2014; 2:119-37.
32. Keyes KM, Eaton NR, Krueger RF et al. Thought disorder in the meta-structure of psychopathology. Psychol Med 2013;43:1673-83.
33. Verona E, Sachs-Ericsson N, Joiner TE. Suicide attempts associated with externalizing psychopathology in an epidemiological sample. Am J Psychiatry 2004;161:444-51.
34. Wright AGC, Krueger RF, Hobbs MJ et al. The structure of psychopathology: toward an expanded quantitative empirical model. J Abnorm Psychol 2013;122:281-94.
35. Krueger RF, Caspi A, Moffitt TE et al. The structure and stability of common mental disorders (DSM-III-R): a longitudinal-epidemiological study. J Abnorm Psychol 1998;107:216-27.
36. Beesdo-Baum K, Höfler M, Gloster AT et al. The structure of common mental disorders: a replication study in a community sample of adolescents and young adults. Int J Methods Psychiatr Res 1998;18:204-20.
37. Wittchen HU, Beesdo-Baum K, Gloster AT et al. The structure of mental disorders re-examined: Is it developmentally stable and robust against additions?. Int J Methods Psychiatr Res 2009;18:189-203.
38. Carragher N, Krueger RF, Eaton NR et al. ADHD and the externalizing spectrum: direct comparison of categorical, continuous, and hybrid models of liability in a nationally representative sample. Soc Psychiatry Psychiatr Epidemiol 2014;49:1307-17.

39. Eaton NR, Krueger RF, Keyes KM et al. Borderline personality disorder comorbidity: relationship to the internalizing-externalizing structure of common mental disorders. Psychol Med 2011;41:1041-50.
40. Lahey BB, Applegate B, Hakes JK et al. Is there a general factor of prevalent psychopathology during adulthood? J Abnorm Psychol 2012;121:971-7.
41. Waldman ID, Poore HE, van Hulle C et al. External validity of a hierarchical dimensional model of child and adolescent psychopathology: tests using confirmatory factor analyses and multivariate behavior genetic analyses. J Abnorm Psychol 2016;125:1053-66.
42. De Jonge P, Wardenaar KJ, Lim CC et al. The cross-national structure of mental disorders: results from the World Mental Health Surveys. Psychol Med 2018;48:2073-84.
43. Kessler RC, Ormel J, Petukhova M et al. Development of lifetime comorbidity in the World Health Organization World Mental Health Surveys. Arch Gen Psychiatry 2011;68:90-100.
44. Conway CC, Mansolf M, Reise SP. Ecological validity of a quantitative classification system for mental illness in treatment-seeking adults. Psychol Assess 2019;31:730-40.
45. Cox BJ, Clara IP, Enns MW. Posttraumatic stress disorder and the structure of common mental disorders. Depress Anxiety 2002;15:168-71.
46. Forbush KT, Watson D. The structure of common and uncommon mental disorders. Psychol Med 2013;43:97-108.
47. Forbush KT, South SC, Krueger RF et al. Locating eating pathology within an empirical diagnostic taxonomy: evidence from a community-based sample. J Abnorm Psychol 2010;119:282-92.
48. James LM, Taylor J. Revisiting the structure of mental disorders: borderline personality disorder and the internalizing/externalizing spectra. Br J Clin Psychol 2008;47:361-80.
49. Krueger RF. The structure of common mental disorders. Arch Gen Psychiatry 1999;56:921-6.
50. Krueger RF, Hicks BM, Patrick CJ et al. Etiologic connections among substance dependence, antisocial behavior and personality: modeling the externalizing spectrum. J Abnorm Psychol 2002;111:411-24.
51. Miller MW, Fogler JM, Wolf EJ et al. The internalizing and externalizing structure of psychiatric comorbidity in combat veterans. J Trauma Stress 2018;21:58-65.
52. Miller MW, Wolf EJ, Reardon A et al. Personality and the latent structure of PTSD comorbidity. J Anxiety Disord 2012;26:599-607.
53. Tuvblad C, Zheng M, Raine A et al. A common genetic factor explains the covariation among ADHD ODD and CD symptoms in 9-10 year old boys and girls. J Abnorm Child Psychol 2009;37:153-67.
54. Young SE, Stallings MC, Corley RP et al. Genetic and environmental influences on behavioral disinhibition. Am J Med Genet 2000;96:684-95.
55. Kotov R, Ruggero CJ, Krueger RF et al. New dimensions in the quantitative classification of mental illness. Arch Gen Psychiatry 2011;68:1003-11.
56. Kendler KS, Aggen SH, Knudsen GP et al. The structure of genetic and environmental risk factors for syndromal and subsyndromal common DSM-IV Axis I and Axis II disorders. Am J Psychiatry 2011;168:29-39.
57. Røysamb E, Kendler KS, Tambs K et al. The joint structure of DSM-IV Axis I and Axis II disorders. J Abnorm Psychol 2011;120:198-209.
58. Conway CC, Brown TA. Evaluating dimensional models of psychopathology in outpatients diagnosed with emotional disorders: a cautionary tale. Depress Anxiety 2018;35:898-902.
59. Conway CC, Starr LR, Espejo EP et al. Stress responsivity and the structure of common mental disorders: transdiagnostic internalizing and externalizing dimensions are associated with contrasting stress appraisal biases. J Abnorm Psychol 2016;125:1079-89.
60. Farmer RF, Seeley JR, Kosty DB et al. Refinements in the hierarchical structure of externalizing psychiatric disorders: patterns of lifetime liability from mid-adolescence through early adulthood. J Abnorm Psychol 2009;118:699-710.
61. Kim H, Eaton NR. The hierarchical structure of common mental disorders: connecting multiple levels of comorbidity, bifactor models, and predictive validity. J Abnorm Psychol 2015;124:1064-78.
62. Slade T, Watson D. The structure of common DSM-IV and ICD-10 mental disorders in the Australian general population. Psychol Med 2006;36:1593-1600.
63. Vollebergh WA, Iedema J, Bijl RV et al. The structure and stability of common mental disorders: the NEMESIS study. Arch Gen Psychiatry 2001;58:597-603.
64. Wright AGC, Simms LJ. A metastructural model of mental disorders and pathological personality traits. Psychol Med 2015;45:2309-19.
65. Kotov R. New dimensions in the quantitative classification of mental illness. Arch Gen Psychiatry 2011;68:1003-11.
66. Kendler KS, Myers JM, Maes HH et al. The relationship between the genetic and environmental influences on common internalizing psychiatric disorders and mental well-being. Behav Genet 2011;41:641-50.
67. Samuel D, Widiger T. A meta-analytic review of the relationships between the five-factor model and DSM-IV-TR personality disorders: a facet level analysis. Clin Psychol Rev 2008;28:1326-42.
68. Saulsman LM, Page AC. The five-factor model and personality disorder empirical literature: a meta-analytic review. Clin Psychol Rev 2004;23:1055-85.
69. Shiner RL, Tackett JL. Personality disorders in children and adolescents. In: Mash EJ, Barkley RA (eds). Child psychopathology. New York: Guilford, 2014:848-96.
70. Frick PJ, Myers TDW. Conduct disorder and callous-unemotional traits. In: Lochman JE, Matthys W (eds). The Wiley handbook of disruptive and impulse-control disorders. Chichester: Wiley-Blackwell, 2018:37-54.
71. Widiger TA, Sellbom M, Chmielewski M et al. Personality in a hierarchical model of psychopathology. Clin Psychol Sci 2019;7:77-92.
72. Krueger RF, South SC. Externalizing disorders: cluster 5 of the proposed meta-structure for DSM-V and ICD-11. Psychol Med 2009;39:2061-70.
73. Markon KE, Krueger RF, Watson D. Delineating the structure of normal and abnormal personality: an integrative hierarchical approach. J Pers Soc Psychol 2005;88:139-57.
74. Rothbart MK, Ahadi SA, Evans DE. Temperament and personality: origins and outcomes. J Pers Soc Psychol 2000;78:122-35.
75. Watson D, Clark LA, Chmielewski M. Structures of personality and their relevance to psychopathology: II. Further articulation of a comprehensive unified trait structure. J Pers 2008;76:1545-86.
76. Tackett JL, Slobodskaya HR, Mar RA et al. The hierarchical structure of childhood personality in five countries: continuity from early childhood to early adolescence. J Pers 2012;80:847-79.
77. Crego C, Widiger TA. Convergent and discriminant validity of alternative measures of maladaptive personality traits. Psychol Assess 2016;28:1561-75.
78. 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.
79. Tackett JL, Martel MM, Kushner SC. Temperament, externalizing disorders, and attention-deficit/hyperactivity disorder. In: Zentner M, Shiner RL (eds). Handbook of temperament. New York: Guilford, 2012:562-80.
80. Watson D, Clark LA. Personality traits as an organizing framework for personality pathology. Personal Ment Health 2020;14:51-75.
81. Watson D, Stasik SM, Ro E et al. Integrating normal and pathological personality: relating the DSM-5 Trait-Dimensional Model to general traits of personality. Assessment 2013;20:312-26.
82. 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.
83. 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 Theory Res Treat 2014;5:43-54.
84. Rojas SL, Widiger TA. Coverage of the DSM-IV-TR/DSM-5 Section II Personality Disorders with the DSM-5 Dimensional Trait Model. J Personal Disord 2017;31:462-82.
85. 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 Theory Res Treat 2018;10:97-104.
86. Burke JD, Boylan K, Rowe R et al. Identifying the irritability dimension of ODD: application of a modified bifactor model across five large community samples of children. J Abnorm Psychol 2014;123:841-51.
87. Coccaro EF, Lee R, McCloskey MS. Relationship between psychopathy, aggression, anger, impulsivity, and intermittent explosive disorder. Aggress Behav 2014;40:526-36.
88. Fernandez E, Johnson SL. Anger in psychological disorders: prevalence, presentation, etiology and prognostic implications. Clin Psychol Rev 2016;46:124-35.
89. Gomez R, Corr PJ. ADHD and personality: a meta-analytic review. Clin Psychol Rev 2014;34:376-88.
90. Herzhoff K, Tackett JL. Subfactors of oppositional defiant disorder: converging evidence from structural and latent class analyses. J Child Psychol Psychiatry 2016;57:18-29.
91. Johnston OG, Cruess DG, Burke JD. Irritability and behavioral symptom dimensions of oppositional defiant disorder in young adults: associations with DSM-5 pathological personality traits. J Psychopathol Behav Assess 2020;42:424-35.

92. Reardon KW, Tackett JL, Lynam D. The personality context of relational aggression: a Five-Factor Model profile analysis. Personal Disord Theory Res Treat 2017;9:228-38.
93. Waldman ID, Rowe R, Boylan K et al. External validation of a bifactor model of oppositional defiant disorder. Mol Psychiatry 2021;26:682-93.
94. Nikolas MA, Burt SA. Genetic and environmental influences on ADHD symptom dimensions of inattention and hyperactivity: a meta-analysis. J Abnorm Psychol 2010;119:1-17.
95. Ficks CA, Dong L, Waldman ID. Sex differences in the etiology of psychopathic traits in youth. J Abnorm Psychol 2014;123:406-11.
96. Rhee SH, Waldman ID. Genetic and environmental influences on antisocial behavior: a meta-analysis of twin and adoption studies. Psychol Bull 2002;128:490-529.
97. Burt SA. Are there meaningful etiological differences within antisocial behavior? Results of a meta-analysis. Clin Psychol Rev 2009;29:163-78.
98. Tackett JL, Krueger RF, Iacono WG et al. Symptom-based subfactors of DSM-defined conduct disorder: evidence for etiologic distinctions. J Abnorm Psychol 2005;114:483-7.
99. Tackett JL, Waldman ID, Lahey BB. Etiology and measurement of relational aggression: a multi-informant behavior genetic investigation. J Abnorm Psychol 2009;118:722-33.
100. Burt SA, Krueger RF, McGue M et al. Sources of covariation among attention-deficit/hyperactivity disorder, oppositional defiant disorder, and conduct disorder: the importance of shared environment. J Abnorm Psychol 2001;110:516-25.
101. Dick DM, Viken RJ, Kaprio J et al. Understanding the covariation among childhood externalizing symptoms: genetic and environmental Influences on conduct disorder, attention deficit hyperactivity disorder, and oppositional defiant disorder symptoms. J Abnorm Child Psychol 2005;33:219-29.
102. Cosgrove VE, Rhee SH, Gelhorn HL et al. Structure and etiology of co-occurring internalizing and externalizing disorders in adolescents. J Abnorm Child Psychol 2011;39:109-23.
103. Hicks BM, Krueger RF, Iacono WG et al. Family transmission and heritability of externalizing disorders: a twin-family study. Arch Gen Psychiatry 2004;61:922-8.
104. Waldman ID, Tackett JL, Van Hulle CA et al. Child and adolescent conduct disorder substantially shares genetic influences with three socioemotional dispositions. J Abnorm Psychol 2011;120:57-70.
105. Kendler KS, Karkowski LM, Neale MC et al. Illicit psychoactive substance use, heavy use, abuse, and dependence in a US population-based sample of male twins. Arch Gen Psychiatry 2000;57:261-9.
106. Torgersen S, Czajkowski N, Jacobson K et al. Dimensional representations of DSM-IV cluster B personality disorders in a population-based sample of Norwegian twins: a multivariate study. Psychol Med 2008;38:1617-25.
107. South SC, Krueger RF, Knudsen GP et al. A population based twin study of DSM–5 maladaptive personality domains. Personal Disord Theory Res Treat 2017;8:366-75.
108. 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.
109. Kendler KS, Aggen SH, Gillespie N et al. The structure of genetic and environmental influences on normative personality, abnormal personality traits, and personality disorder symptoms. Psychol Med 2019;49:1392-9.
110. Verhulst B, Neale MC, Kendler KS. The heritability of alcohol use disorders: a meta-analysis of twin and adoption studies. Psychol Med 2015;45:1061-72.
111. Verweij KJH, Zietsch BP, Lynskey MT et al. Genetic and environmental influences on cannabis use initiation and problematic use: a meta-analysis of twin studies. Addiction 2010;105:417-30.
112. Tsuang MT, Bar JL, Harley RM et al. The Harvard Twin Study of Substance Abuse: what we have learned. Harv Rev Psychiatry 2001;9:267-79.
113. van den Bree MBM, Johnson EO, Neale MC et al. Genetic and environmental influences on drug use and abuse/dependence in male and female twins. Drug Alcohol Depend 1998;52:231-41.
114. Maes HH, Sullivan PF, Bulik CM et al. A twin study of genetic and environmental influences on tobacco initiation, regular tobacco use and nicotine dependence. Psychol Med 2004;34:1251-61.
115. Kendler KS, Myers J, Prescott CA. Specificity of genetic and environmental risk factors for symptoms of cannabis, cocaine, alcohol, caffeine, and nicotine dependence. Arch Gen Psychiatry 2007;64:1313-20.
116. Kendler KS, Prescott CA, Myers J et al. The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. Arch Gen Psychiatry 2003;60:929-37.
117. Gizer IR, Ficks C, Waldman ID. Candidate gene studies of ADHD: a meta-analytic review. Hum Genet 2009;126:51-90.
118. Demontis D, Walters RK, Martin J et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Genet 2019;51:63-75.
119. Dick DM, Li T-K, Edenberg HJ et al. A genome-wide screen for genes influencing conduct disorder. Mol Psychiatry 2004;9:81-6.
120. Demontis D, Walters R, Rajagopal VM et al. Identification of risk variants and characterization of the polygenic architecture of disruptive behavior disorders in the context of ADHD. bioRxiv 2019: 791160.
121. Tielbeek JJ, Johansson A, Polderman TJC et al. Genome-wide association studies of a broad spectrum of antisocial behavior. JAMA Psychiatry 2017;74:1242-50.
122. Witt SH, Streit F, Jungkunz M et al. Genome-wide association study of borderline personality disorder reveals genetic overlap with bipolar disorder, major depression and schizophrenia. Transl Psychiatry 2017;7:e1155.
123. Tielbeek JJ, Vink JM, Polderman TJC et al. Genetic correlation of antisocial behaviour with alcohol, nicotine, and cannabis use. Drug Alcohol Depend 2018;187:296-9.
124. Chang L-H, Whitfield JB, Liu M et al. Associations between polygenic risk for tobacco and alcohol use and liability to tobacco and alcohol use, and psychiatric disorders in an independent sample of 13,999 Australian adults. Drug Alcohol Depend 2019;205:107704.
125. Salvatore JE, Edwards AC, McClintick JN et al. Genome-wide association data suggest ABCB1 and immune-related gene sets may be involved in adult antisocial behavior. Transl Psychiatry 2015;5:e558.
126. Lo M-T, Hinds DA, Tung JY et al. Genome-wide analyses for personality traits identify six genomic loci and show correlations with psychiatric disorders. Nat Genet 2017;49:152-6.
127. Walters RK, Polimanti R, Johnson EC et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. Nat Neurosci 2018;21:1656-69.
128. Kranzler HR, Zhou H, Kember RL et al. Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. Nat Commun 2019;10:1499.
129. Liu M, Jiang Y, Wedow R et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet 2019;51:237-44.
130. Johnson EC, Sanchez-Roige S, Acion L et al. Polygenic contributions to alcohol use and alcohol use disorders across population-based and clinically ascertained samples. Psychol Med (in press).
131. Pasman JA, Verweij KJH, Gerring Z et al. GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal effect of schizophrenia liability. Nat Neurosci 2018;21:1161-70.
132. Minică CC, Verweij KJH, van der Most PJ et al. Genome-wide association meta-analysis of age at first cannabis use. Addiction 2018;113:2073-86.
133. Demontis D, Rajagopal VM, Thorgeirsson TE et al. Genome-wide association study implicates CHRNA2 in cannabis use disorder. Nat Neurosci 2019;22:1066-74.
134. Gelernter J, Sherva R, Koesterer R et al. Genome-wide association study of cocaine dependence and related traits: FAM53B identified as a risk gene. Mol Psychiatry 2014;19:717-23.
135. Cabana-Domínguez J, Shivalikanjli A, Fernàndez-Castillo N et al. Genome-wide association meta-analysis of cocaine dependence: shared genetics with comorbid conditions. Prog Neuropsychopharmacol Biol Psychiatry 2019;94:109667.
136. Cheng Z, Zhou H, Sherva R et al. Genome-wide association study identifies a regulatory variant of RGMA associated with opioid dependence in European Americans. Biol Psychiatry 2018;84:762-70.
137. Zhou H, Rentsch CT, Cheng Z et al. GWAS including 82,707 subjects identifies functional coding variant in OPRM1 gene associated with opioid use disorder. medRxiv 2019:19007039.
138. Polimanti R, Walters RK, Johnson EC et al. Leveraging genome-wide data to investigate differences between opioid use vs. opioid dependence in 41,176 individuals from the Psychiatric Genomics Consortium. Mol Psychiatry 2020;25:1673-87.
139. Grotzinger AD, Rhemtulla M, de Vlaming R et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. Nat Hum Behav 2019;3:513-25.
140. Waldman ID, Poore HE, Luningham JM et al. Testing structural models of psychopathology at the genomic level. World Psychiatry 2020;19:350-9.

141. Karlsson Linnér RK, Mallard TT, Barr PB et al. Using the genetic architecture of substance use disorders to aid in gene identification: findings from the Externalizing Consortium. Presented at the NIDA Genetics and Epigenetics Cross-Cutting Research Team Meeting, Rockville, January 2020.
142. Cicchetti D, Handley ED. Child maltreatment and the development of substance use and disorder. Neurobiol Stress 2019;10:100144.
143. Deater-Deckard K, Dodge KA, Bates JE et al. Multiple risk factors in the development of externalizing behavior problems: group and individual differences. Dev Psychopathol 1998;10:469-93.
144. Pinquart M. Associations of parenting dimensions and styles with externalizing problems of children and adolescents: an updated meta-analysis. Dev Psychol 2017;53:873-932.
145. Samek DR, Goodman RJ, Erath SA et al. Antisocial peer affiliation and externalizing disorders in the transition from adolescence to young adulthood: selection versus socialization effects. Dev Psychol 2016;52:813-23.
146. McKee L, Colletti C, Rakow A et al. Parenting and child externalizing behaviors: are the associations specific or diffuse? Aggress Violent Behav 2008; 13:201-15.
147. Keyes KM, Eaton NR, Krueger RF et al. Childhood maltreatment and the structure of common psychiatric disorders. Br J Psychiatry 2012;200:107-15.
148. Conway CC, Raposa EB, Hammen C et al. Transdiagnostic pathways from early social stress to psychopathology: a 20-year prospective study. J Child Psychol Psychiatry 2018;59:855-62.
149. Vachon DD, Krueger RF, Rogosch FA et al. Assessment of the harmful psychiatric and behavioral effects of different forms of child maltreatment. JAMA Psychiatry 2015;72:1135-42.
150. Forbes MK, Magson NR, Rapee RM. Evidence that different types of peer victimization have equivalent associations with transdiagnostic psychopathology in adolescence. J Youth Adolesc 2020;49:590-604.
151. Rodriguez-Seijas C, Stohl M, Hasin DS et al. Transdiagnostic factors and mediation of the relationship between perceived racial discrimination and mental disorders. JAMA Psychiatry 2015;72:706-13.
152. Conway CC, Hammen C, Brennan PA. Expanding stress generation theory: test of a transdiagnostic model. J Abnorm Psychol 2012;121:754-66.
153. Snyder HR, Young JF, Hankin BL. Chronic stress exposure and generation are related to the p-factor and externalizing specific psychopathology in youth. J Clin Child Adolesc Psychol 2019;48:306-15.
154. Ensink JBM, de Moor MHM, Zafarmand MH et al. Maternal environmental risk factors and the development of internalizing and externalizing problems in childhood: the complex role of genetic factors. Am J Med Genet B Neuropsychiatr Genet 2020;183:17-25.
155. Morgan AB, Lilienfeld SO. A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. Clin Psychol Rev 2000;20:113-36.
156. Ogilvie JM, Stewart AL, Chan RCK et al. Neuropsychological measures of executive function and antisocial behavior: a meta-analysis. Criminology 2011;49:1063-107.
157. Pasion R, Fernandes C, Pereira MR et al. Antisocial behaviour and psychopathy: uncovering the externalizing link in the P3 modulation. Neurosci Biobehav Rev 2018;91:170-86.
158. Bernat EM, Nelson LD, Steele VR et al. Externalizing psychopathology and gain-loss feedback in a simulated gambling task: dissociable components of brain response revealed by time-frequency analysis. J Abnorm Psychol 2011;120:352-64.
159. Friedman NP, Rhee SH, Ross JM et al. Genetic and environmental relations of executive functions to antisocial personality disorder symptoms and psychopathy. Int J Psychophysiol (in press).
160. Romer D, Betancourt LM, Brodsky NL et al. Does adolescent risk taking imply weak executive function? A prospective study of relations between working memory performance, impulsivity, and risk taking in early adolescence. Dev Sci 2011;14:1119-33.
161. Lansing AE, Plante WY, Golshan S et al. Emotion regulation mediates the relationship between verbal learning and internalizing, trauma-related and externalizing symptoms among early-onset, persistently delinquent adolescents. Learn Individ Differ 2019;70:201-15.
162. Becker SP, Leopold DR, Burns GL et al. The internal, external, and diagnostic validity of sluggish cognitive tempo: a meta-analysis and critical review. J Am Acad Child Adolesc Psychiatry 2016;55:163-78.
163. Fragkaki I, Cima M, Meesters C. The association between callous-unemotional traits, externalizing problems, and gender in predicting cognitive and affective morality judgments in adolescence. J Youth Adolesc 2016;45:1917-30.
164. McDonald JB, Bozzay ML, Bresin K et al. Facets of externalizing psychopathology in relation to inhibitory control and error processing. Int J Psychophysiol (in press).
165. Endres MJ, Rickert ME, Bogg T et al. Externalizing psychopathology and behavioral disinhibition: working memory mediates signal discriminability and reinforcement moderates response bias in approach-avoidance learning. J Abnorm Psychol 2011;120:336-51.
166. Brennan GM, Baskin-Sommers AR. Brain-behavior relationships in externalizing: P3 amplitude reduction reflects deficient inhibitory control. Behav Brain Res 2018;337:70-9.
167. Pan XX, Ma HW, Dai XM. Value of integrated visual and auditory continuous performance test in the diagnosis of childhood attention deficit hyperactivity disorder. Chin J Contemp Pediatr 2007;9:210-2.
168. Slobodin O, Cassuto H, Berger I. Age-related changes in distractibility: developmental trajectory of sustained attention in ADHD. J Atten Disord 2018;22:1333-43.
169. Slobodin O, Blankers M, Kapitány-Fövény M et al. Differential diagnosis in patients with substance use disorder and/or attention-deficit/hyperactivity disorder using continuous performance test. Eur Addict Res 2020;26:151-62.
170. Berger I, Slobodin O, Cassuto H. Usefulness and validity of continuous performance tests in the diagnosis of attention-deficit hyperactivity disorder children. Arch Clin Neuropsychol 2017;32:81-93.
171. Zeeuw PD, Aarnoudse-Moens C, Bijlhout J et al. Inhibitory performance, response speed, intraindividual variability, and response accuracy in ADHD. J Am Acad Child Adolesc Psychiatry 2008;47:808-16.
172. Huang-Pollock CL, Karalunas SL, Tam H et al. Evaluating vigilance deficits in ADHD: a meta-analysis of CPT performance. J Abnorm Psychol 2012;121: 360-71.
173. Xu M, Jiang W, Du Y et al. Executive function features in drug-naive children with oppositional defiant disorder. Shanghai Arch Psychiatry 2017;29:228-36.
174. Noordermeer SDS, Luman M, Oosterlaan J. A systematic review and meta-analysis of neuroimaging in oppositional defiant disorder (ODD) and conduct disorder (CD) taking attention-deficit hyperactivity disorder (ADHD) into account. Neuropsychol Rev 2016;26:44-72.
175. Connor DF, Chartier KG, Preen EC et al. Impulsive aggression in attention-deficit/hyperactivity disorder: symptom severity, co-morbidity, and attention-deficit/hyperactivity disorder subtype. J Child Adolesc Psychopharmacol 2010;20:119-26.
176. Frick PJ, Nigg JT. Current issues in the diagnosis of attention deficit hyperactivity disorder, oppositional defiant disorder, and conduct disorder. Annu Rev Clin Psychol 2012;8:77-107.
177. McCloskey MS, Phan KL, Angstadt M et al. Amygdala hyperactivation to angry faces in intermittent explosive disorder. J Psychiatr Res 2016;79:34-41.
178. Coccaro EF, Fanning JR, Keedy SK et al. Social cognition in intermittent explosive disorder and aggression. J Psychiatr Res 2016;83:140-50.
179. Fahlgren MK, Puhalla AA, Sorgi KM et al. Emotion processing in intermittent explosive disorder. Psychiatry Res 2019;273:544-50.
180. Fettich KC, McCloskey MS, Look AE et al. Emotion regulation deficits in intermittent explosive disorder. Aggress Behav 2015;41:25-33.
181. Izquierdo A, Jentsch JD. Reversal learning as a measure of impulsive and compulsive behavior in addictions. Psychopharmacology 2012;219:607-20.
182. Volkow ND, Li T-K. Drug addiction: the neurobiology of behaviour gone awry. Nat Rev Neurosci 2004;5:963-70.
183. Finn PR, Rickert ME, Miller MA et al. Reduced cognitive ability in alcohol dependence: examining the role of covarying externalizing psychopathology. J Abnorm Psychol 2009;118:100-16.
184. Kirisci L, Tarter R, Ridenour T et al. Externalizing behavior and emotion dysregulation are indicators of transmissible risk for substance use disorder. Addict Behav 2015;42:57-62.
185. Verdejo-García A, Pérez-García M. Profile of executive deficits in cocaine and heroin polysubstance users: common and differential effects on separate executive components. Psychopharmacology 2007;190:517-30.
186. Lundqvist T. Cognitive consequences of cannabis use: comparison with abuse of stimulants and heroin with regard to attention, memory and executive functions. Pharmacol Biochem Behav 2005;81:319-30.
187. Smith JL, Mattick RP, Jamadar SD et al. Deficits in behavioural inhibition in substance abuse and addiction: a meta-analysis. Drug Alcohol Depend 2014;145:1-33.
188. Verdejo-García A, Lawrence AJ, Clark L. Impulsivity as a vulnerability marker for substance-use disorders: review of findings from high-risk research, problem gamblers and genetic association studies. Neurosci Biobehav Rev 2008;32:777-810.

189. Ruocco AC. The neuropsychology of borderline personality disorder: a meta-analysis and review. Psychiatry Res 2005;137:191-202.
190. Kalpakci A, Ha C, Sharp C. Differential relations of executive functioning to borderline personality disorder presentations in adolescents. Personal Ment Health 2018;12:93-106.
191. Coolidge FL, Thede LL, Jang KL. Are personality disorders psychological manifestations of executive function deficits? Bivariate heritability evidence from a twin study. Behav Genet 2004;34:75-84.
192. Herpertz SC, Bertsch K. The social-cognitive basis of personality disorders. Curr Opin Psychiatry 2014;27:73-7.
193. Lovett BJ, Sheffield RA. Affective empathy deficits in aggressive children and adolescents: a critical review. Clin Psychol Rev 2007;27:1-13.
194. Domes G, Schulze L, Herpertz SC. Emotion recognition in borderline personality disorder – a review of the literature. J Personal Disord 2009;23:6-19.
195. Frick PJ, White SF. Research Review: The importance of callous-unemotional traits for developmental models of aggressive and antisocial behavior. J Child Psychol Psychiatry 2008;49:359-75.
196. Glenn AL, Johnson AK, Raine A. Antisocial personality disorder: a current review. Curr Psychiatry Rep 2013;15:427.
197. O'Kearney R, Salmon K, Liwag M et al. Emotional abilities in children with oppositional defiant disorder (ODD): impairments in perspective-taking and understanding mixed emotions are associated with high callous-unemotional traits. Child Psychiatry Hum Dev 2017;48:346-57.
198. Jiang W, Li Y, Du Y et al. Emotional regulation and executive function deficits in unmedicated Chinese children with oppositional defiant disorder. Psychiatry Investig 2016;13:277-87.
199. Verdejo-García A, Pérez-García M, Bechara A. Emotion, decision-making and substance dependence: a somatic-marker model of addiction. Curr Neuropharmacol 2006;4:17-31.
200. Verdejo-García A, Rivas-Pérez C, Vilar-López R et al. Strategic self-regulation, decision-making and emotion processing in poly-substance abusers in their first year of abstinence. Drug Alcohol Depend 2007;86:139-46.
201. Wilcox CE, Pommy JM, Adinoff B. Neural circuitry of impaired emotion regulation in substance use disorders. Am J Psychiatry 2016;173:344-61.
202. Herpertz SC, Schneider I, Schmahl C et al. Neurobiological mechanisms mediating emotion dysregulation as targets of change in borderline personality disorder. Psychopathology 2018;51:96-104.
203. Carpenter RW, Trull TJ. Components of emotion dysregulation in borderline personality disorder: a review. Curr Psychiatry Rep 2013;15:335.
204. Baskin-Sommers A, Krusemark E, Ronningstam E. Empathy in narcissistic personality disorder: from clinical and empirical perspectives. Personal Disord 2014;5:323-33.
205. Lee DS. A longitudinal analysis of executive functions, learning-related skills, and mathematics achievement. Dissertation, University of California, Irvine, 2017.
206. Gao Y, Raine A. P3 event-related potential impairments in antisocial and psychopathic individuals: a meta-analysis. Biol Psychol 2009;82:199-210.
207. Begleiter H, Porjesz B, Bihari B et al. Event-related brain potentials in boys at risk for alcoholism. Science 1984;225:1493-6.
208. Iacono WG, Carlson SR, Malone SM et al. P3 event-related potential amplitude and the risk for disinhibitory disorders in adolescent boys. Arch Gen Psychiatry 2002;59:750-7.
209. Patrick CJ, Bernat EM, Malone SM et al. P300 amplitude as an indicator of externalizing in adolescent males. Psychophysiology 2006;43:84-92.
210. Viding E, Sebastian CL, Dadds MR et al. Amygdala response to preattentive masked fear in children with conduct problems: the role of callous-unemotional traits. Am J Psychiatry 2012;169:1109-16.
211. Hicks BM, Bernat E, Malone SM et al. Genes mediate the association between P3 amplitude and externalizing disorders. Psychophysiology 2007;44:98-105.
212. Yancey JR, Venables NC, Hicks BM et al. Evidence for a heritable brain basis to deviance-promoting deficits in self-control. J Crim Justice 2013;41:309-17.
213. Young SE, Friedman NP, Miyake A et al. Behavioral disinhibition: liability for externalizing spectrum disorders and its genetic and environmental relation to response inhibition across adolescence. J Abnorm Psychol 2009;118:117-30.
214. Venables NC, Foell J, Yancey JR et al. Quantifying inhibitory control as externalizing proneness: a cross-domain model. Clin Psychol Sci 2018;6:561-80.
215. Iacono WG, Carlson SR, Taylor J et al. Behavioral disinhibition and the development of substance-use disorders: findings from the Minnesota Twin Family Study. Dev Psychopathol 1999;11:869-900.
216. Patrick CJ, Foell J, Venables NC et al. Substance use disorders as externalizing outcomes. In: Beauchaine T, Hinshaw S (eds). Oxford handbook of externalizing spectrum. Oxford: Oxford University Press, 2016:38-60.
217. Dikman ZV, Allen JJB. Error monitoring during reward and avoidance learning in high- and low-socialized individuals. Psychophysiology 2000;37:43-54.
218. Pailing PE, Segalowitz SJ, Dywan J et al. Error negativity and response control. Psychophysiology 2002;39:198-206.
219. Hall JR, Bernat EM, Patrick CJ. Externalizing psychopathology and the error-related negativity. Psychol Sci 2007;18:326-33.
220. Michelini G, Cheung CHM, Kitsune V et al. The etiological structure of cognitive-neurophysiological impairments in ADHD in adolescence and young adulthood. J Atten Disord 2021;25:91-104.
221. Pasion R, Cruz AR, Barbosa F. Dissociation of boldness and disinhibition psychopathic traits in ERN modulation. Personal Individ Differ 2016;95:6-10.
222. Ribes-Guardiola P, Poy R, Patrick CJ et al. Electrocortical measures of performance monitoring from go/no-go and flanker tasks: differential relations with trait dimensions of the triarchic model of psychopathy. Psychophysiology 2020;57:e13573.
223. Marsh AA. Reduced amygdala response to fearful expressions in children and adolescents with callous-unemotional traits and disruptive behavior disorders. Am J Psychiatry 2008;165:712-20.
224. Frick PJ, Ray JV, Thornton LC et al. Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. Psychol Bull 2013;140:1-57.
225. Brislin SJ, Patrick CJ. Callousness and affective face processing: clarifying the neural basis of behavioral-recognition deficits through the use of brain event-related potentials. Clin Psychol Sci 2019;7:1389-402.
226. Brislin SJ, Yancey JR, Perkins ER et al. Callousness and affective face processing in adults: behavioral and brain-potential indicators. Personal Disord Theory Res Treat 2017;9:122-32.
227. Patrick CJ, Fowles DC, Krueger RF. Triarchic conceptualization of psychopathy: developmental origins of disinhibition, boldness, and meanness. Dev Psychopathol 2009;21:913-38.
228. Vanyukov MM, Tarter RE, Kirillova GP et al. Common liability to addiction and "gateway hypothesis": theoretical, empirical and evolutionary perspective. Drug Alcohol Depend 2012;123:S3-17.
229. Perkins ER, Joyner KJ, Patrick CJ et al. Neurobiology and the Hierarchical Taxonomy of Psychopathology: progress toward ontogenetically informed and clinically useful nosology. Dialogues Clin Neurosci 2020;22:51-63.
230. Bartholow BD, Henry EA, Lust SA. Effects of alcohol sensitivity on P3 event-related potential reactivity to alcohol cues. Psychol Addict Behav 2007;21:555-63.
231. van Hemel-Ruiter ME, de Jong PJ, Ostafin BD et al. Reward sensitivity, attentional bias, and executive control in early adolescent alcohol use. Addict Behav 2015;40:84-90.
232. Blum K, Braverman ER, Holder JM et al. The reward deficiency syndrome: a biogenetic model for the diagnosis and treatment of impulsive, addictive and compulsive behaviors. J Psychoactive Drugs 2000;32:1-112.
233. Joyner KJ, Bowyer CB, Yancey JR et al. Blunted reward sensitivity and trait disinhibition interact to predict substance use problems. Clin Psychol Sci 2019;7:1109-24.
234. Weber S, Habel U, Amunts K et al. Structural brain abnormalities in psychopaths – a review. Behav Sci Law 2008;26:7-28.
235. de-Almeida CP, Wenzel A, de-Carvalho CS et al. Amygdalar volume in borderline personality disorder with and without comorbid post-traumatic stress disorder: a meta-analysis. CNS Spectr 2012;17:70-5.
236. Ruocco AC, Amirthavasagam S, Zakzanis KK. Amygdala and hippocampal volume reductions as candidate endophenotypes for borderline personality disorder: a meta-analysis of magnetic resonance imaging studies. Psychiatry Res Neuroimaging 2012;201:245-52.
237. Pardini DA, Raine A, Erickson K et al. Lower amygdala volume in men is associated with childhood aggression, early psychopathic traits, and future violence. Biol Psychiatry 2014;75:73-80.
238. Dager AD, McKay DR, Kent JW et al. Shared genetic factors influence amygdala volumes and risk for alcoholism. Neuropsychopharmacology 2015;40:412-20.
239. Hill SY, De Bellis MD, Keshavan MS et al. Right amygdala volume in adolescent and young adult offspring from families at high risk for developing alcoholism. Biol Psychiatry 2001;49:894-905.
240. 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.

241. Benegal V, Antony G, Venkatasubramanian G et al. Imaging study: gray matter volume abnormalities and externalizing symptoms in subjects at high risk for alcohol dependence. Addict Biol 2007;12:122-32.
242. Cheetham A, Allen NB, Whittle S et al. Amygdala volume mediates the relationship between externalizing symptoms and daily smoking in adolescence: a prospective study. Psychiatry Res Neuroimaging 2018;276:46-52.
243. Cardinale EM, O'Connell K, Robertson EL et al. Callous and uncaring traits are associated with reductions in amygdala volume among youths with varying levels of conduct problems. Psychol Med 2019;49:1449-58.
244. Zald DH. The human amygdala and the emotional evaluation of sensory stimuli. Brain Res Rev 2003;41:88-123.
245. Karl A, Schaefer M, Malta L et al. A meta-analysis of structural brain abnormalities in PTSD. Neurosci Biobehav Rev 2006;30:1004-31.
246. Woon FL, Sood S, Hedges DW. Hippocampal volume deficits associated with exposure to psychological trauma and posttraumatic stress disorder in adults: a meta-analysis. Prog Neuropsychopharmacol Biol Psychiatry 2010;34:1181-8.
247. Morey RA, Gold AL, LaBar KS et al. Amygdala volume changes in posttraumatic stress disorder in a large case-controlled veterans group. Arch Gen Psychiatry 2012;69:1169-78.
248. Poeppl TB, Donges MR, Mokros A et al. A view behind the mask of sanity: meta-analysis of aberrant brain activity in psychopaths. Mol Psychiatry 2019;24:463-70.
249. Ruocco AC, Amirthavasagam S, Choi-Kain LW et al. Neural correlates of negative emotionality in borderline personality disorder: an activation-likelihood-estimation meta-analysis. Biol Psychiatry 2013;73:153-60.
250. Jones AP, Laurens KR, Herba CM et al. Amygdala hypoactivity to fearful faces in boys with conduct problems and callous-unemotional traits. Am J Psychiatry 2009;166:95-102.
251. Kiehl KA. A cognitive neuroscience perspective on psychopathy: evidence for paralimbic system dysfunction. Psychiatry Res 2006;142:107-28.
252. Yang Y, Raine A. Prefrontal structural and functional brain imaging findings in antisocial, violent, and psychopathic individuals: a meta-analysis. Psychiatry Res Neuroimaging 2009;174:81-8.
253. Blair RJR. The amygdala and ventromedial prefrontal cortex: functional contributions and dysfunction in psychopathy. Philos Trans R Soc B Biol Sci 2008;363:2557-65.
254. Zald DH, McHugo M, Ray KL et al. Meta-analytic connectivity modeling reveals differential functional connectivity of the medial and lateral orbitofrontal cortex. Cereb Cortex 2014;24:232-48.
255. Boccia M, Dacquino C, Piccardi L et al. Neural foundation of human moral reasoning: an ALE meta-analysis about the role of personal perspective. Brain Imaging Behav 2017;11:278-92.
256. Shamay-Tsoory SG. The neural bases for empathy. Neuroscientist 2011;17:18-24.
257. Koenigs M, Tranel D. Pseudopsychopathy: a perspective from cognitive neuroscience. In: Zald D, Rauch S (eds). Orbitofrontal cortex. Oxford: Oxford University Press, 2006:597-619.
258. Cheetham A, Allen NB, Whittle S et al. Orbitofrontal cortex volume and effortful control as prospective risk factors for substance use disorder in adolescence. Eur Addict Res 2017;23:37-44.
259. Schoenbaum G, Shaham Y. The role of orbitofrontal cortex in drug addiction: a review of preclinical studies. Biol Psychiatry 2008;63:256-62.
260. Zilverstand A, Huang AS, Alia-Klein N et al. Neuroimaging impaired response inhibition and salience attribution in human drug addiction: a systematic review. Neuron 2018;98:886-903.
261. Qin K, Zhang F, Chen T et al. Shared gray matter alterations in individuals with diverse behavioral addictions: a voxel-wise meta-analysis. J Behav Addict 2020;9:44-57.
262. Rubia K. "Cool" inferior frontostriatal dysfunction in attention-deficit/hyperactivity disorder versus "hot" ventromedial orbitofrontal-limbic dysfunction in conduct disorder: a review. Biol Psychiatry 2011;69:e69-87.
263. Hart H, Radua J, Nakao T et al. Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: exploring task-specific, stimulant medication, and age effects. JAMA Psychiatry 2013;70:185-98.
264. Morein-Zamir S, Robbins TW. Fronto-striatal circuits in response-inhibition: relevance to addiction. Brain Res 2015;1628:117-29.
265. Hoogman M, Muetzel R, Guimaraes JP et al. Brain imaging of the cortex in ADHD: a coordinated analysis of large-scale clinical and population-based samples. Am J Psychiatry 2019;176:531-42.
266. Norman LJ, Carlisi C, Lukito S et al. Structural and functional brain abnormalities in attention-deficit/hyperactivity disorder and obsessive-compulsive disorder: a comparative meta-analysis. JAMA Psychiatry 2016;73:815-25.
267. Alegria AA, Radua J, Rubia K. Meta-analysis of fMRI studies of disruptive behavior disorders. Am J Psychiatry 2016;173:1119-30.
268. Luijten M, Machielsen MWJ, Veltman DJ et al. Systematic review of ERP and fMRI studies investigating inhibitory control and error processing in people with substance dependence and behavioural addictions. J Psychiatry Neurosci 2014;39:149-69.
269. Xiao P, Dai Z, Zhong J et al. Regional gray matter deficits in alcohol dependence: a meta-analysis of voxel-based morphometry studies. Drug Alcohol Depend 2015;153:22-8.
270. Sato JR, Salum GA, Gadelha A et al. Default mode network maturation and psychopathology in children and adolescents. J Child Psychol Psychiatry 2016;57:55-64.
271. Whittle S, Vijayakumar N, Simmons JG et al. Internalizing and externalizing symptoms are associated with different trajectories of cortical development during late childhood. J Am Acad Child Adolesc Psychiatry 2020;59:177-85.
272. Tervo-Clemmens B, Quach A, Calabro FJ et al. Meta-analysis and review of functional neuroimaging differences underlying adolescent vulnerability to substance use. NeuroImage 2020;209:116476.
273. Volkow ND, Wang G-J, Fowler JS et al. Addiction circuitry in the human brain. Annu Rev Pharmacol Toxicol 2012;52:321-36.
274. Willuhn I, Wanat MJ, Clark JJ et al. Dopamine signaling in the nucleus accumbens of animals self-administering drugs of abuse. Curr Top Behav Neurosci 2010;3:29-71.
275. Ellison-Wright I, Ellison-Wright Z, Bullmore E. Structural brain change in attention deficit hyperactivity disorder identified by meta-analysis. BMC Psychiatry 2008;8:51.
276. 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.
277. Bickel WK, Koffarnus MN, Moody L et al. The behavioral- and neuro-economic process of temporal discounting: a candidate behavioral marker of addiction. Neuropharmacology 2014;76:518-27.
278. Buckholtz JW, Treadway MT, Cowan RL et al. Mesolimbic dopamine reward system hypersensitivity in individuals with psychopathic traits. Nat Neurosci 2010;13:419-21.
279. Plichta MM, Scheres A. Ventral-striatal responsiveness during reward anticipation in ADHD and its relation to trait impulsivity in the healthy population: a meta-analytic review of the fMRI literature. Neurosci Biobehav Rev 2014;38:125-34.
280. Shanmugan S, Wolf DH, Calkins ME et al. Common and dissociable mechanisms of executive system dysfunction across psychiatric disorders in youth. Am J Psychiatry 2016;173:517-26.
281. Xia CH, Ma Z, Ciric R et al. Linked dimensions of psychopathology and connectivity in functional brain networks. Nat Commun 2018;9:3003.
282. Abram SV, Wisner KM, Grazioplene RG et al. Functional coherence of insula networks is associated with externalizing behavior. J Abnorm Psychol 2015;124:1079-91.
283. Sadeh N, Spielberg JM, Logue MW et al. Linking genes, circuits, and behavior: network connectivity as a novel endophenotype of externalizing. Psychol Med 2019;49:1905-13.
284. Cecil CAM, Walton E, Viding E. DNA methylation, substance use and addiction: a systematic review of recent animal and human research from a developmental perspective. Curr Addict Rep 2015;2:331-46.
285. Prom-Wormley EC, Ebejer J, Dick DM et al. The genetic epidemiology of substance use disorder: a review. Drug Alcohol Depend 2017;180:241-59.
286. Provençal N, Suderman MJ, Guillemin C et al. Association of childhood chronic physical aggression with a DNA methylation signature in adult human T cells. PLoS One 2014;9:e89839.
287. Provençal N, Booij L, Tremblay RE. The developmental origins of chronic physical aggression: biological pathways triggered by early life adversity. J Exp Biol 2015;218:123-33.
288. Waltes R, Chiocchetti AG, Freitag CM. The neurobiological basis of human aggression: a review on genetic and epigenetic mechanisms. Am J Med Genet B Neuropsychiatr Genet 2016;171:650-75.
289. Barker ED, Walton E, Cecil CAM et al. A methylome-wide association study of trajectories of oppositional defiant behaviors and biological overlap with attention deficit hyperactivity disorder. Child Dev 2018;89:1839-55.
290. Knopik VS, Marceau K, Bidwell LC et al. Prenatal substance exposure and offspring development: does DNA methylation play a role? Neurotoxicol Teratol 2019;71:50-63.

291. Barker ED, Walton E, Cecil CAM. Annual research review: DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology. J Child Psychol Psychiatry 2018;59:303-22.
292. Cecil CAM, Walton E, Smith RG et al. DNA methylation and substance-use risk: a prospective, genome-wide study spanning gestation to adolescence. Transl Psychiatry 2016;6:e976.
293. Barker ED, Cecil CAM, Walton E et al. Inflammation-related epigenetic risk and child and adolescent mental health: a prospective study from pregnancy to middle adolescence. Dev Psychopathol 2018;30:1145-56.
294. Slopen N, Kubzansky LD, Koenen KC. Internalizing and externalizing behaviors predict elevated inflammatory markers in childhood. Psychoneuroendocrinology 2013;38:2854-62.
295. Coccaro EF, Lee R, Coussons-Read M. Elevated plasma inflammatory markers in individuals with intermittent explosive disorder and correlation with aggression in humans. JAMA Psychiatry 2014;71:158-65.
296. Marsland AL, Prather AA, Petersen KL et al. Antagonistic characteristics are positively associated with inflammatory markers independently of trait negative emotionality. Brain Behav Immun 2008;22:753-61.
297. Mitchell RHB, Goldstein BI. Inflammation in children and adolescents with neuropsychiatric disorders: a systematic review. J Am Acad Child Adolesc Psychiatry 2014;53:274-96.
298. Anand D, Colpo GD, Zeni G et al. Attention-deficit/hyperactivity disorder and inflammation: what does current knowledge tell us? A systematic review. Front Psychiatry 2017;8:228.
299. Leffa DT, Torres ILS, Rohde LA. A review on the role of inflammation in attention-deficit/hyperactivity disorder. Neuroimmunomodulation 2018; 25:328-33.
300. Wang T-Y, Lee S-Y, Hu M-C et al. More inflammation but less brain-derived neurotrophic factor in antisocial personality disorder. Psychoneuroendocrinology 2017;85:42-8.
301. Ashare RL, Wetherill RR. The intersection of sex differences, tobacco use, and inflammation: implications for psychiatric disorders. Curr Psychiatry Rep 2018;20:75.
302. Nennig SE, Schank JR. The role of NFkB in drug addiction: beyond inflammation. Alcohol Alcohol 2017;52:172-9.
303. Harricharan R, Abboussi O, Daniels WMU. Addiction: a dysregulation of satiety and inflammatory processes. In: Calvey T, Daniels WMU (eds). Progress in brain research. Amsterdam: Elsevier, 2017:65-91.
304. Scassellati C, Bonvicini C, Faraone SV et al. Biomarkers and attention-deficit/hyperactivity disorder: a systematic review and meta-analyses. J Am Acad Child Adolesc Psychiatry 2012;51:1003-1019.e20.
305. Haltigan JD, Roisman GI, Susman EJ et al. Elevated trajectories of externalizing problems are associated with lower awakening cortisol levels in mid adolescence. Dev Psychol 2011;47:472-8.
306. Loney BR, Butler MA, Lima EN et al. The relation between salivary cortisol, callous-unemotional traits, and conduct problems in an adolescent nonreferred sample. J Child Psychol Psychiatry 2006;47:30-6.
307. Lovallo WR. Cortisol secretion patterns in addiction and addiction risk. Int J Psychophysiol 2006;59:195-202.
308. Marceau K, Abel E. Mechanisms of cortisol - substance use development associations: hypothesis generation through gene enrichment analysis. Neurosci Biobehav Rev 2018;92:128-39.
309. Zorn JV, Schür RR, Boks MP et al. Cortisol stress reactivity across psychiatric disorders: a systematic review and meta-analysis. Psychoneuroendocrinology 2017;77:25-36.
310. Harro J, Oreland L. The role of MAO in personality and drug use. Prog Neuropsychopharmacol Biol Psychiatry 2016;69:101-11.
311. Kiive E, Fischer K, Harro M et al. Platelet monoamine oxidase activity in association with adolescent inattentive and hyperactive behaviour: a prospective longitudinal study. Personal Individ Differ 2007;43:155-66.
312. Moore TM, Scarpa A, Raine A. A meta-analysis of serotonin metabolite 5-HIAA and antisocial behavior. Aggress Behav 2002;28:299-316.
313. Seo D, Patrick CJ, Kennealy PJ. Role of serotonin and dopamine system interactions in the neurobiology of impulsive aggression and its comorbidity with other clinical disorders. Aggress Violent Behav 2008;13:383-95.
314. Duke AA, Bègue L, Bell R et al. Revisiting the serotonin-aggression relation in humans: a meta-analysis. Psychol Bull 2013;139:1148-72.
315. Mervielde I, De Clercq B, De Fruyt F et al. Temperament, personality, and developmental psychopathology as childhood antecedents of personality disorders. J Personal Disord 2005;19:171-201.
316. Tackett JL, Balsis S, Oltmanns TF et al. A unifying perspective on personality pathology across the life span: developmental considerations for the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders. Dev Psychopathol 2009;21:687-713.
317. Widiger TA, De Clercq B, De Fruyt F. Childhood antecedents of personality disorder: an alternative perspective. Dev Psychopathol 2009;21:771-91.
318. Clark LA. Temperament as a unifying basis for personality and psychopathology. J Abnorm Psychol 2005;114:505-21.
319. Shiner RL. A developmental perspective on personality disorders: lessons from research on normal personality development in childhood and adolescence. J Personal Disord 2005;19:202-10.
320. Martel MM, Smith TE, Lee CA. Personality development and externalizing psychopathology. In: McAdams DP, Shiner Rl, Tackett JL (eds). Handbook of personality development. New York: Guilford, 2019:534-50.
321. Morizot J. The contribution of temperament and personality traits to criminal and antisocial behavior development and desistance. In: Morizot J, Kazemian L (eds). The development of criminal and antisocial behavior: theory, research and practical applications. Cham: Springer, 2015:137-65.
322. Lahey BB, Applegate B, Chronis AM et al. Psychometric characteristics of a measure of emotional dispositions developed to test a developmental propensity model of conduct disorder. J Clin Child Adolesc Psychol 2008;37:794-807.
323. Caspi A, Henry B, McGee RO et al. Temperamental origins of child and adolescent behavior problems: from age three to age fifteen. Child Dev 1995;66: 55-68.
324. Caspi A. Behavioral observations at age 3 years predict adult psychiatric disorders: longitudinal evidence from a birth cohort. Arch Gen Psychiatry 1996;53:1033-9.
325. Clercq BD, Fruyt FD. A five-factor model framework for understanding childhood personality disorder antecedents. J Pers 2012;80:1533-63.
326. Nielsen IKM. The impact of temperamental dimensions on change in symptoms of oppositional defiant disorder from preschool to first grade. Master Thesis, Norwegian University of Science and Technology, 2014.
327. Stringaris A, Maughan B, Goodman R. What's in a disruptive disorder? Temperamental antecedents of oppositional defiant disorder: findings from the Avon Longitudinal Study. J Am Acad Child Adolesc Psychiatry 2010;49:474-83.
328. Tremblay RE, Galera C, Orri M et al. Developmental trajectories of aggression and other problematic behaviors associated with IED. In: Coccaro EF, McCloskey MS (eds). Intermittent explosive disorder. Cambridge: Academic Press, 2019:3-15.
329. Forbes MK, Rapee RM, Camberis A-L et al. Unique associations between childhood temperament characteristics and subsequent psychopathology symptom trajectories from childhood to early adolescence. J Abnorm Child Psychol 2017;45:1221-33.
330. Waller R, Wagner NJ, Barstead MG et al. A meta-analysis of the associations between callous-unemotional traits and empathy, prosociality, and guilt. Clin Psychol Rev 2020;75:101809.
331. Chanen AM, Kaess M. Developmental pathways to borderline personality disorder. Curr Psychiatry Rep 2012;14:45-53.
332. Stepp SD, Whalen DJ, Pedersen SL. The externalizing pathway to borderline personality disorder in youth. In: Sharp C, Tackett JL (eds). Handbook of borderline personality disorder in children and adolescents. New York: Springer, 2014:247-63.
333. Ivanov I, Schulz KP, London ED et al. Inhibitory control deficits in childhood and risk for substance use disorders: a review. Am J Drug Alcohol Abuse 2008;34:239-58.
334. Rioux C, Castellanos-Ryan N, Parent S et al. The interaction between temperament and the family environment in adolescent substance use and externalizing behaviors: support for diathesis-stress or differential susceptibility? Dev Rev 2016;40:117-50.
335. Creemers HE, Verhulst FC, Huizink AC. Temperamental risk factors for adolescent cannabis use: a systematic review of prospective general population studies. Subst Use Misuse 2009;44:1833-54.
336. Moffitt TE. Life-course-persistent and adolescence-limited antisocial behavior: a developmental taxonomy. Psychol Rev 1993;100:674-701.
337. Loeber R, Hay D. Key issues in the development of aggression and violence from childhood to early adulthood. Annu Rev Psychol 1997;48:371-410.
338. Kuperman S, Schlosser SS, Kramer JR et al. Risk domains associated with an adolescent alcohol dependence diagnosis. Addiction 2001;96:629-36.
339. Lynam DR, Charnigo R, Moffitt TE et al. The stability of psychopathy across adolescence. Dev Psychopathol 2009;21:1133-53.
340. Wichers M, Gardner C, Maes HH et al. Genetic innovation and stability in externalizing problem behavior across development: a multi-informant twin study. Behav Genet 2013;43:191-201.
341. Olson SL, Bates JE, Sandy JM et al. Early developmental precursors of ex-

ternalizing behavior in middle childhood and adolescence. J Abnorm Child Psychol 2000;28:119-33.
342. Helmond P, Overbeek G, Brugman D et al. A meta-analysis on cognitive distortions and externalizing problem behavior: associations, moderators, and treatment effectiveness. Crim Justice Behav 2015;42:245-62.
343. Maughan DR, Christiansen E, Jenson WR et al. Behavioral parent training as a treatment for externalizing behaviors and disruptive behavior disorders: a meta-analysis. Sch Psychol Rev 2005;34:267-86.
344. Reyno SM, McGrath PJ. Predictors of parent training efficacy for child externalizing behavior problems – a meta-analytic review. J Child Psychol Psychiatry 2006;47:99-111.
345. Barnes TN, Smith SW, Miller MD. School-based cognitive-behavioral interventions in the treatment of aggression in the United States: a meta-analysis. Aggress Violent Behav 2014;19:311-21.
346. Kremer KP, Maynard BR, Polanin JR et al. Effects of after-school programs with at-risk youth on attendance and externalizing behaviors: a systematic review and meta-analysis. J Youth Adolesc 2015;44:616-36.
347. Mingebach T, Kamp-Becker I, Christiansen H et al. Meta-meta-analysis on the effectiveness of parent-based interventions for the treatment of child externalizing behavior problems. PLoS One 2018;13:e0202855.
348. Faraone SV, Biederman J, Spencer TJ et al. Comparing the efficacy of medications for ADHD using meta-analysis. Medscape Gen Med 2006;8:4.
349. Faraone SV, Buitelaar J. Comparing the efficacy of stimulants for ADHD in children and adolescents using meta-analysis. Eur Child Adolesc Psychiatry 2010;19:353-64.
350. Faraone SV, Glatt SJ. A comparison of the efficacy of medications for adult attention-deficit/hyperactivity disorder using meta-analysis of effect sizes. J Clin Psychiatry 2010;71:754-63.
351. Ipser J, Stein DJ. Systematic review of pharmacotherapy of disruptive behavior disorders in children and adolescents. Psychopharmacology 2007;191: 127-40.
352. Lundahl BW, Kunz C, Brownell C et al. A meta-analysis of motivational interviewing: twenty-five years of empirical studies. Res Soc Work Pract 2010;20:137-60.
353. Comer JS, Chow C, Chan PT et al. Psychosocial treatment efficacy for disruptive behavior problems in very young children: a meta-analytic examination. J Am Acad Child Adolesc Psychiatry 2013;52:26-36.
354. Epstein RA, Fonnesbeck C, Potter S et al. Psychosocial interventions for child disruptive behaviors: a meta-analysis. Pediatrics 2015;136:947-60.
355. Battagliese G, Caccetta M, Luppino OI et al. Cognitive-behavioral therapy for externalizing disorders: a meta-analysis of treatment effectiveness. Behav Res Ther 2015;75:60-71.
356. 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.
357. Cottler L. The DSM-IV field trial for substance use disorders: major results. Drug Alcohol Depend 1995;38:59-69.
358. Narrow WE, Clarke DE, Kuramoto SJ et al. DSM-5 field trials in the United States and Canada, Part III: Development and reliability testing of a cross-cutting symptom assessment for DSM-5. Am J Psychiatry 2013;170:71-82.
359. 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.
360. Quilty LC, Ayearst L, Chmielewski M et al. The psychometric properties of the personality inventory for DSM-5 in an APA DSM-5 field trial sample. Assessment 2013;20:362-9.
361. Wright AGC, 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 2014;124:199-207.
362. Chmielewski M, Ruggero CJ, Kotov R et al. Comparing the dependability and associations with functioning of the DSM–5 Section III trait model of personality pathology and the DSM-5 Section II personality disorder model. Personal Disord Theory Res Treat 2017;8:228-36.
363. Sunderland M, Slade T, Krueger RF. Examining the shared and unique relationships among substance use and mental disorders. Psychol Med 2015;45:1103-13.
364. Ruggero CJ, Kotov R, Hopwood CJ et al. Integrating the Hierarchical Taxonomy of Psychopathology (HiTOP) into clinical practice. J Consult Clin Psychol 2019;87:1069-84.
365. Bornstein RF, Natoli AP. Clinical utility of categorical and dimensional perspectives on personality pathology: a meta-analytic review. Personal Disord 2019;10:479-90.
366. Tellegen A, Waller N. Exploring personality through test construction: development of the Multidimensional Personality Questionnaire. In: Boyle GJ, Matthews G, Saklofske DH (eds). The SAGE handbook of personality theory and assessment. Thousand Oaks: Sage, 2008:261-92.
367. Blonigen DM, Patrick CJ, Gasperi M et al. Delineating the construct network of the Personnel Reaction Blank: associations with externalizing tendencies and normal personality. Psychol Assess 2010;23:18-30.
368. Venables NC, Patrick CJ. Validity of the Externalizing Spectrum Inventory in a criminal offender sample: relations with disinhibitory psychopathology, personality, and psychopathic features. Psychol Assess 2011;24:88-100.
369. Soe-Agnie SE, Patrick CJ, Nijman HLI et al. Validation of the full and brief Externalizing Spectrum Inventory in Dutch forensic inpatients. J Forensic Psychiatry Psychol 2016;27:77-91.
370. Sunderland M, Slade T, Krueger RF et al. Efficiently measuring dimensions of the externalizing spectrum model: development of the Externalizing Spectrum Inventory-Computerized Adaptive Test (ESI-CAT). Psychol Assess 2016;29:868-80.
371. Ben-Porath YS, Tellegen A. Minnesota Multiphasic Personality Inventory-2-Restructured Form. www.pearsonassessments.com.
372. Moray LC. Professional manual for the Personality Assessment Inventory, 2nd ed. Lutz: Psychological Assessment Resources, 2007.
373. Lee TTC, Sellbom M, Hopwood CJ. Contemporary psychopathology assessment: mapping major personality inventories onto empirical models of psychopathology. In: Bowden SC (ed). Neuropsychological assessment in the age of evidence-based practice: diagnostic and treatment evaluations. New York: Oxford University Press, 2017:65-94.
374. Sellbom M. Elucidating the validity of the externalizing spectrum of psychopathology in correctional, forensic, and community samples. J Abnorm Psychol 2016;125:1027-38.
375. Sellbom M. Mapping the MMPI-2-RF specific problems scales onto extant psychopathology structures. J Pers Assess 2017;99:341-50.
376. Sellbom M. The MMPI-2-Restructured Form (MMPI-2-RF): assessment of personality and psychopathology in the twenty-first century. Annu Rev Clin Psychol 2019;15:149-77.
377. McCrae RR, Costa PT Jr, Martin TA. The NEO-PI-3: a more readable revised NEO personality inventory. J Pers Assess 2005;84:261-70.
378. Simms LJ, Goldberg LR, Roberts JE et al. Computerized Adaptive Assessment of Personality Disorder: introducing the CAT-PD project. J Pers Assess 2011;93:380-9.
379. Wright AGC, Thomas KM, Hopwood CJ et al. The hierarchical structure of DSM-5 pathological personality traits. J Abnorm Psychol 2012;121:951-7.
380. Achenbach TM, Ruffle TM. The Child Behavior Checklist and related forms for assessing behavioral/emotional problems and competencies. Pediatr Rev 2000;21:265-71.
381. Goodman R. Psychometric properties of the Strengths and Difficulties Questionnaire. J Am Acad Child Adolesc Psychiatry 2001;40:1337-45.
382. Shaffer D, Fisher P, Lucas C et al. NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. J Am Acad Child Adolesc Psychiatry 2000;39:28-38.
383. Carragher N, Teesson M, Sunderland M et al. The structure of adolescent psychopathology: a symptom-level analysis. Psychol Med 2016;46:981-94.
384. Lahey BB, Zald DH, Perkins SF et al. Measuring the hierarchical general factor model of psychopathology in young adults. Int J Methods Psychiatr Res 2018;27:e1593.
385. McElroy E, Belsky J, Carragher N et al. Developmental stability of general and specific factors of psychopathology from early childhood to adolescence: dynamic mutualism or p-differentiation? J Child Psychol Psychiatry 2018;59:667-75.
386. Conway KP, Levy J, Vanyukov M et al. Measuring addiction propensity and severity: the need for a new instrument. Drug Alcohol Depend 2010;111:4-12.
387. McKenna H, Treanor C, O'Reilly D et al. Evaluation of the psychometric properties of self-reported measures of alcohol consumption: a COSMIN systematic review. Subst Abuse Treat Prev Policy 2018;13:6.
388. Suris A, Lind L, Emmett G et al. Measures of aggressive behavior: overview of clinical and research instruments. Aggress Violent Behav 2004;9:165-227.

DOI:10.1002/wps.20844

Racism and mental health

The COVID-19 pandemic, with its striking inequities in mortality rates between Whites and stigmatized racial/ethnic groups in the US and UK, and the recent global protests about police violence have raised questions about and increased interest in the potential impacts of racism on health and particularly on mental health.

Racism is an organized societal system in which the dominant White group categorizes individuals into "races" and devalues, disempowers, and differentially allocates resources to ethnic groups considered to be inferior[1]. The ideology of inferiority permeates societal systems and institutions ensuring that racism is not limited to individual beliefs and behaviors[1]. Racism operates through institutional, interpersonal and cultural pathways. Here we provide a brief overview of these levels of racism and how they can adversely affect mental health.

Institutional or structural racism can be defined as racial discrimination that is embedded in institutional structures and policies[1]. Examples of institutional racism include residential segregation, racialized immigration policy, and racialized incarceration. For example, in the US, residential segregation, the physical separation of races by enforced residence in particular places, is a central determinant of racial inequities in health. Residing in areas with concentrated poverty and social disadvantage can adversely affect mental health by leading to high levels of exposure to stressors (psychosocial, physical and chemical) and reduced access to opportunities and resources, including schooling, employment, and health services. Empirical analyses reveal that eliminating residential segregation in the US would erase racial differences in income, education and unemployment, and reduce racial differences in single motherhood by two thirds[1].

Immigration policies often reproduce ideologies of belonging and othering that are patterned by race/ethnicity and can adversely affect racialized immigrant groups. Research documents that exclusionary immigration policies with restrictions on rights and aggressive anti-immigrant policy enforcement have negative effects on mental health[2].

Racialized incarceration also has mental health consequences. The US have the largest incarcerated population globally, with an overrepresentation of Blacks and Latinos[3]. This has facilitated a historic shift from mental illness being treated in hospitals to being treated in carceral systems, which has led to jails and prisons in the US becoming the largest providers of mental health care. A national study in the US found that prior arrest history was associated with the prevalence of major depressive disorder among African Americans and Caribbean Blacks[4]. In addition, other US research reveals that aggressive policing, such as the killing of unarmed African Americans, leads to declining mental health for the entire Black population in the state in which the incident occurs[1].

Self-reported interpersonal discrimination is the most studied domain of racism in the mental health literature. A review of literature reviews and meta-analyses published between 2013 and 2019 on discrimination and health identified eight reviews focused on mental health[5]. Although most studies came from the US, there were studies from some 20 countries. This body of research indicates that discrimination was positively associated with increased risk of major mental disorders and inversely related to positive mental health outcomes such as life satisfaction and self-esteem. The accumulation of experiences of discrimination over time was associated with increasing risk of mental health problems. Exposure to discrimination was also associated with adverse changes in personality over time, such as increases in neuroticism.

Although the majority of studies have been cross-sectional, a growing number of prospective studies link discrimination to mental health risk. Some studies have also documented that the association between discrimination and mental health is robust to adjustment for potential psychological confounders such as neuroticism[6]. In addition, racial discrimination is also linked to worse mental health and increase in risky behavior for children and adolescents[7]. In addition to direct exposure to racial discrimination, vicarious exposure, through parental or caregiver experiences of discrimination, is also associated with worse mental health outcomes[8].

Cultural racism refers to the racist ideologies that are present in the media, stereotypes, and norms of society that undergird institutional and interpersonal racism[1]. It can affect mental health in multiple ways. First, cultural racism can initiate and sustain policies that create conditions which are harmful to mental health, such as housing decisions to maintain residential segregation which facilitates increased exposure to traumatic experiences and a broad range of physical and social stressors. Second, some members of stigmatized racial groups internalize the negative racial stereotypes of the culture, which in turn can lead to increased psychological distress and substance use. Third, cultural racism can also trigger stereotype threat – expectations and anxieties activated by a stigmatized group when negative stereotypes about their group are made salient. Research reveals that stereotype threat can lead to increased anxiety, reduced self-regulation, and impaired decision making, which can also affect patient-provider communication and adherence to medical advice[1].

Furthermore, cultural racism can lead to individual-level unconscious bias in clinicians that can trigger discrimination adversely affecting the quality of clinical care. For example, research has documented racial differences in the application of psychiatric diagnostic criteria, so that Latinos are diagnosed with anxiety disorders more frequently than White people reporting the same symptoms[4]. Similarly, clinicians exposed to the same symptoms are more likely to diagnose African Americans with psychotic disorders than mood disorders compared to Whites[4].

Future research is needed to better understand the intersection of racial discrimination with other forms of group discrimination (e.g., gender-related) and identify how multiple forms of discrimination may impact mental health. Emerging evidence

indicates that multiple forms of discrimination, such as racism and heterosexism, are associated with increased risk of mental health problems[9]. Additionally, our current understanding is limited about the potential intergenerational impacts of racism and their related epigenetic effects, with emerging evidence suggesting that these processes are likely to be operative[1].

Research on racism and mental health, to date, has focused more on documenting that racism matters than on identifying interventions to minimize the adverse effects of exposure to racism and reduce the occurrence of racism in the first place. Some evidence suggests that psychosocial resources such as social ties and religious involvement can reduce some of the negative effects of discrimination on mental health[6]. However, effectively addressing the multifactorial impacts of racism on mental health will require multilevel societal interventions that seek to build racial equity into homes, schools, neighborhoods and workplaces to minimize current racial economic gaps and improve socioeconomic and living conditions for the disadvantaged.

Interventions around resiliency and cultural/structural competency in the medical field have shown some promise, but more concerted attention is needed to address the multiple and interconnected systems through which racism operates[1,3]. Diversifying the mental health workforce in terms of including underrepresented racial/ethnic groups and professional experience (e.g., medicine, social work, religion) is also a necessary step towards addressing inequities in mental health care[3]. Comprehensive, coordinated, strategic initiatives are needed both within and outside of psychiatry and medicine to better understand, prevent and effectively intervene on the effects of racism on mental health.

David R. Williams[1,2], Onisha S. Etkins[1]
[1]Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA, USA; [2]Department of African and African American Studies, Harvard University, Cambridge, MA, USA

1. Williams DR, Lawrence JA, Davis BA. Annu Rev Publ Health 2019;40:105-25.
2. Hatzenbuehler ML, Prins SJ, Flake M et al. Soc Sci Med 2017;174:169-78.
3. Medlock MM, Shtasel D, Trinh N-HT et al (eds). Racism and psychiatry: contemporary issues and interventions. Berlin: Springer, 2018.
4. Anglin DM, Lighty Q, Yang LH et al. Psychiatry Res 2014;219:114-21.
5. Williams DR, Lawrence JA, Davis BA et al. Health Serv Res 2019;54:1374-88.
6. Lewis TT, Cogburn CD, Williams DR. Annu Rev Clin Psychol 2015;11:407-40.
7. Priest N, Paradies Y, Trenerry B et al. Soc Sci Med 2013;95:115-27.
8. Heard-Garris NJ, Cale M, Camaj L et al. Soc Sci Med 2018;199:230-40.
9. Vargas SM, Huey SJ, Miranda J. Am J Orthopsychiatry 2020;90:374-90.

DOI:10.1002/wps.20845

The epidemic of fentanyl misuse and overdoses: challenges and strategies

Fentanyl, a synthetic opioid with analgesic and anesthetic properties, is currently associated with one of the deadliest addiction crises in the US. Misuse of fentanyl (and fentanyl analogues) has been estimated to be responsible for 48,000 (out of a total of 83,335) overdose deaths in the 12 months ending in June 2020[1], a rate that has increased more than 29 fold since 2012, when the annual fatalities from fentanyl and its analogues were 1,615.

The cases of overdoses and deaths in the US are linked to illegally manufactured fentanyl, which rapidly penetrated the US illicit market since 2013. Though not as pervasively as in the US, increases in overdose deaths due to illicit fentanyl and its analogues have also occurred in Canada, in several European countries (including Estonia, Germany, Finland and the UK) and in Australia[2].

Fentanyl is relatively easy to synthesize and manufacture, and less difficult to traffic than heroin, since it requires much smaller volumes to transport across borders. It is, therefore, hugely profitable to drug dealers (50-100 times more than heroin), which can be expected to result in an expansion of the illicit fentanyl market across the globe.

The majority of opioid-related overdose deaths in the US are the result of fentanyl being ingested as a substitute for heroin or with drugs such as cocaine and methamphetamine that had been adulterated (cut) with the opioid, frequently without users being aware of this. Fentanyl, when used by itself or in combination with other drugs, can be taken orally, injected, snorted or smoked. Most heroin users do not report actively seeking fentanyl, and some are afraid of it but might have no choice because of the higher costs of uncontaminated heroin or its unavailability.

When fentanyl is used to adulterate other drugs (heroin, prescription opioids, psychostimulants), it increases their lethality. In the case of psychostimulants, this occurs not only due to the synergistic effects on the cardiopulmonary system, but also because stimulant users, who have no tolerance to opioids, are at very high risk of overdosing when ingesting fentanyl.

The unique pharmacological effects of fentanyl have contributed to its widespread misuse and are also the ones that make it a valuable therapeutic for anesthesia and for severe pain management. Fentanyl binds to mu-opioid receptors (MOR), which mediate the analgesic and the rewarding effects of opioid drugs, such as morphine and heroin, as well as their respiratory depressing actions[3]. However, fentanyl is much more potent at activating MOR-associated signaling than morphine (80-100 fold) or heroin (30-50 fold), and its higher lipophilicity leads to higher and faster brain uptake than for those other drugs. These properties underlie fentanyl's high potency as an analgesic and its rapid actions, which are beneficial for the treatment of breakthrough pain or other severe pain conditions. However, they are also responsible for its powerful rewarding effects, which can rapidly

result in physical dependence and in addiction, and for its severe and abrupt inhibition of respiration, which increases the risk for overdose.

The treatment of fentanyl addiction (fentanyl opioid use disorder or fOUD) is the same as for other opioid use disorders (OUD). It is based on the use of medications such as methadone (full MOR agonist), buprenorphine (partial MOR agonist) and naltrexone (MOR antagonist)[4]. These medications are the gold standard for OUD treatment, and multiple studies have shown that they prevent overdoses and relapse in patients exposed to fentanyl.

However, clinical cases and anecdotal reports indicate that it is much more challenging to treat patients with fOUD than with other OUD. There are greater difficulties in initiating buprenorphine treatment, resulting from buprenorphine-precipitated withdrawal[5] and lower rates of abstinence and retention after six months of buprenorphine treatment[6]. The slow clearance of fentanyl as a result of its accumulation in fatty tissues may require slower detoxification prior to buprenorphine or naltrexone induction, and the higher rates of tolerance and physical dependence associated with repeated fentanyl use might necessitate higher doses of methadone or buprenorphine than for other OUDs. Treatment of withdrawal symptoms during fentanyl detoxification might be aided, as for other opioids, by the use of the alpha-adrenergic drugs lofexidine and clonidine[7]. Overall, much more clinical research is needed to investigate how to optimally treat fOUD.

Like other opioids, fentanyl can result in overdoses due to its respiratory depressant effects. Signs of overdose include slow irregular breathing, slowing of circulation, sedation, acute respiratory distress, seizures, and coma. With repeated opioid exposure, individuals develop tolerance to the respiratory depressant effects of opioids (tolerance also develops for analgesia and reward), allowing them to tolerate much higher doses than naïve individuals[8]. Because tolerance to opioids decreases with interruption of use, whether during voluntary detoxification or incarceration, the relapse to opioid use after treatment discontinuation or after release from jail/prison is particularly dangerous.

Even for those who have developed tolerance to opioids, the very high potency of fentanyl, the impossibility of precisely dosing and the frequency with which drugs are mixed in the black market contribute to the high overdose risk associated with its misuse. As for other opioids, the treatment of fentanyl overdoses requires the timely delivery of naloxone (MOR antagonist) either via parenteral or intranasal administration[3]. Naloxone, which also has a very high affinity for MOR, displaces fentanyl from the receptor, thereby restoring breathing (as well as precipitating an acute opioid withdrawal).

Clinical cases and case reports have indicated that overdoses from fentanyl frequently require multiple naloxone administrations, due to the shorter duration of the action of naloxone ($t_{1/2}$: 1.3-2.4 hours) than that of fentanyl ($t_{1/2}$: 7 to 8 hours), prolonged further by the slow clearance rates of fentanyl in frequent users. Additionally, when fentanyl is injected rapidly, it can result in chest wall rigidity, which interferes further with breathing and exacerbates the risk of death; these effects are not MOR-mediated and might reflect noradrenergic and cholinergic mechanisms[9].

All this generates the need for further development of fentanyl overdose treatments, including higher-dose naloxone formulations, autoinjectors that automatically release naloxone with an impending overdose, longer-acting opioid antagonists (i.e., nalmefene), treatments against chest wall rigidity, and medications to stimulate respiration and oxygenation to help overdoses from the combination of opioids with alcohol, benzodiazepines or stimulants.

Modeling studies have revealed that the epidemic of opioid overdose deaths, including those from fentanyl, can be reversed by multi-pronged approaches that expand access to medications to treat opioid use disorders and increase retention in medication treatment, and by widely expanding access to naloxone for overdose reversals. It will also require strengthening the education of health care professionals in pain management, in safe use of opioids, and in how to screen and treat substance use disorders (including OUD).

Allocation of resources to implement these interventions is necessary, and timely surveillance systems that can serve as early warning signals for the presence of fentanyl or other opioids in a community would also be beneficial. In parallel, prevention interventions are needed to protect against opioid misuse initiation, recognizing that socioeconomic factors have contributed to the opioid crisis and that addressing them is necessary for preventing OUD and other substance use disorders in the long term.

Nora D. Volkow
National Institute on Drug Abuse, National Institutes of Health, Bethesda, MD, USA

1. US Centers for Disease Control and Prevention. Increase in fatal drug overdoses across the United States driven by synthetic opioids before and during the COVID-19 pandemic. CDC Health Alert Network, December 17, 2020.
2. United Nations. World drug report 2020. https://wdr.unodc.org/wdr2020/.
3. Comer SD, Cahill CM. Neurosci Biobehav Rev 2019;106:49-57.
4. National Academies of Sciences, Engineering, and Medicine. Medications for opioid use disorder save lives. Washington: National Academies Press, 2019.
5. Silverstein SM, Daniulaityte R, Martins SS et al. Int J Drug Policy 2019;74:76-83.
6. Wakeman SE, Chang Y, Regan S et al. J Addict Med 2019;13:253-7.
7. Srivastava AB, Mariani JJ, Levin FR. Lancet 2020;395:1938-48.
8. Volkow ND, McLellan AT. N Engl J Med 2016;374:1253-63.
9. Torralva R, Janowsky A. J Pharmacol Exp Ther 2019;371:453-75.

DOI:10.1002/wps.20846

The need for publicly funded research on therapeutic use of psychedelic drugs

A psychedelic drug is one that "produces thought, mood and perceptual changes otherwise rarely experienced except in dreams, contemplative and religious exaltation, flashes of vivid involuntary memory, and acute psychosis"[1]. It does so "without causing physical addiction, craving, major physiological disturbances, delirium, disorientation or amnesia"[1].

The "classic psychedelics" include mescaline, psilocybin, lysergic acid diethylamide (LSD), dimethyltryptamine (DMT), and plant-based substances such as ibogaine and ayahuasca. Their chemical structures differ, but they all act on the 5-HT2A serotonin receptor[1]. 3,4-methylenedioxymethamphetamine (MDMA) is also included, although it does not produce the perceptual effects of the classic psychedelics[2].

Over the past two decades, there has been a revival of clinical research on the therapeutic use of psilocybin and MDMA[2,3]. This research has been encouraged by the US Food and Drug Administration (FDA) because in phase 2 trials these drugs have produced substantial benefits, respectively, in patients with treatment-resistant depression and post-traumatic stress disorder (PTSD)[3]. Funding for psychedelic research has largely been philanthropic, because the pharmaceutical industry is not interested in drugs that are off patent.

The new psychedelic research that is being done in leading universities in the US and Europe includes randomized controlled trials conducted to the standard required for FDA approval[3]. Psilocybin has been chosen rather than LSD, because it has a shorter period of action (4-6 hours vs. 8-12 hours), its pharmacology is better understood, it is less likely to produce "bad trips", and it does not carry the cultural baggage of LSD[3]. Clinical trials have also been done on MDMA-assisted psychotherapy in PTSD.

If phase 3 trials confirm the results of phase 1 and 2 studies, psilocybin is likely to be approved for treatment-resistant depression, and for depression and anxiety in patients with terminal cancer. MDMA-assisted psychotherapy may also be approved to treat PTSD.

A major challenge in conducting randomized placebo-controlled trials of psychedelics is that it is impossible for patients and therapists not to be aware of who has been given a psychedelic drug[4]. Recent trials have used an "active placebo", such as methylphenidate or dextroamphetamine, or used low, moderate and high doses of the psychedelic drug to see if treatment effects are related to dose[5].

It has been argued[6] that psilocybin has a low abuse potential, because it does not produce euphoria in humans or self-administration in animals, and there are much lower rates of regular use of this drug in population surveys than for cannabis, cocaine and opioids. Furthermore, users rapidly develop tolerance to its effects and so do not persist in using it.

Studerus et al[4] reported very few acute, subacute and long-term effects of psilocybin in 110 participants in laboratory studies followed up for 8-16 months. This was a select group in that persons with a family or personal history of psychiatric disorders were excluded and 40% had used a psychedelic drug at least once. The short-term adverse effects were minor: fatigue, headache, lack of energy, and difficulty concentrating the day after. Eleven individuals reported "negative changes in psychological well-being and/or mental functions" after the psilocybin session. One reported "persistent emotional instability, anxiety and depressive feelings" that he "attributed to suppressed memories" released by the drug. He recovered after receiving psychotherapy.

Psilocybin has been described as a "disruptive" treatment because a single dose produces an immediate clinical response – unlike selective serotonin reuptake inhibitors (SSRIs) that require two weeks of treatment – and its benefits are sustained for six months in a substantial proportion of patients[2,3]. It also appears to act by different mechanisms than SSRIs and serotonin and norepinephrine reuptake inhibitors (SNRIs)[2].

The major limitations of the current evidence for psychedelic drugs are interconnected. In the absence of pharmaceutical industry interest, limited support from philanthropic sources has funded the research, restricting trials to relatively small samples of patients because of the cost of doing larger studies. The persons who have done the research believe in the therapeutic value of psychedelic drugs. This is to be expected, given the history of psychedelics and the reputational challenges in conducting clinical research on them.

If psychedelic drugs are introduced in clinical practice, there is a risk that their use will get ahead of the evidence on their safety and efficacy, in much the same way that "medical cannabis" has done[7]. If psilocybin is approved for treatment-resistant depression, patients and prescribers are likely to demand its use as a first-line treatment for severe depression rather than requiring that patients first fail to respond to SSRIs and other antidepressants. It is unclear whether the FDA and other drug regulators will require trials of psilocybin as a first-line treatment. There may also be demands to use psilocybin off-label to treat anxiety disorders. If MDMA-assisted psychotherapy is approved to treat PTSD, there may be demands to use MDMA off-label to treat other anxiety and depressive disorders. If the criteria for who is a qualified therapist are relaxed, MDMA may be used to treat unhappiness, anxiety and existential angst.

The evidence may be used to argue for compassionate access to other psychedelic drugs, such as LSD, mescaline and DMT. It is uncertain if the use of psychedelics will remain under medical supervision for approved disorders, or whether their use will be advocated for spiritual and other nonmedical purposes. A combination of libertarian and utilitarian arguments may be used to justify the legalization of adult use of these drugs for any purpose, because they cause little harm to users and have a low abuse potential[8].

There may also be demands for compassionate access to plant-based psychedelic drugs in advance of any research evidence. US

states may pass citizen-initiated referenda to legalize the medical use of psychedelic mushrooms and plants, such as ibogaine and ayahuasca, by appealing to the putative "entourage" effects of whole plants and the misconception that medicines derived from plants are safer than "synthetic" pharmaceuticals[9].

For all these reasons, we need public funding of independent evaluations of the efficacy of psychedelic drugs. Trials should involve larger numbers of patients who are representative of those clinical disorders for which these drugs may be used, and should include longer-term follow-up evaluations of safety and sustainability of favorable outcomes.

Wayne Hall
University of Queensland National Centre for Youth Substance Use Research, Brisbane, QLD, Australia

1. Grinspoon L, Bakalar J. Psychedelic drugs reconsidered. New York: Basic Books, 1979.
2. Johnson MW, Hendricks PS, Barrett FS et al. Pharmacol Ther 2019;197:83-102.
3. Bogenschutz MP, Ross S. Curr Top Behav Neurosci 2018;36:361-91.
4. Studerus E, Kometer M, Hasler F et al. J Psychopharmacol 2011;25:1434-52.
5. Johnson MW, Griffiths RR, Hendricks PS et al. Neuropharmacology 2018; 142:143-66.
6. Strassman RJ. J Nerv Ment Dis 1984;172:577-95.
7. Hall W, Lynskey M. World Psychiatry 2020;19:179-86.
8. Yaden D, Anderson D, Mattar M. In: Ellens J, Robers T (eds). Psychedelic policy quagmire: health, law, freedom, and society. Santa Barbara: Praeger, 2015:267-93.
9. Ribeiro S. In: Labate BC, Cavnar C (eds). Plant medicines, healing and psychedelic science. Cham: Springer International, 2018:133-49.

DOI:10.1002/wps.20847

Rationale for and usefulness of the inclusion of gaming disorder in the ICD-11

Video games are among the most popular consumer electronic products in the world. They are having a growing mass appeal both as an interactive recreational activity, in which one can engage individually or with other players, and as a passive entertainment in the form of viewership of broadcasted gaming events, including e-sports and live-streamed games (e.g., twitch.tv). Modern games offer a diverse range of unique and highly immersive experiences. Portable consoles and smart devices have promoted the ubiquity of video games by making them easily accessible almost anywhere.

Gaming can produce numerous benefits for many players, including the fulfilment of psychological needs of social relatedness, autonomy and competence. However, over the last three decades, there has been increasing research interest in the phenomenon of problematic gaming. Survey studies and clinical case reports have highlighted that some individuals experience difficulties in regulating their engagement in gaming activities and play to an excessive degree, resulting in mental and physical symptoms as well as functional impairment[1,2]. A meta-analysis[3] reported that the worldwide prevalence of problematic gaming, as defined by standard addiction criteria, can be estimated to be 1-2%.

Internet gaming disorder was considered as a potential mental disorder for the DSM-5, but the decision was for it to be listed only as a condition for further study. The DSM-5 criteria were consistent with substance use and addictive disorders, including reference to loss of control, tolerance, and withdrawal. Gaming disorder is now included in the ICD-11 among "disorders due to addictive behaviours". Here we outline the approach taken in the ICD-11.

In the ICD-11, gaming disorder is defined as a dysfunctional pattern of gaming, characterized by: a) impaired control (e.g., failed attempts to cut or diminish gaming involvement; gaming performed in a more prolonged or intensive way than planned); b) an increasing priority given to gaming to the extent that it takes precedence over other life interests and daily activities; and c) a continued involvement in gaming despite negative consequences for the individual and his/her acquaintances. To meet the diagnosis, the maladaptive gaming pattern has to be either continuous or episodic and recurrent, be manifested over an extended period of time (typically 12 months), and cause psychological distress or significant impairment in personal, family, social, professional, and/or other important areas of functioning.

Several features are key to emphasize. First, the guidelines include only a few essential requirements, making them practical for use in multiple settings by different health care practitioners. Second, the guidelines do not include withdrawal and tolerance, as these are not relevant to gaming[4]. Third, the emphasis on functional impairment is key for differentiating between people with gaming disorder and the large proportion of individuals engaged in intense or persistent patterns of gaming (e.g., 20-30 hours per week) without experiencing associated negative consequences[5].

The decision to introduce gaming disorder in the ICD-11 was guided by epidemiological, clinical and neurobiological studies, as well as data obtained from treatment providers[1,2]. These lines of evidence have consistently shown that problematic gaming behaviours are associated with a range of negative outcomes (e.g., depressed mood, poorer work performance and school grades, worse sleep, interpersonal conflicts). In addition, there is a growing treatment demand internationally for gaming-related problems, particularly among adolescents and young adults, and an increasing number of clinical trials involving self-referred patients seeking help for these problems[6]. The treatment literature, while still developing, indicates that some therapies targeting the mechanisms underlying gaming disorder and promoting adaptive coping strategies can have positive long-term outcomes[7].

Although there is increasing agreement among researchers and practitioners, in the areas of psychiatry, clinical psychology and public health, that gaming-related harms constitute an im-

portant mental health issue[1,2,8], a key concern is the potential for this diagnosis to lead to inappropriate medicalization, policies and treatment[9]. In particular, some researchers have argued that the introduction of the diagnostic category of gaming disorder may encourage the pathologization of all forms of gaming behaviours, including safe or adaptive gaming activities. Certainly, it is important for clinical guidelines to carefully define and delineate harmful and pathological involvement in video games from those behaviours consistent with a healthy passion or hobby. Such considerations are crucial to ensure the clinical validity and utility of a clinical diagnosis[5]. In the ICD-11, this important demarcation includes an explicitly stated reference to functional impairment caused by gaming.

A recent Delphi study[4] provides further support for the ICD-11 approach to gaming disorder. This study involved a representative and international panel of experts asked to critically evaluate, in terms of the available evidence base, all of the proposed gaming disorder criteria according to their diagnostic validity (defined as the extent to which a specific criterion is a feature of the condition), clinical utility (defined as the extent to which a specific criterion is able to distinguish normal from problematic behaviour), and prognostic value (defined as the extent to which a specific criterion is crucial in predicting chronicity of the condition). Following the structured and iterative Delphi expert consensus method, the study indicated that there was strong agreement on the ICD-11 guidelines for gaming disorder, and that these guidelines would enable clinically valid and relevant diagnosis of gaming disorder without pathologizing healthy gaming.

The inclusion of gaming disorder in the ICD-11 is an important step toward meeting global challenges related to harmful overuse of digital technologies. This includes the development of a public health framework that identifies and promotes steps to reduce gaming-related harms[8]. Moreover, the recognition of gaming disorder promotes the value of multiple research efforts, aimed at testing the efficacy and effectiveness of preventive and clinical interventions, and elucidating the etiological mechanisms (e.g., personality, environmental and neurobiological factors) that affect the onset, maintenance and progression of the condition. Research efforts to be promoted are also those aimed at rethinking how to map the effects of gaming on children and adolescents, in particular with regard to the most popular game genres.

The recognition of gaming disorder is likely to encourage steps toward greater social responsibility measures, either enforced by governments and/or developed from within the gaming industry itself. Gaming products are currently largely unregulated, despite concerns that some in-game purchasing systems (e.g., "loot-boxes") are similar to electronic gambling and may financially exploit vulnerable players. Important assistance that the industry can provide includes telemetry data-sharing, disclosure of product design features, and/or access to special populations (e.g., highly engaged users).

The above collaborative efforts will ultimately help individuals who are vulnerable to or affected by gaming-related problems, while recognizing the popular cultural status and the enjoyment of gaming experiences for most people.

Joël Billieux[1,2], Dan J. Stein[3], Jesus Castro-Calvo[4], Susumu Higushi[5], Daniel L. King[6]

[1]Institute of Psychology, University of Lausanne, Lausanne, Switzerland; [2]Centre for Excessive Gambling, Lausanne University Hospitals, Lausanne, Switzerland; [3]South African 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; [4]Department of Personality, Assessment, and Psychological Treatments, University of Valencia, Valencia, Spain; [5]National Hospital Organization Kurihama Medical and Addiction Center, Yokosuka, Japan; [6]College of Education, Psychology, and Social Work, Flinders University, Adelaide, SA, Australia

1. Saunders JB, Hao W, Long J et al. J Behav Addict 2017;6:271-9.
2. Rumpf HJ, Achab S, Billieux J et al. J Behav Addict 2018;7:556-61.
3. Stevens MWR, Dorstyn D, Delfabbro P et al. Aust N Z J Psychiatry (in press).
4. Castro-Calvo J, King DL, Stein D et al. Addiction (in press).
5. Billieux J, Flayelle M, Rumpf HJ et al. Curr Addict Rep 2019;6:323-30.
6. Müller KW, Dreier M, Duven E et al. J Clin Psychiatry 2017;78:e244-51.
7. Wölfling K, Müller KW, Dreier M et al. JAMA Psychiatry 2019;76:1018-25.
8. Stein DJ, Billieux J, Bowden-Jones H et al. World Psychiatry 2018;17:363-4.
9. Van Rooij AJ, Ferguson CJ, Carras MC et al. J Behav Addict 2018;7:1-9.

DOI:10.1002/wps.20848

Preventive psychiatry: a blueprint for improving the mental health of young people

Paolo Fusar-Poli[1-3], Christoph U. Correll[4-7], Celso Arango[8-10], Michael Berk[11-14], Vikram Patel[15,16], John P.A. Ioannidis[17-19]

[1]Early Psychosis: Interventions and Clinical-detection (EPIC) Lab, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; [2]OASIS Service, South London and Maudsley NHS Foundation Trust, London, UK; [3]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy; [4]Department of Psychiatry, Zucker Hillside Hospital, Northwell Health, Glen Oaks, NY, USA; [5]Department of Psychiatry and Molecular Medicine, Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, USA; [6]Center for Psychiatric Neuroscience, Feinstein Institute for Medical Research, Manhasset, NY, USA; [7]Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin Berlin, Berlin, Germany; [8]Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón, Madrid, Spain; [9]Health Research Institute (IiGSM), School of Medicine, Universidad Complutense de Madrid, Madrid, Spain; [10]Biomedical Research Center for Mental Health (CIBERSAM), Madrid, Spain; [11]Institute for Mental and Physical Health and Clinical Translation, School of Medicine, Deakin University, Barwon Health, Geelong, VIC, Australia; [12]Department of Psychiatry, University of Melbourne, Melbourne, VIC, Australia; [13]Orygen Youth Health, University of Melbourne, Melbourne, VIC, Australia; [14]Florey Institute for Neuroscience and Mental Health, University of Melbourne, Melbourne, VIC, Australia; [15]Department of Global Health and Social Medicine, Harvard University T.H. Chan School of Public Health, Boston, MA, USA; [16]Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Boston, MA, USA; [17]Stanford Prevention Research Center, Department of Medicine, Stanford University, Stanford, CA, USA; [18]Department of Biomedical Data Science, Stanford University, Stanford, CA, USA; [19]Department of Epidemiology and Population Health, Stanford University, Stanford, CA, USA

Preventive approaches have latterly gained traction for improving mental health in young people. In this paper, we first appraise the conceptual foundations of preventive psychiatry, encompassing the public health, Gordon's, US Institute of Medicine, World Health Organization, and good mental health frameworks, and neurodevelopmentally-sensitive clinical staging models. We then review the evidence supporting primary prevention of psychotic, bipolar and common mental disorders and promotion of good mental health as potential transformative strategies to reduce the incidence of these disorders in young people. Within indicated approaches, the clinical high-risk for psychosis paradigm has received the most empirical validation, while clinical high-risk states for bipolar and common mental disorders are increasingly becoming a focus of attention. Selective approaches have mostly targeted familial vulnerability and non-genetic risk exposures. Selective screening and psychological/psychoeducational interventions in vulnerable subgroups may improve anxiety/depressive symptoms, but their efficacy in reducing the incidence of psychotic/bipolar/common mental disorders is unproven. Selective physical exercise may reduce the incidence of anxiety disorders. Universal psychological/psychoeducational interventions may improve anxiety symptoms but not prevent depressive/anxiety disorders, while universal physical exercise may reduce the incidence of anxiety disorders. Universal public health approaches targeting school climate or social determinants (demographic, economic, neighbourhood, environmental, social/cultural) of mental disorders hold the greatest potential for reducing the risk profile of the population as a whole. The approach to promotion of good mental health is currently fragmented. We leverage the knowledge gained from the review to develop a blueprint for future research and practice of preventive psychiatry in young people: integrating universal and targeted frameworks; advancing multivariable, transdiagnostic, multi-endpoint epidemiological knowledge; synergically preventing common and infrequent mental disorders; preventing physical and mental health burden together; implementing stratified/personalized prognosis; establishing evidence-based preventive interventions; developing an ethical framework, improving prevention through education/training; consolidating the cost-effectiveness of preventive psychiatry; and decreasing inequalities. These goals can only be achieved through an urgent individual, societal, and global level response, which promotes a vigorous collaboration across scientific, health care, societal and governmental sectors for implementing preventive psychiatry, as much is at stake for young people with or at risk for emerging mental disorders.

Key words: Young people, prevention, mental disorders, preventive psychiatry, psychosis, bipolar disorder, anxiety, depression, evidence-based medicine, neurodevelopment, children, adolescents

(World Psychiatry 2021;20:200–221)

According to the latest World Health Organization (WHO) Global Burden of Disease study, about one billion people of the total global population (7.5 billion) are affected by any mental disorder[1], including psychotic, bipolar or common mental disorders such as depression and anxiety. Overall, about 50% of mental disorders start by the age of 14, and 75% start by the age of 24[2,3]. Young people account for 41% of the current global population (0-14 years: 25.4% and 15-24 years: 15.5%[4]). Justifiably, mental disorders have been called "the chronic diseases of the young"[5].

After their onset, mental disorders often persist, disrupting the capacity for young people to fulfil their potential[6,7], limiting access to mental[8] and physical[9-12] health care, and exposing them to poor education and reduced occupational opportunities[13], stigma, social isolation, discrimination, and violation of human rights[14-16]. Young individuals suffering from mental disorders have higher morbidity and mortality risks for any reason (including suicide[17]) than the general population, translating into a striking 10-20 years reduction in life expectancy[18].

The mental health of the younger generation, and indeed of our future, is already fragile and threatened by exceptional worldwide forces such as an ongoing pandemic, population migrations, economic uncertainties, the sustainability of ecosystems and climate changes[19]. An urgent individual, societal, and global level response is needed to reduce the incidence and burden of mental disorders in young people[6,20]. Preventive approaches in psychiatry lagged behind somatic medicine[21] and emerged only a few decades ago, increasingly gaining traction. At the same time, future advancements require ongoing efforts to identify and overcome their limitations.

This paper addresses these issues, with a focus on reducing the incidence of psychotic, bipolar and common mental disorders. We first summarize the conceptual foundations of preventive psychiatry and then appraise the evidence supporting

different preventive approaches in young people, as well as their current limitations. The knowledge reviewed is then used to develop a blueprint for future preventive research and practice to improve the mental health of young people.

DEFINING PREVENTIVE PSYCHIATRY

This section reviews core preventive psychiatry concepts and frameworks that hold relevance for assessing the evidence and limitations of prevention in young populations and informing future research.

Public health framework

"Possible measures of prevention"[22] for mental disorders have been advocated since the late 19th century. In the early 20th century, an individual with the lived experience of a mental disorder initiated the mental hygiene movement[23], which generated new community practices for preventing mental disorders in young people[24], establishing preliminary public health principles[25] of preventive psychiatry[26]. Therefore, historically, service users and the community have been key actors in the development of preventive psychiatry, a discipline which is closely intertwined with societal and cultural values.

Early work by Leavell and Clark (middle of 20th century) introduced a classification of prevention in medicine[27], which was tailored on the pre-pathogenesis (primary prevention: health promotion and specific protection) and pathogenesis (secondary and tertiary prevention) phases of syphilis[28]. Caplan, in 1964, classified prevention in mental health as follows: a) primary prevention, which "aims at reducing the incidence of new cases of mental disorder and disability in a population"; b) secondary prevention, which "aims at reducing the duration of cases (and therefore the prevalence) of mental disorders, which will inevitably occur in spite of the programs of primary prevention"; c) tertiary prevention, which "aims at reducing the community rate of residual defect, which is a sequel to acute mental illness"[29].

In 1978, Strasser introduced a fourth level of "primordial prevention" to denote activities that prevented the penetration and appearance of risk factors (risk factors increase the likelihood of clinical events, while protective factors decrease this likelihood) into the population itself, as opposed to primary prevention which addresses risk factors to prevent diseases[30]. Finally, Bradford Hill defined nine criteria that may be considered in navigating the difficult question of causation versus plain association: strength of association, consistency across different situations, specificity and temporality between exposure and outcomes, biological gradient, biological plausibility, coherence with present knowledge, experiment (in laboratory and randomized trials), and analogy with similar classes of exposures and outcomes[31,32].

Gordon's framework

The original formulation of the public health framework was disease-oriented, relying on mechanistic linearity of infectious diseases and identification of a clear-cut biological onset. It also ignored epidemiological knowledge on statistical associations between risk/protective factors and clinical events, as well as multifactorial aetiopathologies with a long period of latency[33]. Furthermore, several disorders may be risk factors for other disorders, so all treatments could potentially be labelled as preventive interventions.

In 1983, Gordon[33] addressed these issues in the context of physical illnesses, reserving the term prevention for those individuals who were not "suffering from any discomfort or disability from the disease or disorder to be prevented", thus excluding tertiary prevention as well as antecedents such as clinical high-risk syndromes (see below). Furthermore, Gordon noted that the public health definitions of prevention had little correspondence to interventions offered, and proposed an alternative threefold classification based on the costs and benefits of delivering the intervention: a) universal prevention, "a measure that is desirable for everybody", including actions for the general public which, in many cases, can be "applied without professional advice or assistance"; b) selective prevention, "a procedure [which] can be recommended only when the individual is a member of a subgroup of the population whose risk of becoming ill is above average"; c) indicated preventive measures, that "are advisable only for persons who, on examination, are found to manifest a risk factor, condition, or abnormality that identifies them, individually, as being at sufficiently high risk to require the preventive intervention"[33].

As illustrated in Figure 1, while targeted approaches (i.e., selective and/or

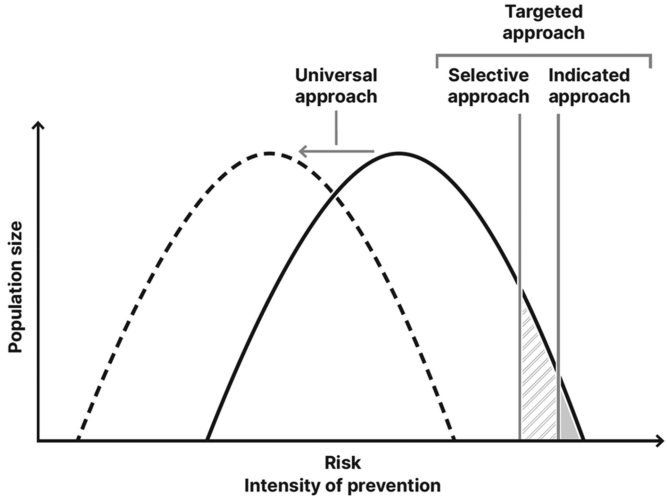

Figure 1 Universal, selective and indicated prevention. Selective and indicated approaches aim to reduce risk amongst those with the most to gain, and therefore reach a small proportion of the population. Universal approaches aim to shift the risk profile of the whole population.

indicated) aim to reduce risk among those with the most to gain, and therefore reach a small proportion of the population, universal approaches aim to shift the risk profile of the whole population.

US Institute of Medicine framework

Gordon's classification was not designed for use in mental disorders. In 1994, the US Institute of Medicine[34] noted that the definition of caseness is more difficult to establish in psychiatry than in somatic medicine, and that the presence of symptoms and dysfunctions is frequent even if diagnostic criteria (ICD/DSM) for mental disorder are not met. Prevention was thus refined as "reducing incidence, prevalence, recurrence of mental disorders, the time spent with symptoms, or the risk condition for a mental illness, preventing or delaying recurrences and also decreasing the impact of illness in the affected person, their families and the society"[34]. The Institute allowed indicated interventions to target antecedents of the disorder, such as clinical high-risk syndromes[34].

It was also acknowledged that, although some people receiving indicated preventive interventions may already have comorbid mental disorders, if they are selected into the intervention based on having early symptoms, then the intervention is still considered preventive[34]. Kessler and Price[35] later refined the concept as primary prevention of secondary psychiatric comorbidities.

The Institute also defined prevention screening to identify risk exposure at population level (for universal prevention efforts, e.g. poverty, violence, lack of health care) or at-risk group/individual level (for selective prevention efforts, e.g. maternal depression or childhood abuse), or to identify core/distinctive characteristics in high-risk individuals (for indicated prevention, e.g. attenuated symptoms, functional impairment or early phenotypic features). Core requisites of prevention screening are identifiable risk/protective factors linked to a disorder, availability of a validated screening tool, an effective intervention to address the identified factors and improve outcomes, solid guidelines on care pathways following screening, wide acceptability to the population, and dynamic implementation of screening procedures[34].

WHO framework

In the current WHO framework (Table 1), universal, selective and indicated preventive interventions are all included within primary prevention[36], and indicated approaches are allowed to target antecedents/clinical high-risk syndromes (see below).

The WHO classifies the management of mental disorders as a continuum encompassing prevention (complementary universal, selective and indicated approaches), treatment (secondary prevention and early or standard treatment), and rehabilitation (tertiary prevention and long-term care). The conceptual boundaries between preventive "interventions" (in "individuals") and "treatments" (in "patients"), particularly in early management[37], are porous at times and associated with several empirical, ethical and societal aspects.

Prevention of mental disorders vs. promotion of good mental health

The WHO broadly defines good mental health 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"[36]. Therefore, mental health is much more than the absence of mental disorders.

Good mental health and mental disorder, although interrelated, are not on a one-dimensional continuum. For example, empirical evidence has associated individual levels of creativity with psy-

Table 1 World Health Organization's classification of preventive approaches for mental disorders[36]

Public health classification of prevention	Gordon's classification of prevention[33], modified by the US Institute of Medicine[34]
Primary prevention seeks to prevent the onset (incidence) of a disorder or illness.	
	Universal prevention is defined as those interventions that are targeted at the general public or a whole population group that has not been identified on the basis of increased risk.
	Selective prevention targets individuals or subgroups of the population whose risk of developing a mental disorder is significantly higher than average, as evidenced by biological, psychological or social risk factors.
	Indicated prevention targets high-risk people who are identified as having minimal but detectable signs or symptoms foreshadowing mental disorder, or biological markers indicating predisposition for mental disorders, but who do not meet diagnostic criteria for disorder at that time.
Secondary prevention seeks to lower the rate of established cases of the disorder or illness in the population (prevalence) through early detection and treatment of diagnosable diseases.	
Tertiary prevention includes interventions that reduce disability, enhance rehabilitation and prevent relapses and recurrences of the illness.	

chotic or bipolar disorders[38,39], and this association has recently been confirmed at a genetic level[40]. Conversely, individuals without mental disorders do not necessarily have good mental health. Normally developing young people can display reactive mild anxiety or depression as physiological adaptive strategies aimed at harm avoidance and extinction of maladaptive behaviours[41].

Therefore, mental health promotion can be implemented across all stages illustrated in Figure 2 (e.g., from healthy people to individuals affected with chronic mental disorders)[34], and not only during the pre-pathological phase (i.e., within primary preventive approaches, as suggested by Leavell and Clark[27]). Promotion of good mental health could also be enhanced by improving physical health, given the close relatedness between these two domains[42].

Neurodevelopmental prevention of mental disorders in young people

As noted by Clark[28], prevention "requires knowledge of the natural history" of a disease. Psychotic disorders are infrequent before the age of 14[43]; their incidence peaks in the age group of 15-35 and declines after the age of 35[44]. The average age of onset for bipolar disorder is 23 years, with a wide range (9 to 37)[45]. The median onset age is earlier for anxiety disorders (11 years of age) versus major depression (32 years)[2]. The range of the age of onset of depressive disorders is typically wider than for many other mental disorders[46].

The pathophysiology of psychotic disorders is generally understood to originate from several genetic and non-genetic risk/protective factors (and their interactions) that impact the neurodevelopment[7,47,48]. Early abnormalities of maturational changes appear from the ectodermal phase to the first year after birth (first-wave hits)[49]. A further phase of significant neurobiological changes is from mid-childhood through pubescence to mid-20s (second-wave hits)[47], when the risk of disorder onset is the highest. Similar neurobiological models have been investigated for bipolar disorder[50,51] and depression[52].

Clinical staging models[53] integrate these epidemiological and neurobiological findings (Figure 2)[47]. The clinical staging model for psychosis is the most established[54,55], but similar models have also emerged for bipolar[56-59], depressive[60,61] and anxiety[62-65] disorders. The premorbid stage starts during the perinatal period and is often asymptomatic and generally associated with preserved functioning (Figure 2). Accumulation of further risk factors from infancy to young adulthood could lead to the emergence of a clinical high-risk stage (Figure 2), characterized by attenuated symptoms that do not meet the diagnostic threshold for mental disorders but are typically associated with some degree of functional impairment. These attenuated symptoms can then progress to a fully symptomatic mental disorder, and then persist into adulthood, especially if treated sub-optimally, leading to a relapsing stage and eventually a chronic stage (Figure 2).

The period from the prenatal/perinatal phase to the onset of the first episode of the disorder may represent the most compelling window of preventive opportunity[7,55]. By integrating the preventive framework within a neurodevelopmentally sensitive clinical staging model, primary prevention (universal, selective and indicated) and promotion of good mental (and physical) health[7] emerge as core strategies to target this critical window (Figure 2).

EVIDENCE SUPPORTING PRIMARY PREVENTION AND MENTAL HEALTH PROMOTION IN YOUNG PEOPLE

This section reviews the evidence supporting indicated, selective and universal preventive interventions and promotion of good mental health, reflecting the in-

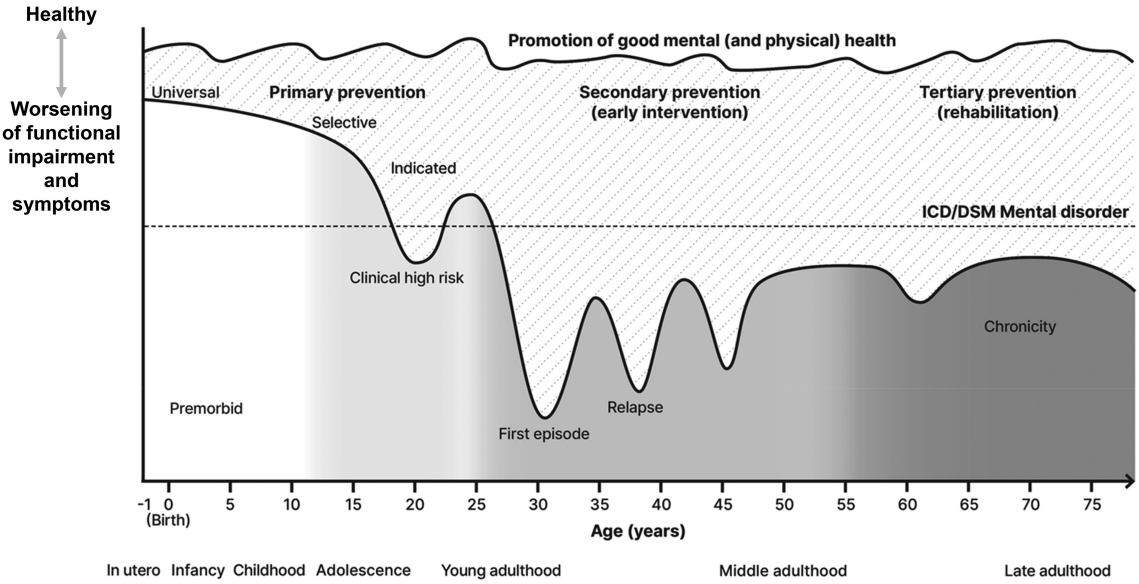

Figure 2 Neurodevelopmental continuum model for prevention of psychosis, bipolar disorder and common mental disorders, and promotion of good mental and physical health

creasing width of these approaches from relatively small subgroups to the wider population (Figure 1).

Indicated preventive interventions

The available evidence supporting indicated preventive interventions for psychotic, bipolar and common mental disorders is summarized in Table 2.

Psychosis

Indicated prevention of psychosis originated in Australia about twenty-five years ago[66] and subsequently gained traction globally, leading to the implementation of specialized services[67] taking care – according to a survey carried out in 2017-2018 – of over than 22,000 young individuals across Western Europe (51.1%), North America (17.0%), East Asia (17.0%), Australia (6.4%), South America (6.4%) and Africa (2.1%)[67]. The consolidation of this paradigm in clinical practice has impacted national[68] and international[69] clinical guidelines and diagnostic manuals (e.g., DSM-5 attenuated psychosis syndrome[70]), although not everywhere[71].

Young (typically 14-35 years old, mean age 21 years[72]) individuals at clinical high risk for psychosis (CHR-P)[73,74] accumulate several risk factors for the disorder[44,75,76], which can lead to functional impairments[77] and the emergence of attenuated psychotic symptoms[78] (which last on average 2 years[72]). Because of these problems, these individuals often seek help[79], including at specialized CHR-P clinical services when available[67,80,81].

Detection of CHR-P individuals is unsystematic and mostly based on referrals made on suspicion of psychosis risk by several agencies and idiosyncratic sampling strategies. This recruitment phase nevertheless leads to substantial risk enrichment in help-seeking samples[82]. Although several screening instruments for CHR-P have been tested, their validation is currently limited[83].

In CHR-P clinics, help-seeking individuals undergo a semi-structured psychometric assessment with validated instruments, which deliver a group-level estimate for predicting psychosis (i.e., at risk vs. not at risk)[74]. The CHR-P criteria are robustly associated with psychosis onset (odds ratio, OR=9.32)[44] within high-risk clinical sam-

Table 2 Level of evidence for available indicated interventions to prevent (reduce the incidence) of psychotic, bipolar and common (depression/anxiety) mental disorders in young people

	Psychotic disorders	Bipolar disorder	Depression/Anxiety disorders
Target	Clinical high risk for psychosis (CHR-P)[72]****	Bipolar at-risk states[97,105]*, bipolar prodrome[103,104]*	Not available
Detection			
Referral, risk enrichment	On suspicion of psychosis risk, 15% at 3 years[82]***	On suspicion of bipolar risk	Screening in schools, universities or primary care[114,115]***
Screening instruments (sensitivity, specificity, positive predictive value, negative predictive value)	Several, but poor validation (67-100%, 39-100%, 24-100%, 58-100%)[83]**	BPSS-AS-P (data not available)[103]*	Some, but none validated (data not available)[114]***
Duration of attenuated symptoms	709 days[72]****	107.9 months[98]***	Not available
Mean age (SD) or range	21 (3.2) years[72]****	16-23 years[103-105]*	18-25 years[113]***
Prognosis			
Assessment instruments (accuracy)	CAARMS[287]***, SIPS[288]***, DSM-5 APS [288]*** (0.90 pooled at 38 months)[74]*** Not recommended outside clinical samples[74]***	BPSS-FP (data not available)[104]*, SIBARS (0.7 at 18 months)[105]* Used in clinical samples only[104,105]*	Not available
Transition risk	17% at 1 year; 22% at 3 years (BLIPS>APS>GRD)[88]***	14% at 1 year[289]*; 23% at 2 years[105]*	Not available
Intervention			
Type of intervention (efficacy)	Needs-based interventions, psychotherapy, pharmacotherapy, combinations (no evidence for superior efficacy in preventing psychosis or improving other outcomes)[93,94]***,[72,251]****	Family-focused therapy (reduced time to recovery, no effect on incidence of bipolar disorder)[106]* Individual psychotherapy (no efficacy on affective symptoms)[107]*	Psychotherapy/psychoeducation (reduced severity of depressive/anxiety symptoms[113,115]***, but not with digital psychoeducation[119]*** and not in humanitarian settings[120]***; no evidence of effect on incidence of depressive/anxiety disorders[113,115]***)

* single study, ** systematic review, *** meta-analysis, **** umbrella review. APS – attenuated psychotic symptoms, BLIPS – brief limited intermittent psychotic symptoms, BPSS-AS-P – Bipolar Prodrome Symptom Scale - Abbreviated Screen for Patients, BPSS-FP – Bipolar Prodrome Symptom Interview and Scale-Full Prospective, CAARMS – Comprehensive Assessment of At Risk Mental States, GRD – genetic risk and deterioration syndrome, SIBARS – Semistructured Interview for Bipolar At Risk States, SIPS – Structured Interview for Psychosis-Risk Syndromes.

ples (but not in the general population[84]), while they cannot predict new cases of bipolar or common mental disorders[85,86].

In CHR-P samples, most (~85%) individuals present with attenuated psychotic symptoms (APS), ~10% with short-lived frank psychotic symptoms (brief and limited intermittent psychotic symptoms, BLIPS), and ~5% with schizotypal traits or a relative affected with psychosis coupled with functional decline (genetic risk and deterioration, GRD)[72]. Since most (68%) individuals with BLIPS also meet ICD-10 criteria for an acute and transient psychotic disorder[87], interventions in CHR-P people extend beyond primary indicated prevention (for APS and GRD) into secondary prevention (for BLIPS). The overall risk of developing psychosis (22% at 3 years) differs across these three subgroups[88].

Transition to psychosis is associated with clinically meaningful real-world outcomes[89] and is modulated by baseline levels of attenuated positive psychotic (OR=2.56) and negative (OR=2.68) symptoms, while good functioning reduces the risk (OR=0.59)[44].

Indicated prevention implemented in CHR-P services (the NICE-recommended intervention is cognitive behavioural therapy[68]) has the potential to ameliorate presenting symptoms, delay or prevent the onset of psychosis, reduce health care access and duration of untreated psychosis (secondary prevention)[55,90]. Furthermore, CHR-P services routinely incorporate comprehensive needs-based interventions focusing on psychosocial, vocational and familial requirements, as well as several public health initiatives such as outreach campaigns in collaboration with the local community (e.g., non-governmental organizations, youth centres, schools, colleges, faith groups; low-income, racial/ethnic, sexual and gender minorities) to foster mental health literacy (e.g., reducing illicit substances use, enhancing self-coping strategies) and promote good mental (e.g., resilience, positive lifestyle behaviours) and physical health[72].

Earlier meta-analyses of randomized controlled trials suggested a significant preventive effect for psychological interventions[91,92]. However, the most updated network meta-analysis[93] found no robust evidence to favour any of these indicated interventions compared to each other or needs-based interventions. A second independent pairwise meta-analysis by the Cochrane group confirmed these findings, concluding that "there was no convincing unbiased, high-quality evidence" to suggest that any type of intervention is more effective than others, including needs-based interventions[94] (another meta-analysis was recently published[95], but used older data than the above-mentioned ones[93,94]).

Bipolar disorder

Indicated prevention in bipolar disorder was developed, following the CHR-P template, only fifteen years ago[96,97], and is rapidly emerging[98-101]. The supporting evidence lags behind that for CHR-P[99,102].

Detection of a symptomatic clinical high risk for bipolar disorder is complicated by its inherent episodicity, long duration and the complex nature and definition of the disorder[98]. Individuals at clinical risk for bipolar disorder are represented by young help-seeking clinical samples[97] (mean age 16-23 years[103-105]), including a subset of CHR-P individuals[105], who present with attenuated bipolar-risk features (which last on average 9 years[98]). Self-administered screening instruments have been developed, but require further validation[103].

With respect to assessment, early manifestations – such as sleep disturbance, anxiety, irritability, cyclothymic features, manic or hypomanic symptoms, and depression – are non-specific[100]. Emerging semi-structured interviews can rate sub-threshold manic, depressive and general symptoms[104] to define high-risk subgroups in clinical samples: sub-threshold mania, depression and cyclothymic features, genetic risk and depression, genetic risk and cyclothymic features, sub-threshold mixed episode, mood swings[105]. The prospective validity of these instruments awaits validation, despite some promising pilot findings[105].

Interventional research is in its infancy. Two randomized controlled trials conducted in young people presenting with genetic risk for (schizo)affective disorder and attenuated affective symptoms suggested a potential beneficial effect of family-focused and cognitive behavioural therapy on time to recovery from attenuated symptoms[106], but no efficacy in terms of reducing the severity of affective symptoms[107] or preventing the onset of bipolar disorder[106].

Common mental disorders

Indicated prevention of depression and anxiety disorders in young people still represents a "blind spot in health care"[108-110] and has been less investigated than selective/universal approaches[111]. There is also some degree of overlap with indicated prevention for bipolar disorder, because subthreshold/frank depressive episodes (especially the atypical phenotype) and cyclothymic features or genetic risk for depression coupled with bipolar-like features are already subsumed in the clinical criteria for bipolar risk[112].

Young people[113] at clinical high risk for depression/anxiety disorders have been detected through psychometric screening for sub-threshold symptoms in schools, universities or primary care[114,115], typically following selective/universal screening[116]. However, results do not suggest that such screening is ready for wider use. Beyond these attempts, there are no established clinical high-risk criteria to assess young people with an increased risk of depression (without bipolar risk features) or anxiety disorders and predict their outcomes.

Early meta-analyses not focusing on young individuals showed that indicated psychological interventions, generally based on cognitive behavioural therapy, can reduce the incidence of depression[114,117], and that these interventions can be effectively delivered digitally in middle-aged adults[118]. However, the most recent meta-analysis focusing on young people with baseline sub-threshold depression (along with selective/universal approaches) found that none of the included psychological intervention studies measured the incidence of emerging depression[113]. Another recent meta-analysis confirmed that there is no evidence favouring digital psychoeducation over no intervention to

improve depressive symptoms in young people[119].

Meta-regression analyses showed that psychological/psychoeducational interventions might be effective in reducing the severity of some anxiety symptoms in young people, but no conclusion could be drawn concerning prevention of the onset of anxiety disorders[115]. A meta-analysis showed that indicated psychological/social interventions are not effective to prevent anxiety/depression in people living in low- and middle-income countries affected by humanitarian crises[120].

Selective preventive interventions

Selective preventive interventions in the premorbid stage of psychotic, bipolar and common mental disorders (summarized in Table 3) would require screening and reducing the exposures to identified detrimental factors in at-risk groups before symptoms and help-seeking behaviour manifest[76].

This approach would require robust aetiopathological knowledge of the association between specific genetic and non-genetic factors and incidence of these disorders (and effective interventions). However, comprehensive explanatory pathophysiology is not established in psychiatry, and no singular putative causal factor fully meets Bradford Hill criteria, so that current diagnostic manuals (ICD-11/DSM-5) refer to mental syndromes (i.e., disorders) and not pathophysiological processes (i.e., diseases).

Genetic factors

Many genetic variants have been identified that modulate the risk for psychotic,

Table 3 Level of evidence for at-risk group exposures and available selective interventions to prevent (reduce the incidence) of psychotic, bipolar and common (depression/anxiety) mental disorders in young people

	Psychotic disorders	Bipolar disorder	Depression/Anxiety disorders
At-risk group exposures (association with the disorder)	**Genetic risk/protective factors:** 22q11.2 deletion syndrome (prevalence 10-41%[132]*, risk 37% at 32 months[135]*) Offspring (RR=7.54)[121]*** Twins (monozygotic concordance rate 40%)[123]* First-degree relatives (one proband: OR=7.69; two probands: OR=11.11)[127]*** **Non-genetic risk/protective factors:** Black-Caribbean ethnicity in England (OR=4.87)[44]**** Ethnic minority in low ethnic density area (OR=3.71)[44]**** Second-generation immigrants (OR=1.68)[44]**** Trait anhedonia (OR=4.41)[44]**** Minor physical anomalies (OR=5.30)[44]**** Premorbid IQ (OR=0.47)[44]**** Olfactory identification ability (OR=0.19)[44]**** Several prenatal/perinatal factors (OR=0.86 to 3.05)[150]*** Physical activity (OR=0.728)[235]**** Smoking (OR=1.99)[235]**** **Peripheral biomarkers** Decreased pyridoxal (vitamin B6) levels (data not available)[147]****	**Genetic risk/protective factors:** Offspring (RR=4.06)[121]*** Twins (monozygotic concordance rate 45%)[124]* First-degree relatives (one proband: RR=6.10, two probands: RR=29.1)[128]* **Non-genetic risk/protective factors:** Irritable bowel syndrome (OR=2.48)[144]**** Childhood adversity (OR=2.86)[144] **** Physical activity (OR=0.49)[235]**** Smoking (OR=1.46)[235]**** Poor sleep (OR=1.79)[235]**** **Peripheral risk/protective biomarkers:** Elevated awakening cortisol levels (g=0.25)[147]****	**Genetic risk/protective factors:** Offspring (depression: RR=2.38[121]***; anxiety: RR=1.76 [122]***) Twins (monozygotic concordance rate – depression: 46%[126]*; anxiety: 13-73%[125]*) First-degree relatives (anxiety: OR=4.1-6.1[129]*; depression: one proband OR=2.14, two probands OR=3.23[130]***) **Non-genetic risk/protective factors:** Sedentary behaviour (RR=1.25)[145]**** Sexual dysfunction (OR=2.71)[145]**** Four or five metabolic risk factors (OR=2.06)[145]**** Obesity (OR=1.35)[145]**** Job strain (OR=1.77)[145]**** Physical abuse in childhood (OR=1.98)[145]**** Early physical trauma (OR=2.59)[146]**** Physical activity (OR=0.837)[235]**** Smoking (OR=1.73)[235]**** Healthy diet (OR=0.77)[235]**** Poor sleep (OR=2.27)[235]****
Type of intervention (efficacy)	Screening for family history of psychotic disorder (data not available)[132]* Screening pregnant/postnatal women for emerging psychopathology (data not available)[148]***	Psychoeducation for young people at risk (improved affective symptoms but no evidence of effect on incidence of bipolar disorder)[155]***	Screening for family history of depression and psychoeducation (improved depressive symptoms and reduced incidence of depression in offspring)[136]*** Screening for post-partum depression and psychoeducation/psychotherapy (inconclusive evidence)[152]*** Psychological interventions in women disclosing partner violence (improved anxiety but not depression)[153]*** Psychological/psychoeducation (improved anxiety symptoms[115]***, but not as school-based interventions[156]*** and not in humanitarian settings[120]***; no evidence of effect in preventing depression/anxiety disorders[115]***) Physical exercise in at-risk youths (reduced severity of depression[157]*** and incidence of anxiety[174]***)
	Behavioural counselling to prevent illicit substance use in at-risk adolescents and young adults (no evidence of efficacy)[154]***		

* single study, ** systematic review, *** meta-analysis, **** umbrella review. OR – odds ratio, RR – risk ratio

bipolar or common mental disorders, but almost all of them have very small, and thus clinically unclear, effects for selective screening. Polygenic risk scores have been developed to overcome these limitations by analyzing genetic variants *en masse*[48], but the variance explained is still too small for implementation in selective prevention and does not provide singular neurobiological targets.

For example, offspring of patients affected with psychosis, bipolar disorder or depression have a greater risk of developing these disorders (32% by adulthood)[121,122]. Monozygotic twins[123-126] and first-degree relatives (depending on the number of probands)[127-130] also have an increased likelihood of developing these disorders. However, only 17.4% of the association between family history of psychosis and the disorder is mediated through a modelled polygenic risk score[131]. The only molecular risk factor for psychosis that may have a preventive relevance is the 22q11.2 deletion syndrome, which is characterized by high rates of schizophrenia (prevalence from 10% in adolescents to 41% in young adults)[132].

Overall, familial vulnerability (along with 22q11.2 deletion syndrome) represents the most implementable target for selective screening intervention in health care[133]. It is more established for psychosis[134], but it is emerging for bipolar disorder. One possible intervention could be monitoring and psychometric assessment for a CHR-P/bipolar-risk state when symptoms or functional disability develop[135].

The associated preventive capacity is, however, limited: while a meta-analysis found that selective psychoeducational interventions may have a small effect on reducing the severity and incidence of depression in the offspring of patients[136], the preventive efficacy of other psychosocial interventions in young individuals with a familial vulnerability for psychotic[137], bipolar[138] or anxiety disorders is currently unknown.

Non-genetic factors

Similarly, non-genetic factors have not yet entered selective screening[139,140]. This situation is mostly due to the intrinsic complexity of the psyche itself[141], and conflicting research findings that are characterized by several biases such as high heterogeneity, excess significance, selective reporting of statistically significant (i.e., "positive") results and no adjustment for multiple confounders[142,143]. Table 3 lists non-genetic factors, along with their meta-analytic strength of association (according to established criteria to classify the evidence) with psychotic[44], bipolar[144], depressive[145] and anxiety[146] disorders.

Among 733,316 measurements on 162 different peripheral biomarkers for psychosis, bipolar disorder and depression, only two were found to be reliably associated with these disorders[147](see Table 3). Studies targeting inflammatory biomarkers using anti-inflammatory therapies like aspirin[148] or targeting individual nutrients such as vitamin D[149] to prevent depression have not turned out to be effective approaches, at least in adults, dampening hopes in youth[133].

Within risk/protective factors listed in Table 3 (their distinction from biomarkers may be challenging without clear pathophysiological knowledge), the majority exert their role before the age of 25 years, and some are potentially modifiable in vulnerable groups. For example, the evidence concerning several prenatal/perinatal risk factors laid the rationale for screening pregnant/postnatal women for emerging psychopathology in order to detect an incipient risk of psychosis or post-partum depression[150,151]. However, the risk may not be high enough to make such screening clinically useful. Furthermore, a meta-analysis investigating psychological/ psychoeducational selective interventions (along with universal/indicated ones) to prevent post-partum depression in pregnant/postnatal women[151] found considerable cost-effectiveness uncertainty[152].

Women disclosing current or recent intimate partner violence exposure represent another vulnerable group. A meta-analysis found that selective psychological interventions can reduce their anxiety (but not depression) even in in low/middle-income countries[153].

Another potentially modifiable risk factor selectively targeted across psychotic, bipolar and common mental disorders has been the initiation of illicit and non-medical drug use among adolescents and young adults. However, a recent meta-analysis by the US Preventive Service Task Force found no evidence to favour selective (as well as population-level/universal) behavioural counselling[154].

Selective psychological/social interventions are not effective to prevent anxiety/depression in humanitarian settings[120], and there is scarce preventive research in other vulnerable subgroups such as racial/ethnic, sexual and gender minorities.

Selective approaches have also been tested in various subgroups of at-risk youths. A recent meta-analysis reviewed the efficacy of selective (along with universal) interventions for young people (across different settings), finding that psychoeducation may be the most effective preventive intervention for improving affective symptoms (Hedges' g=0.6), but there was no efficacy on the incidence of mood disorders[155]. Another meta-regression analysis showed that selective (as well as universal) psychological/psychoeducational interventions delivered across different settings (e.g., community schools and colleges, primary care clinics) might be effective in reducing some anxiety symptoms in young people, although findings were inconclusive regarding prevention of depression/anxiety disorders[115]. Recent sensitivity (network) meta-analyses found little evidence that selective (and universal/indicated) school-based educational interventions are effective for the prevention of common mental disorders in young people[156].

Importantly, a recent umbrella review has documented that an exercise intervention may be effective in reducing depressive symptoms in at-risk youths[157]. However, even the possible benefits at the level of symptoms may be due to selective reporting and other biases for what are largely subjective outcomes in unmasked trials.

Universal preventive interventions

As shown in Figure 1, universal preventive strategies (summarized in Table 4)

Table 4 Level of evidence for population-level exposures and available universal interventions to prevent (reduce the incidence) of psychotic, bipolar and common (depression/anxiety) mental disorders in young people

	Psychotic disorders	Bipolar disorder	Depression/Anxiety disorders
Population-level exposures (association with the disorder)	**Surrogate markers**: psychotic experiences (risk of psychosis 0.5-1 per year)[290]***	**Surrogate markers**: K6/10 (data not available)[163]*	**Surrogate markers**: K6/10 (data not available)[163]*
	Neurodevelopmental biomarkers (data not available)[164,165,167]*		
	Social determinants of mental disorders (data not available)[173]**** Demographic (community diversity, population density, longevity, survival) Economic (economic recessions, economic inequalities, macroeconomic policy) Neighbourhood (infrastructure, neighbourhood deprivation, built environment settings) Environmental events (natural/industrial disasters, war or conflict, climate change, forced migration) Social/cultural (community social capital, social stability, culture)		
Type of intervention (efficacy)	Screening for psychotic experiences (data not available)[161]*	Screening for bipolar experiences (data not available)[163]*	Screening for depressive/anxiety experiences (data not available)[163]*
	Perinatal phosphatidylcholine (modulated biomarkers of neonatal brain development[164]*; fewer attention problems and less social withdrawal[165]*); perinatal folate acid (improved executive functioning)[167]*; vitamin D, polyunsaturated fatty acids (inconclusive evidence)[166]**	Psychoeducation for young people (improved affective symptoms but no evidence of effect on incidence of bipolar disorder)[155]***	Psychological/psychoeducation (improved anxiety symptoms[115]***, but not as school-based interventions[156]*** and not in humanitarian settings[120]***; no evidence on preventing depression/anxiety disorders[115]***) Public health strategies on school climate (improved depressive symptoms)[171]* Physical exercise (reduced incidence of anxiety disorders)[174]***
	Reduction of gender-based violence, child maltreatment, racial discrimination and xenophobia; basic income grants and improved employment; safe neighbourhoods; reductions in violence; early response to environmental events; action on protecting vulnerable ecosystems; improved education (data not available)[173]****		
	Behavioural counselling to prevent illicit substance use in adolescents and young adults (no evidence of efficacy)[154]***		

* single study, ** systematic review, *** meta-analysis, **** umbrella review. K6/10 – Kessler Distress Scale 6- or 10-item

would theoretically allow a population-wide reduction in incidence/burden of psychosis, bipolar and common mental disorders in young people, producing wider societal-level benefits compared to indicated/selective measures.

Universal strategies may take the form of a safe intervention that: a) decreases exposures to population-level risk factors (most of the at-risk group exposures listed in Table 3 could be, in principle, considered as well for population-level universal approaches) and/or b) increases exposure to population-level protective factors. However, pathophysiological knowledge is limited, and there is a lack of methods to readily assess the efficiency of such interventions.

In line with Gordon's observations, psychosis and bipolar disorder are characterized by a low incidence and long latency between exposures and the manifestation of the disorders (the latter point also applies to common mental disorders). Demonstrating an impact on the incidence of these disorders would be, if at all feasible, long and expensive[158].

Decreasing exposures to population-level risk factors

A possible avenue may be to use surrogate population-level markers that may predict the effect of universal interventions on the incidence of disorders and that are convenient to measure. For example, "psychotic experiences"[159] are relatively frequent at the population level (prevalence about 8% in young adults aged 24[160]) and can be measured through self-administered questionnaires (e.g., Prodromal Questionnaire, PQ)[161]. These mostly transitory sub-threshold manifestations are not to be conflated with clinical psychotic symptoms (see below)[162], but could represent a potential surrogate marker of psychosis (risk of psychosis: 0.5-1% per year[160]). Other self-administered instruments, such as the Kessler Psychological Distress Scale (6 or 10 items, K6/10)[163], could theoretically be used as surrogate markers for bipolar, depressive and anxiety disorders. However, to date, there is no preventive capacity associated with these surrogate markers.

Similarly, neurodevelopmental surrogate biomarkers have been used to test dietary phosphatidylcholine supplementation in healthy pregnant women[164]. Phosphatidylcholine is an agonist at alpha-7 nicotinic receptors, which are involved in the final maturation of GABA inhibitory synapses before birth, and have been implicated in schizophrenia[164]. A first randomized controlled trial confirmed the effect of perinatal phosphatidylcholine on an electrophysiological biomarker of foetal development[164]. A subsequent study demonstrated that, at 40 months, phosphatidylcholine impacted neurocognitive biomarkers, leading to fewer attention problems and less social withdrawal compared with the placebo group, thus potentially altering the risk of later development of psychosis[165].

Another dietary intervention involved folic acid supplementation in pregnancy (folate is important in neurogenesis, cell growth and proliferation, and myelination), which has become one of the most important public health advances in medi-

cine[166]. A randomized controlled trial demonstrated that folate supplementation could improve some neurocognitive biomarkers in children 8.5 years later[167]. Other compounds for use in pregnancy (vitamin D[168], polyunsaturated fatty acids[166]) have been suggested, but no randomized controlled trials have been conducted, and the overall evidence is inconclusive[166].

Several other compounds have demonstrated hints of efficacy on experimental neurodevelopmental biomarkers (e.g., neonatal N-acetylcysteine[62], sulphoraphane[169], modulation of microbiota[170]) and are under investigation in humans (not listed in Table 3). However, these surrogate markers have not been well validated, and thus it is unknown whether they would indeed translate to preventive benefits. Furthermore, it is important to have fully preregistered protocols, including details on which biomarkers will be collected and how/when they will be analyzed. The large number of markers and analytical options allows for a situation where spurious "positive" results may emerge/be more likely published.

Beyond surrogate markers, universal psychoeducation and psychological interventions [115,155,156] have been frequently tested (blended with selective interventions, Table 3) for young people. Psychological interventions may improve affective symptoms[155], while psychotherapy/psychoeducation may improve some anxiety symptoms[115] (but not as school-based education intervention[156]). Multi-component public health and youth engagement strategies impacting the overall school climate (rather than individual behaviour change) may improve depressive symptoms (along with physical health outcomes)[171,172]. However, there is no evidence that they can impact the incidence of depression/anxiety disorders[115]. As noted above, universal interventions were not effective to prevent illicit substance use in the general adolescent and young adult population[154], or to prevent common mental disorders in humanitarian settings[120].

To date, the most established population-level exposures encompass social determinants of mental disorders, which have become the cornerstone of public health prevention. A large umbrella review has summarized about 300 (mostly observational) papers on social determinants of psychotic, bipolar and common mental disorders, and empirically linked them with the Sustainable Development Goals promoted by the United Nations Member States in 2015 (demographic, economic, neighbourhood, environmental events, social and cultural domains)[173]. For example, there is strong evidence that adverse social and economic circumstances – including poverty, income inequality, interpersonal and collective violence, and forced migration – are key risk determinants of psychotic disorders[173].

The umbrella review identified several interventions that lie at the interface between universal, primordial and promotion approaches and could potentially lead to high benefit for young people: reduction of gender-based violence, child maltreatment, racial discrimination and xenophobia, basic income grants and improved employment, safe neighbourhoods, reductions in violence, early response to environmental events, action on protecting vulnerable ecosystems, and improved education[173]. However, the review acknowledged that future trials should demonstrate the direct effect of these interventions on psychotic, bipolar or common mental disorders; furthermore, many implementation challenges remain unresolved[173].

Increasing exposures to population-level protective factors

Current evidence is mostly limited to the promotion of good mental health (reviewed below). Other approaches have focused on universal physical exercise interventions in young people, to foster resilience and additionally relieve the associated physical health burden. A recent umbrella review has demonstrated that an exercise intervention may be potentially effective in reducing the incidence of anxiety[174] in the general young population. Universal exercise interventions have also been suggested for psychotic[175] and bipolar[176] disorder. Interventions promoting positive lifestyle behaviours are under development (see below). However, there is not yet solid evidence demonstrating that these interventions can prevent psychotic, bipolar or common mental disorders (see below).

Promotion of good mental health

Promotion of good mental health (not summarized in Tables 2-4) has received less research attention than prevention of mental disorders, mostly because operationalization of outcomes have been fragmented[41]. Mental health promotion is also highly sensitive to different systems, cultures or clinical practices that differ in values. However, core domains of good mental health have been empirically proposed[177], encompassing mental health literacy, attitude towards mental disorders, self-perceptions and values, cognitive skills, academic/occupational performance, emotions, behaviours, self-management strategies, social skills, family and significant relationships, physical health, sexual health, meaning of life, and quality of life[41].

The consistency and magnitude of available interventions to promote good mental health in young people are similarly patchy and conflicting, comprising psychoeducation (including parent training)[178,179], psychotherapy[180,181], and less frequently physical therapy[182], pet[183] or art[184] therapy.

A meta-analysis appraised the efficacy of these interventions aimed to promote good mental health in asymptomatic young people[185]. Compared to controls, available interventions significantly improved mental health literacy (Hedges' g=0.685), emotions (g=0.541), self-perceptions and values (g=0.490), quality of life (g=0.457), cognitive skills (g=0.428), social skills (g=0.371), physical health (g=0.285), sexual health (g=0.257), academic/occupational performance (g=0.211) and attitude towards mental disorders (g=0.177)[185]. Another recent umbrella review showed that positive psychology could increase subjective well-being[186]. Although several interventions could be effective, evidence was of modest quality, and it is unknown whether these interventions can later impact the incidence of psychotic, bipolar or common mental disorders.

FUTURE DIRECTIONS OF RESEARCH AND PRACTICE

In this section, we integrate the conceptual frameworks with the evidence reviewed and suggest ten core ways toward advancing research and practice to prevent psychotic, bipolar and common mental disorders in young people.

Universal or targeted? Integrating preventive frameworks

An intense debate has lately centred on the antithesis between targeted and universal interventions for young people. Some authors[187] have split the field into proponents[188-190], opponents[191-193], and those with ambivalent attitudes[194] towards targeted interventions. A frequent criticism is that indicated prevention implemented in CHR-P clinics should be replaced by universal/public health approaches, aimed for example to decrease cannabis use (an environmental risk factor for psychosis) in young people[195]. Similar criticisms are emerging for the indicated prevention of clinical high-risk states for bipolar disorder[196,197]. The overarching supporting argument is that targeted interventions represent a "prevention paradox"[187], because they can only benefit a small minority of young people[198].

It is true that CHR-P clinics can currently detect only a minority of individuals who will later develop psychosis[199] (similarly, early intervention services can only detect about half of first episode cases[200]), but research innovations to overcome this limitation are under development[198,201]. Notably, this criticism overlooks the fundamental conceptual point illustrated in Figure 1: targeted approaches are expected *a priori* to target the tip of the iceberg of the population-level risk, and are thus complementary and not antithetical to universal approaches. Furthermore, mainstreaming universal approaches to reduce cannabis abuse holds only theoretical foundation, because these approaches are not empirically effective in children, adolescents and young adults[154].

Future research and clinical practice should better incorporate the continuum model for preventive psychiatry illustrated in Figure 2, which integrates universal, selective and indicated approaches to synergistically and complementarily maximize their efficiency in young people, and indeed across the age spectrum. For example, school-based interventions to prevent anxiety and depression in children and young people are conceived as multilevel, systems-based interventions[156] that encompass different modalities. Another example concerns the quest for effective suicide prevention initiatives in young people, where no single strategy clearly stands above the others, and combinations of individual- and population-level strategies have been recommended[202]. A further example may be the implementation of a stepped or sequential assessment framework encompassing face-to-face CHR-P or bipolar-risk assessment (indicated prevention) following universal screening with self-assessment instruments (e.g., PQ, K6/10)[203], and the enhancement of public health approaches already partially implemented by CHR-P services in the local community. Available meta-analyses show that targeted and universal interventions can be blended together in young people to help preventing postnatal depression[152] or anxiety[115].

In line with these arguments, the Lancet Commission on Global Mental Health called for a joint global initiative on preventive psychiatry integrating public health/universal and targeted approaches[173]. However, if single interventions are not effective, it is yet unclear how exactly their combination could be optimally effective.

Advancing multivariable, transdiagnostic, multi-endpoint epidemiological knowledge

As noted by Leavell and Clark[27], robust genetic and environmental epidemiological knowledge is required to inform evidence-based preventive approaches. We have demonstrated above that this knowledge is currently limited, and several advancements are needed.

To date, non-genetic factors have been mostly measured in univariate analyses that cannot control for their intercorrelation. Future epidemiological studies are required to augment polygenic risk prediction by collecting multiple non-genetic exposures in the same individuals, using poly-environmental risk scores (e.g., psychosis poly-risk score, PPS) recently developed[204], and exploring their interaction with lifestyle behaviours (see below).

Environmental exposures can be measured with digital health technologies (electronic medical records, mobile apps)[205], but pose more challenges to measure massively: for example, measurement error, missing data and selection biases may be prominent, and operational definitions of environmental exposures may vary across and even within datasets. Collaborative harmonization efforts should mitigate these obstacles and integrate polygenetic and poly-environmental information to better map the complex pathophysiology of psychotic, bipolar and common mental disorders.

Another area of future research is the identification of protective and resilience factors. To date, the disease-centric model of research has inhibited the investigation of resilience factors that predict good outcomes (and that, therefore, cannot simply be defined as the inverse of risk factors). Shared definitions of good outcomes should also be developed, in particular with respect to promotion of good mental health, which is currently too fragmented. For example, in the CHR-P field, there is a current refocus on good outcomes beyond psychosis onset (e.g., functional status, remission, quality of life[206]). Importantly, these outcomes hold transdiagnostic potential to accommodate multi-endpoint numerators across psychotic, bipolar and common mental disorders (as well as across physical health disorders) that are essential to justify the denominator of preventive (universal/selective/indicated) effort and cost. For example, social functioning is a shared domain across schizophrenia, depression and neurodegenerative disorders such as Alzheimer's disease[207].

Transdiagnostic approaches have been suggested to complement current psychiatric nosography[208], which is intrinsically limited, in particular in young people[209,210], by integrating clinical staging models and optimizing preventive efforts. However, to date, transdiagnostic approaches have been limited by several methodological

caveats that should be addressed by future research.

First, there are frequently reporting inconsistencies (e.g., definition of the gold-standard DSM/ICD diagnoses, outcome measures, and type of transdiagnostic approach) and low quality of studies, with few findings externally replicated[211]. Future studies could use the TRANSD recommendations, that may help improving the reporting of transdiagnostic research[212]. Second, while psychotic, bipolar and common mental disorders exhibit both multifinality (the same aetiological agents can result in different mental health disorders) and equifinality (multiple agents can lead to the same disorder), knowledge into shared risk/protective factors (Table 3) is still limited. The latter are mostly limited to social determinants of mental disorders, childhood adversity and familial vulnerability (and physical health/lifestyle behaviours discussed below). For example, risk of mood disorders is significantly increased among offspring of parents with schizophrenia (relative risk, RR=1.62), while the risk of schizophrenia is significantly increased in offspring of parents with bipolar disorder (RR=6.42)[121]. However, there is also evidence for diagnostic specificity: machine learning reclassification studies demonstrated a distinction between schizophrenia and mood disorders[213]; treatment requirements and outcomes also differ[55]. Similarly, while early neurocognitive functioning has been suggested as a promising transdiagnostic biomarker[214], some studies suggest that it is more specific to psychosis than to common mental disorders[215].

No convincing evidence supports the existence of a truly transdiagnostic biomarker[147]. Evidence supporting a transdiagnostic clinical staging model that cuts across psychotic, bipolar and common mental disorders[216,217] is similarly limited to a few studies[218], with scarce empirical validation[219]. There are also concerns that the natural course of bipolar[220] and depressive[221] disorders does not necessarily or consistently follow a clinical staging model. However, future research in this field is expected. For example, pervasively reduced neocortical thickness was recently found to be shared across psychotic and common mental disorders, representing a potentially transdiagnostic marker of general psychopathology (termed "p factor")[222]. Thus, universal prevention of all these disorders may, in theory, overlap greatly.

Synergically preventing common and infrequent mental disorders

Refined transdiagnostic preventive approaches could facilitate targeting more prevalent common mental disorders to synergistically prevent the more infrequent psychotic and bipolar disorders, whose incidence may have been progressively declining[198], although not everywhere[223]. Notably, the notion that psychotic symptoms are not infrequent but rather common among young individuals is caused by the trivialization of their contextual significance and operationalization, resulting in non-specificity[224]. For example, surrogate markers, such as psychotic experiences, are frequently conflated with the APS of the CHR-P state[225] (without explaining what makes a symptom truly "psychotic"[225]). Unlike self-assessed psychotic experiences, APS require detection by an experienced and trained clinician to distinguish pathological from non-pathological phenomena[226], and they are neither common features nor distributed continuously in the general population, accounting for only 0.3% of individuals[227].

Overall, 66% of the incidence of clinical psychosis in the population is accounted for by preceding mood disorders[187]. This finding is not new: Conrad's phenomenological clinical-stage model of psychosis onset[228] established early mood dysregulation as the underlying core feature. At the same time, a substantial proportion (37%) of the population-level incidence of psychosis is explained by the CHR-P stage, independently from mood disorders[187]. The majority of CHR-P individuals have comorbid non-psychotic mental disorders (which do not increase the risk for psychosis but tend to persist over time[229]), mostly common mental disorders: 41% depressive disorders and 15% anxiety disorders[72,230]. These findings demonstrate that the CHR-P state is already partially transdiagnostic (some cases of psychosis may originate outside it[198]), potentially capturing a psychosis dimension that emerges from anxiety or depressive disorders.

These considerations may inform the future configuration of preventive health care services. Conventional mental health services are not generally engineered to detect and prevent psychosis onset from anxiety or depressive disorders, as claimed by some authors[195,231]. Young people at risk for psychosis or bipolar disorder typically present with blurred and unspecific symptoms that are too mild to fulfil the entry criteria of conventional mental health services. An alternative approach may be to enhance the transdiagnostic potential of current preventive (e.g., CHR-P) services, implementing the detection of emerging bipolar and depressive (and anxiety) disorders and better integrating them with primary care to facilitate the prevention of physical health burden[3]. Such initiatives are emerging[232].

Furthermore, the needs-based support and the public health campaigns routinely offered by CHR-P services could be expanded to better address the social determinants of psychotic and common mental disorders at the population level[233]. CHR-P services also represent a successful global template for transitional mental health services and applied clinical research[232] that fully integrate between adolescence and young adulthood[67]. This overcomes the historical paediatric-adult bifurcation, in which children and adolescent mental health services are usually cut at the age of 15 or 18 (the transitional period), when young people are most liable to mental disorders. This current two-tier clinical research system is developmentally inappropriate (psychopathology and brain maturation see no abrupt transition among adolescence and early adulthood) to advance preventive psychiatry for young individuals, and leads many of them to fall through cracks.

To overcome these issues, broader youth-friendly mental health services that ensure low-threshold entry into pathways to care are currently advocated, but solid effectiveness evidence is still needed[3], and caution is advised to not over-pathologize the potentially non-specific or transient occur-

rence of common mental health problems in young people[234].

Preventing physical and mental health burden together

Despite the interconnectedness between mental and physical health problems (e.g., several shared risk factors)[42], the severe physical health burden associated with emerging mental disorders in young people is not yet systematically incorporated in preventive approaches. A recent umbrella review of the top-tier evidence has demonstrated that some lifestyle behaviours – such as low levels of physical activity, sleep disturbances, adverse dietary patterns, and tobacco smoking – are associated with an increased risk of psychotic, bipolar and depressive/anxiety disorders[235]. Future research could employ the TRANSD criteria to ascertain the transdiagnostic potential of lifestyle behaviours: for example, poor sleep is associated with bipolar disorder and depression/anxiety but not psychosis, while poor diet is associated with depressive disorders only[235].

Prevention for these risk factors is currently driven by initiatives siloed in other non-communicable disorders, such as cancer and obesity. However, these factors are also common across physical disorders: pursuing physical health and positive lifestyle behaviours is a tantalizing population-level strategy for universal prevention, making sense for concurrently reducing the risk of many other physical diseases[42]. The numerator of cost and risk is thus offset by a denominator of multiple psychiatric and physical disease endpoints[236].

Experience from smoking prevention suggests that similar public health population-level interventions are far more effective than individual-level approaches. However, current preventive capacity is limited[235] (e.g., selective/universal physical exercise may prevent common mental disorders[157,174], but these findings need to be consolidated), and future research should establish the most effective physical health/lifestyle interventions in young people.

Implementing stratified/personalized prognosis

Modern advancements in the field of individualized prediction modelling aim to consolidate stratified (tailored to subgroups) or precision (tailored to the individual subject) preventive psychiatry in young people[237]. Several individualized risk prediction models for forecasting the onset of psychosis, bipolar and depression/anxiety in young people[238] (see Table 5) have been externally validated in terms of prognostic accuracy, which is an essential step to address the extent to which predictions can be generalized to the data from plausibly related settings.

Despite these progresses, prognostic accuracy for most of these models is not sufficient to prove clinical utility and implementability across different scenarios[239]. In fact, a systematic review has found that only about 5% of the total pool of risk prediction models published in psychiatry is externally validated, and that only 0.2% are being considered for implementation (most models may not cross the implementation threshold, as they would not improve outcomes), highlighting a profound replication and translational gap[240]. For example, across all prognostic models reviewed in Table 5, only the transdiagnostic risk calculator has been piloted for real-world implementation in clinical practice[241].

To overcome these limitations, the next generation of research should prioritize further refinements and replications of existing algorithms. Given their complexity, the weighting of the predictors may vary considerably with context (e.g., adolescent vs. young adult, geographic contexts). For those models that may reach higher levels of proof for clinical utility, the implementation pathway is a perilous journey undermined by several obstacles, related to individuals involved (e.g., making their data available or accepting the outputs of the risk calculators), clinicians (e.g., adherence to the recommendations made by prediction models, communicating risks), providers (e.g., confidentiality of data, interpretability of outputs) and funders/organizations (implementing standard prediction procedures)[238].

Implementation science itself is contested and complex, and there is no solid general implementation framework and practical guidance for preventive psychiatry. The next generation of research in this field should develop a coherent and pragmatic implementation framework and associated international infrastructures[177].

The latter necessitate collaborative data sharing efforts and international, large-scale, harmonized and multimodal (e.g., psychopathological, neurobiological, neurocognitive) clinical research databases, integrated with digital technologies (e.g., electronic medical records), as well as specific support from funders and stakeholders[237]. Harmonization is likely to be most successful for future datasets that are prospectively collected. However, efforts should also be made to standardize (to the extent possible) existing datasets that already include large amounts of data.

Establishing evidence-based preventive interventions

Another area of future research is the development of evidence-based preventive interventions to overcome the current divergence between "political" literature, which tends to deliver an overoptimistic message, and evidence-based literature, which emphasizes methodological biases and the inconsistency of the available findings. For example, two independent meta-analyses found no evidence (as opposed to evidence of absence) to favour specific interventions for preventing psychosis in CHR-P individuals[93,94]. Without providing any meta-analytical counter-evidence, some authors have complained that evidence needs to be contextualized, because the "potential for improvement is a key message for patients, families, and practitioners"[242]. The Cochrane authors replied that their meta-analysis was not a criticism of the valuable preventive aims, but only scientific grading of the available evidence[243].

Along these lines, Caplan first noted that, although there was little empirical evidence to support primary prevention and little knowledge of the aetiology of mental disorders, "there appears to be validity to

Table 5 Externally validated, individualized prognostic models for forecasting the onset of psychotic, bipolar (BD), and major depressive (MD)/generalized anxiety (GAD) disorders in young people

	Outcome	Predictors	Development sample size (mean age, location); performance (measure)	External validation sample size (mean age, location); performance (measure)
Cannon et al[291]	Psychosis onset in CHR-P	Age, family history, unusual thoughts and suspiciousness, lower verbal learning and memory performance, slower speed of processing, decline in social functioning	596 (18.5, US); 0.71 (C-index)[291]	176 (16.6, US); 0.79 (AUC)[292] 199 (19.1, China); 0.63 (AUC)[293] 68 (18.59, US); 0.71 (AUC)[294]
Zhang et al[295]	Psychosis onset in CHR-P	Functional decline, positive symptoms, negative symptoms, general symptoms	349 (20.4, China); 0.744 (AUC)[295]	100 (age not available, China); 0.804 (AUC)[295] 68 (18.59, US); 0.65 (AUC)[294]
Fusar-Poli et al[199]	Transdiagnostic psychosis onset in secondary mental health care patients	Age, sex, ethnicity, age by gender, ICD-10 index diagnosis	33,820 (34.4, UK); 0.80 (C-index)[199]	54,716 (32.0, UK); 0.79 (C-index)[199] 13,702 (40.9, UK); 0.73 (C-index)[296] 33,710 (22.7, UK), 0.79 (C-index)[297] 2,430,333 (34.2, US); 0.68 (C-index)[298]
		Refined version including 14 symptoms extracted with natural language processing	28,297 (34.8, UK); 0.86 (C-index)[299]	63,854 (33); 0.85 (C-index)[299]
King et al[300]	MD onset in primary care	Age, sex, country educational status, difficulties in work, history of depression in first-degree relatives, experience of discrimination, lifetime major depression episode, mental quality of life, physical quality of life	5,216 (48.9, UK, Spain, Slovenia, Portugal, The Netherlands); 0.79 (C-index)[300]	1,732 (47.0, Chile); 0.71(C-index)[300] 29,621 (43.8 US); 0.71(AUC)[301]
King et al[302]	GAD and MD onset in primary care	Age, sex, country, difficulties in paid and unpaid work, history of depression in first-degree relatives, follow-up period, lifetime major depression episode, mental quality of life, physical quality of life	4,905 (age not available, UK, Spain, Slovenia, Portugal); 0.75 (C-index)[302]	5,140 (age not available, Netherlands, Estonia, Chile); 0.71-0.81 (C-index)[302] 24,626 (age not available, US); 0.62 (AUC)[303]
Birmhaer et al[304]	Onset of BD-I or BD-II from sub-threshold BD symptoms	Age, sex, mania, depression, anxiety, emotional lability, functioning, duration of BD, ethnicity, family history of BD	140 (11.9, US); 0.71 (AUC)[304]	58 (11.9, US); 0.75 (AUC)[304]
Raket et al[201]	Onset of psychosis (schizophrenia) from primary and secondary care	Demographics and dynamic medical events (diagnoses, prescriptions, procedures, encounters and admissions, observations, and laboratory test results)	102,030 (42, US); 0.856 (AUC)[201]	4,770 (age not available, US); 0.799 (AUC)[201]

CHR-P – clinical high risk for psychosis, AUC – area under the curve

the assumptions"[244] of primary prevention, which ought not to be suspended while awaiting the results of evidence-based medicine. This tension extends beyond the CHR-P paradigm: other evidence-based syntheses have disconfirmed initial promising findings relating to the indicated/selective/universal prevention of anxiety and depression[113,115,156] or reduction of substance abuse in young people[154], and these debates are even more pronounced for public health approaches targeting social determinants of mental disorders. The goal to prevent psychotic, bipolar and common mental disease is noble, but this alone does not justify the use of interventions where there is no demonstrated effectiveness. Preventive breakthroughs that do not show cost-effectiveness (see below) are also unlikely to be implemented in health care systems and in the general population, and this would be for good reasons. Future research is also needed to better customize the effectiveness of preventive interventions to several vulnerable groups such as refugees, prisoners, persons in humanitarian contexts; lesbian, gay, bisexual and transgender persons; persons who are being bullied or exposed to violence, and those who have recently been bereaved.

Future research should also explore methodological innovations. The lack of evidence to favour several preventive interventions[113,115,154,156] may indicate that a one-size-fits-all approach is not effective and obfuscates the efficacy for specific subgroups of individuals. Future individual-participant data level network meta-analyses are under planning[245] and may help deconstructing the effect of different individual- or subgroup- level factors. As new interventions in this field are being tested at a rapid pace, living meta-analyses may be particularly useful to update the emerging evidence. However, subgroup effect claims have a notoriously poor record of validation across medicine[246,247]. Moreover, even if present, they would require very large sample sizes to be able to document and validate them in a rigorous fashion. Even large individual-level meta-analyses may not identify effect modification in most medical interventions[248].

Subgroup effects and intervention effect heterogeneity require rigorous documentation and validation before being adopted[249,250].

Another explanation for the lack of evidence may be that dilution of risk enrichment and infrequent events may have led to reduced statistical power to find a difference between a preventive intervention and a control group (e.g., more than 2,000 CHR-P individuals are needed to detect a 50% reduction in risk to psychosis[251]). Stratification algorithms to control for risk enrichment and inform trial recruitment are under development[252], and harmonization of large-scale datasets within international research consortia is expected to increase the statistical power.

Future meta-analytical approaches could also exclude low-quality studies instead of pooling all available data (which has frequently been advocated[242]), most of which may be of insufficient quality[253-255]. Future interventional studies should investigate the efficacy of emerging preventive compounds (e.g., oxytocin, N-acetylcysteine, cannabinoids), screening procedures (e.g., maternal screening, bipolar risk screening, screening in low/middle-income countries) or refined psychoeducation interventions (e.g., for asymptomatic bipolar familial risk, reduction of alcohol and illicit substance abuse). Innovative adaptive trial designs[256] that integrate with stepped preventive care should also be considered.

Developing an ethical framework for preventive psychiatry

Preventive medicine in young people brings some ethical challenges. For example, the potential cost, inconvenience, social stigma and other harms of a false-positive designation in young people may be high[257]. These concerns are corroborated by lack of valid biomarkers of risk (remarkably, there are no approved biomarkers in all of psychiatry) and adverse effects of antipsychotics[258,259] or other psychotropic agents. Antipsychotics are not recommended for preventing psychosis[68], and these molecules are likely to be inappropriately prescribed to young people at risk outside preventive programmes[260]. Psychological/psychosocial interventions may also be associated with adverse effects. Population interventions to prevent substance abuse in children or common mental disorders in humanitarian settings have been shown to worsen outcomes[261] and to be not more acceptable than the waiting-list condition[120]. Notably, similar ethical issues have been raised in preventive medicine: for example, handling pre-diabetes (intermediate hyperglycaemia) has been challenged for the risk of false positives, as many people do not progress to diabetes[262].

Another question is the extent to which sharing a risk designation with young people and their families may produce harmful stigma (non-maleficence: first do no harm) or offer benefits (beneficence: helping the youth) and autonomy[263]. One perspective is that, in the absence of solid evidence for effectiveness and with potential for harm, preventive services may be seriously questioned. However, evidence shows that stigma is lower in service users than in their health care professionals[264], and caused by the service user's experience of symptoms rather than induced by the clinician's designation[265]. Stigma seems also associated with at-risk features even when no at-risk label is attached[266], and level of stigma in preventive services is comparable to that associated with depression[267]. Thus, sharing an at-risk designation may not only be helpful (beneficence), but honour the ethical principle that young people have the right to receive information relevant to their health (autonomy)[268], in particular given the very real morbidity (e.g., functional impairments of CHR-P individuals[77]), risks (e.g., up to 40% risk of developing persistent psychosis at 2 years for BLIPS[269]), and their active help-seeking behaviours.

A counter-argument is that, if no effective preventive intervention can be provided, then knowing in advance may not be helpful outside clinical monitoring (which can reduce the duration of untreated disorder[90]). However, young people accessing preventive (e.g., CHR-P) services benefit from an integrated package of vocational, psychosocial and familial support interventions which would otherwise not be available to them. Prognostic communication in these services is nuanced, tailoring it to each individual, illustrating the varying outcomes that might be possible (remission/response, persistence, worsening), and that currently there is no certain way to distinguish those possibilities for any given individual[263]. Furthermore, service user groups are actively involved in designing dissemination materials and advising on service delivery[3]. The effectiveness of these approaches needs better study but, in principle, they may help young people endorse the need for several precision/preventive psychiatry concepts, including how testing may lead to tailored interventions[270].

Future collaborative research should set up an ethical framework for implementing preventive psychiatry in young people, involving health care workers, policy makers, service users and their families, and putting emphasis on the subjective experience of the youth[271]. This call would realize the vision of predictive, preventive, personalized and participatory ("P4") psychiatry and help ensure that future progresses occur in an ethically acceptable manner that optimizes benefits and minimizes harms for young people[268].

Improving prevention through education and training

The recent systematic development of science-based prognosis and prevention for young people should be reinforced by a comprehensive educational action targeting several stakeholders. Editors and reviewers of scientific journals should become aware that improving reproducibility standards is needed to maximize the efficiency and trustworthiness of preventive research for young people[237]. Power and other statistical issues need to be considered when interpreting the results of studies.

Another problem is that the current division of medical training leads to often contrasting approaches among adolescent versus adult health care workers, enhancing the cultural divide among the specialities. Innovative curricula could be developed to train new "transitional" health care workers, which could also incorporate core conceptual and methodological issues pertaining to the science of prognosis and preventive in-

terventions (for example, universal/public health approaches may require economic or social understanding that extends beyond medical knowledge[33]).

Future curricula should also rectify the ongoing erosion of psychiatric training on psychopathology and phenomenology, boosted by checklist and algorithmic approaches in the context of pressured health care, to avoid blurring the borders between pathology (e.g., psychotic symptoms) and variants of the normal (e.g., psychotic experiences)[272] in young people.

Furthermore, knowledge and resources in the prevention of mental disorders and mental health promotion in young people are unevenly distributed around the world, and global training initiatives are needed to support countries that are still lacking capacity and expertise. This goal could be achieved through international networks of collaborating research centres[273]. Finally, policy makers should be educated on the achievements and limitations surrounding the prevention of mental disorders in young people, and funders supported to design preventive calls.

Consolidating the cost-effectiveness of preventive psychiatry

Due to high health care cost and impaired ability to work, psychotic, bipolar and common mental disorders in young people lead to a huge economic burden, estimated at US$ 16.3 trillion by 2030 (including neurological and substance use disorders), exceeding cardiovascular disease, chronic respiratory disease, cancer and diabetes, and accounting for more than half of the global economic burden attributable to non-communicable diseases[274]. However, the current global median expenditure on mental health is low (only US$ 2.5 per person annually, accounting for less than 2% of government health expenditure globally[275]), and this may be a major reason for the wide gap between young people's mental health needs and the provision of preventive interventions[275].

Cost-effectiveness of preventive interventions is essential to avoid adding pressure on already overstretched health care budgets. Available evidence indicates that cost-effective preventive interventions may include perinatal screening-plus-intervention programmes[276], stepped care for the prevention of anxiety (but not depression)[277], and prevention of psychosis in CHR-P individuals (savings of US$ 844 per prevented psychosis[278]). However, economic evaluations of preventive approaches in young people remain relatively neglected[279]. Moreover, these cost-effectiveness estimates may also be biased in favour of the tested interventions for various reasons (e.g., involvement of authors who are supportive of the interventions and/or practice them themselves, uncertain and inflated estimates of effectiveness, lack of reasonable estimates on most of the potential harms, difficulty to translate some harms, such as stigma, into quantitative parameters).

Future research should try to remedy some of these shortcomings and address economic evidence gaps in perinatal bipolar or anxiety disorders[275], adolescent mental health[280], interventions in low-resource settings (cost-effectiveness evidence may not be transferable across different countries) and youth mental health services[275]. Future economic preventive studies should also consider a long-term time horizon, in the light of the potentially progressive nature of these disorders and the wider familial and societal impact outside health care. Finally, the interconnectedness of socio-economic, religious, cultural, ethnic inequalities and cost-effectiveness[275] should be better addressed.

Decreasing inequalities to prevent mental disorders

Prevention of mental disorders in young people has not yet solidified as global research or programmatic focus[281]. We have demonstrated above that universal public health approaches targeting the social determinants of mental disorders hold the greatest potential for reducing the risk profile of the whole population. For example, the wide adoption of neo-liberal economic policies and globalization has increased wealth inequality (e.g., in the US, the top 10% of the population averages nine times as much income as the bottom 90%), which is robustly associated with psychotic and depressive disorders[282].

Effective actions may include reducing income inequality, such as progressive taxation policies and a basic universal income, in combination with promotion of good mental health and provision of packages of care with demonstrated effectiveness[283]. However, the effectiveness of poverty alleviation strategies is uncertain and requires further research; conversely, selective or indicated approaches in subgroups with mental health issues have the potential to improve economic outcomes[284].

Future public health research will also require advancements in epidemiologic methods of causal inference, improvements in data quality and availability[233], robust randomized controlled trials to demonstrate specific effectiveness on psychotic, bipolar and mental disorders, qualitative research to customize interventions around context/culture, and mixed-method implementation science to assess the scaling up of interventions[173].

Future public health approaches require committed and sustained efforts to address a range of other barriers, a strong health sector responsibility, and a vigorous leadership role in bringing society-wide attention and cross-government actions together. This point is particularly critical, given the experience of smoking prevention, whose success was predicated on successive hard-fought public policy battles.

Governments should tackle unacceptable inequalities in young people's mental health[285], and invest on improving the social determinants of their mental health: education, employment, social care, housing, criminal justice, poverty alleviation, social security/welfare benefits, community development, and immigration[275]. These inequalities are likely to increase with the current COVID-19 pandemic, which will force to change mental health services, focusing even more on flexible systems that include prevention[286].

Addressing these inequalities should be the shared responsibility of professionals across systems of care, representing the fundamental pillar of an individualized approach to youth mental health[233]. Such

primordial-like type of prevention argues for universal health care coverage and parity between physical and mental health[275]. It is hoped that more progress in this direction can be achieved in the decade to come, as much is at stake for young people at risk for and with emerging psychotic, bipolar and common mental disorders.

ACKNOWLEDGEMENTS

P. Fusar-Poli is supported by the European Union Seventh Framework Program (PSYSCAN project), US National Institute of Mental Health (ProNET project) and Wellcome Trust (STEP project). M. Berk is supported by an Australian National Health and Medical Research Council (NHMRC) Senior Principal Research Fellowship (grant no. 1156072). C. Arango is funded by the Spanish Ministry of Science and Innovation (grant nos. SAM16PE07CP1, PI16/02012, PI19/024), co-financed by European Regional Development Funds from the European Commission (grant no. B2017/BMD-3740 AGES-CM-2), European Union Seventh Framework Program (EU-GEI, PSYSCAN and METSY projects); and European Union H2020 Program (PRISM and AIMS-2-TRIALS projects).

REFERENCES

1. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. 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. Kessler RC, Berglund P, Demler O et al. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005; 62:593-602.
3. Fusar-Poli P. Integrated mental health services for the developmental period (0 to 25 years): a critical review of the evidence. Front Psychiatry 2019;10:355.
4. United Nations. World population prospects 2019. https://population.un.org/wpp/DataQuery/.
5. Insel TR, Fenton WS. Psychiatric epidemiology: it's not just about counting anymore. Arch Gen Psychiatry 2005;62:590-2.
6. McGorry P. Building the momentum and blueprint for reform in youth mental health. Lancet Psychiatry 2019;6:459-61.
7. Millan MJ, Andrieux A, Bartzokis G et al. Altering the course of schizophrenia: progress and perspectives. Nat Rev Drug Discov 2016;15:485-515.
8. Wang PS, Angermeyer M, Borges G et al. Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007;6:177-85.
9. Vancampfort D, Wampers M, Mitchell AJ et al. A meta-analysis of cardio-metabolic abnormalities in drug naive, first-episode and multi-episode patients with schizophrenia versus general population controls. World Psychiatry 2013;12:240-50.
10. Liu NH, Daumit GL, Dua T et al. Excess mortality in persons with severe mental disorders: a multilevel intervention framework and priorities for clinical practice, policy and research agendas. World Psychiatry 2017;16:30-40.
11. Vancampfort D, Firth J, Schuch FB et al. Sedentary behavior and physical activity levels in people with schizophrenia, bipolar disorder and major depressive disorder: a global systematic review and meta-analysis. World Psychiatry 2017;16:308-15.
12. Correll CU, Solmi M, Veronese N et al. Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. World Psychiatry 2017;16:163-80.
13. Mojtabai R, Stuart EA, Hwang I et al. Long-term effects of mental disorders on educational attainment in the National Comorbidity Survey ten-year follow-up. Soc Psychiatry Psychiatr Epidemiol 2015;50:1577-91.
14. Stuart H. Fighting the stigma caused by mental disorders: past perspectives, present activities, and future directions. World Psychiatry 2008;7:185-8.
15. Puras D, Gooding P. Mental health and human rights in the 21st century. World Psychiatry 2019; 18:42-3.
16. Gerlinger G, Hauser M, De Hert M et al. Personal stigma in schizophrenia spectrum disorders: a systematic review of prevalence rates, correlates, impact and interventions. World Psychiatry 2013;12:155-64.
17. Wasserman D, Cheng Q, Jiang GX. Global suicide rates among young people aged 15-19. World Psychiatry 2005;4:114-20.
18. Chesney E, Goodwin GM, Fazel S. Risks of all-cause and suicide mortality in mental disorders: a meta-review. World Psychiatry 2014;13:153-60.
19. Christensen H, Reynolds CFR, Cuijpers P. Protecting youth mental health, protecting our future. World Psychiatry 2017;16:327-8.
20. Sommer IE, Arango C. Moving interventions from after to before diagnosis. World Psychiatry 2017;16:275-6.
21. Arango C. Psychiatrists are experts when it comes to missing boats. Will prevention be the next one? Eur Arch Psychiatry Clin Neurosci 2019;269:869-70.
22. Robinson G. On the prevention and treatment of mental disorders. London: Longman, Brown, Green, Longmans & Roberts, 1859.
23. Parry M. From a patient's perspective: Clifford Whittingham Beers' work to reform mental health services. Am J Public Health 2010;100: 2356-7.
24. Anonymous. Mental hygiene in childhood. BMJ 1904;2:1414-5.
25. Emerson CP. Mental hygiene and its relation to public health. Am J Public Health 1925;15:1062-6.
26. White WA. Preventive principles in the field of mental medicine. J Am Public Health Assoc 1911; 1:82-9.
27. Leavell H, Clark E. Preventive medicine for the doctor in his community: an epidemiologic approach, 1st ed. New York: McGraw-Hill, 1958.
28. Clark EG. Natural history of syphilis and levels of prevention. Br J Vener Dis 1954;30:191-7.
29. Caplan G, Grunebaum H. Perspectives on primary prevention. A review. Arch Gen Psychiatry 1967;17:331-46.
30. Strasser T. Reflections on cardiovascular diseases. Interdiscip Sci Rev 1978;3:225-30.
31. Hofler M. The Bradford Hill considerations on causality: a counterfactual perspective. Emerg Themes Epidemiol 2005;2:11.
32. Lucas RM, McMichael AJ. Association or causation: evaluating links between "environment and disease". Bull World Health Organ 2005;83:792-5.
33. Gordon RS Jr. An operational classification of disease prevention. Public Health Rep 1983;98: 107-9.
34. US Institute of Medicine. Reducing risks for mental disorders: frontiers for preventive intervention research. Washington: National Academies Press, 1994.
35. Kessler RC, Price RH. Primary prevention of secondary disorders: a proposal and agenda. Am J Community Psychol 1993;21:607-33.
36. World Health Organization. Prevention of mental disorders. Effective interventions and policy options. Geneva: Department of Mental Health and Substance Abuse, 2004.
37. Correll CU, Galling B, Pawar A et al. Comparison of early intervention services vs treatment as usual for early-phase psychosis: a systematic review, meta-analysis, and meta-regression. JAMA Psychiatry 2018;75:555-65.
38. Burkhardt E, Pfennig A, Breitling G et al. Creativity in persons at-risk for bipolar disorder – A pilot study. Early Interv Psychiatry 2019;13:1165-72.
39. Parnas J, Sandsten KE, Vestergaard CH et al. Mental illness among relatives of successful academics: implications for psychopathology-creativity research. World Psychiatry 2019;18:362-3.
40. Keller MC, Visscher PM. Genetic variation links creativity to psychiatric disorders. Nat Neurosci 2015; 18:928-9.
41. Fusar-Poli P, Salazar de Pablo G, De Micheli A et al. What is good mental health? A scoping review. Eur Neuropsychopharmacol 2020;31:33-46.
42. Jacka FN, Mykletun A, Berk M. Moving towards a population health approach to the primary prevention of common mental disorders. BMC Med 2012;10:149.
43. Kessler RC, Amminger GP, Aguilar-Gaxiola S et al. Age of onset of mental disorders: a review of recent literature. Curr Opin Psychiatry 2007; 20:359-64.
44. Radua J, Ramella-Cravaro V, Ioannidis JPA et al. What causes psychosis? An umbrella review of risk and protective factors. World Psychiatry 2018;17:49-66.
45. Joslyn C, Hawes DJ, Hunt C et al. Is age of onset associated with severity, prognosis, and clinical features in bipolar disorder? A meta-analytic review. Bipolar Disord 2016;18:389-403.
46. Kessler RC, Bromet EJ. The epidemiology of depression across cultures. Annu Rev Public Health 2013;34:119-38.
47. Jones PB. Adult mental health disorders and their age at onset. Br J Psychiatry 2013;202(Suppl. 54):s5-10.
48. Uher R, Zwicker A. Etiology in psychiatry: embracing the reality of poly-gene-environmental causation of mental illness. World Psychiatry 2017;16:121-9.
49. Paus T, Keshavan M, Giedd JN. Why do many psychiatric disorders emerge during adolescence? Nat Rev Neurosci 2008;9:947-57.
50. Kloiber S, Rosenblat JD, Husain MI et al. Neurodevelopmental pathways in bipolar disorder. Neurosci Biobehav Rev 2020;112:213-26.
51. Parellada M, Gomez-Vallejo S, Burdeus M et al.

Developmental differences between schizophrenia and bipolar disorder. Schizophr Bull 2017;43:1176-89.
52. Hagan CC, Graham JM, Wilkinson PO et al. Neurodevelopment and ages of onset in depressive disorders. Lancet Psychiatry 2015;2:1112-6.
53. Fava GA, Kellner R. Staging: a neglected dimension in psychiatric classification. Acta Psychiatr Scand 1993;87:225-30.
54. McGorry PD, Hickie IB, Yung AR et al. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. Aust N Z J Psychiatry 2006;40:616-22.
55. Fusar-Poli P, McGorry PD, Kane JM. Improving outcomes of first-episode psychosis: an overview. World Psychiatry 2017;16:251-65.
56. Berk M, Post R, Ratheesh A et al. Staging in bipolar disorder: from theoretical framework to clinical utility. World Psychiatry 2017;16:236-44.
57. Salagre E, Dodd S, Aedo A et al. Toward precision psychiatry in bipolar disorder: staging 2.0. Front Psychiatry 2018;9:641.
58. Kapczinski F, Magalhaes PV, Balanza-Martinez V et al. Staging systems in bipolar disorder: an International Society for Bipolar Disorders task force report. Acta Psychiatr Scand 2014;130:354-63.
59. Berk M, Berk L, Dodd S et al. Stage managing bipolar disorder. Bipolar Disord 2014;16:471-7.
60. Verduijn J, Milaneschi Y, van Hemert AM et al. Clinical staging of major depressive disorder: an empirical exploration. J Clin Psychiatry 2015;76:1200-8.
61. Reneses B, Aguera-Ortiz L, Sevilla-Llewellyn-Jones J et al. Staging of depressive disorders: relevance of resistance to treatment and residual symptoms. J Psychiatr Res 2020;129:234-40.
62. Bokma WA, Batelaan NM, Hoogendoorn AW et al. A clinical staging approach to improving diagnostics in anxiety disorders: is it the way to go? Aust N Z J Psychiatry 2020;54:173-84.
63. Clarke PJ, Hickie IB, Scott E et al. Clinical staging model applied to young people presenting with social anxiety. Early Interv Psychiatry 2012;6:256-64.
64. Fontenelle LF, Yucel M. A clinical staging model for obsessive-compulsive disorder: is it ready for prime time? EClinicalMedicine 2019;7:65-72.
65. Voshaar RC, Beekman AT, Pachana N. Clinical staging and profiling of late-life anxiety disorders; the need for collaboration and a life-span perspective. Int Psychogeriatr 2015;27:1057-9.
66. Yung AR, Yuen HP, McGorry PD et al. Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States. Aust N Z J Psychiatry 2005;39:964-71.
67. Kotlicka-Antczak M, Podgorski M, Oliver D et al. Worldwide implementation of clinical services for the prevention of psychosis: the IEPA early intervention in mental health survey. Early Interv Psychiatry (in press).
68. National Institute for Health and Care Excellence. Psychosis and schizophrenia in adults: prevention and management. London: National Institute for Health and Care Excellence, 2014.
69. Schmidt SJ, Schultze-Lutter F, Schimmelmann BG et al. EPA guidance on the early intervention in clinical high risk states of psychoses. Eur Psychiatry 2015;30:388-404.
70. Fusar-Poli P, Carpenter WT, Woods SW et al. Attenuated psychosis syndrome: ready for DSM-5.1? Annu Rev Clin Psychol 2014;10:155-92.
71. Arango C, Bernardo M, Bonet P et al. When the healthcare does not follow the evidence: the case of the lack of early intervention programs for psychosis in Spain. Rev Psiquiatr Salud Ment 2017;10:78-86.
72. Fusar-Poli P, Salazar de Pablo G, Correll C et al. Prevention of psychosis: advances in detection, prognosis and intervention. JAMA Psychiatry 2020;77:755-65.
73. Fusar-Poli P. The Clinical High-Risk State for Psychosis (CHR-P), Version II. Schizophr Bull 2017;43:44-7.
74. Fusar-Poli P, Cappucciati M, Rutigliano G et al. At risk or not at risk? A meta-analysis of the prognostic accuracy of psychometric interviews for psychosis prediction. World Psychiatry 2015; 14:322-32.
75. Oliver D, Reilly T, Baccaredda Boy O et al. What causes the onset of psychosis in individuals at clinical high risk? A meta-analysis of risk and protective factors. Schizophr Bull (in press).
76. Fusar-Poli P, Tantardini M, De Simone S et al. Deconstructing vulnerability for psychosis: meta-analysis of environmental risk factors for psychosis in subjects at ultra high-risk. Eur Psychiatry 2017;40:65-75.
77. Fusar-Poli P, Rocchetti M, Sardella A et al. Disorder, not just state of risk: meta-analysis of functioning and quality of life in people at high risk of psychosis. Br J Psychiatry 2015;207:198-206.
78. Fusar-Poli P, Borgwardt S, Bechdolf A et al. The psychosis high-risk state: a comprehensive state-of-the-art review. JAMA Psychiatry 2013;70:107-20.
79. Falkenberg I, Valmaggia L, Byrnes M et al. Why are help-seeking subjects at ultra-high risk for psychosis help-seeking? Psychiatry Res 2015;228: 808-15.
80. Fusar-Poli P, Estrade A, Spencer TJ et al. Pan-London Network for Psychosis-Prevention (PNP). Front Psychiatry 2019;10:707.
81. Fusar-Poli P, Byrne M, Badger S et al. Outreach and support in south London (OASIS), 2001-2011: ten years of early diagnosis and treatment for young individuals at high clinical risk for psychosis. Eur Psychiatry 2013;28:315-26.
82. Fusar-Poli P, Schultze-Lutter F, Cappucciati M et al. The dark side of the moon: meta-analytical impact of recruitment strategies on risk enrichment in the clinical high risk state for psychosis. Schizophr Bull 2016;42:732-43.
83. Addington J, Stowkowy J, Weiser M. Screening tools for clinical high risk for psychosis. Early Interv Psychiatry 2015;9:345-56.
84. Fusar-Poli P. Why ultra high risk criteria for psychosis prediction do not work well outside clinical samples and what to do about it. World Psychiatry 2017;16:212-3.
85. Webb JR, Addington J, Perkins DO et al. Specificity of incident diagnostic outcomes in patients at clinical high risk for psychosis. Schizophr Bull 2015;41:1066-75.
86. Fusar-Poli P, Rutigliano G, Stahl D et al. Long-term validity of the At Risk Mental State (ARMS) for predicting psychotic and non-psychotic mental disorders. Eur Psychiatry 2017;42:49-54.
87. Fusar-Poli P, Cappucciati M, De Micheli A et al. Diagnostic and prognostic significance of brief limited intermittent psychotic symptoms (BLIPS) in individuals at ultra high risk. Schizophr Bull 2017;43:48-56.
88. Fusar-Poli P, Cappucciati M, Borgwardt S et al. Heterogeneity of psychosis risk within individuals at clinical high risk: a meta-analytical stratification. JAMA Psychiatry 2016;73:113-20.
89. Fusar-Poli P, De Micheli A, Patel R et al. Real-world clinical outcomes two years after transition to psychosis in individuals at clinical high risk: electronic health record cohort study. Schizophr Bull 2020;46:1114-25.
90. Oliver D, Davies C, Crossland G et al. Can we reduce the duration of untreated psychosis? A systematic review and meta-analysis of controlled interventional studies. Schizophr Bull 2018;44:1362-72.
91. Stafford MR, Jackson H, Mayo-Wilson E et al. Early interventions to prevent psychosis: systematic review and meta-analysis. BMJ 2013; 346:f185.
92. Stafford MR, Jackson H, Mayo-Wilson E et al. Errata: Early interventions to prevent psychosis: systematic review and meta-analysis. BMJ 2013;346:f762.
93. Davies C, Cipriani A, Ioannidis JPA et al. Lack of evidence to favor specific preventive interventions in psychosis: a network meta-analysis. World Psychiatry 2018;17:196-209.
94. Bosnjak Kuharic D, Kekin I, Hew J et al. Interventions for prodromal stage of psychosis. Cochrane Database Syst Rev 2019;11:CD012236.
95. Devoe DJ, Farris MS, Townes P et al. Interventions and transition in youth at risk of psychosis: a systematic review and meta-analyses. J Clin Psychiatry 2020;81:17r12053.
96. Bechdolf A, Ratheesh A, Wood SJ et al. Rationale and first results of developing at-risk (prodromal) criteria for bipolar disorder. Curr Pharm Des 2012;18:358-75.
97. Bechdolf A, Nelson B, Cotton SM et al. A preliminary evaluation of the validity of at-risk criteria for bipolar disorders in help-seeking adolescents and young adults. J Affect Disord 2010;127:316-20.
98. Van Meter AR, Burke C, Youngstrom EA et al. The bipolar prodrome: meta-analysis of symptom prevalence prior to initial or recurrent mood episodes. J Am Acad Child Adolesc Psychiatry 2016;55: 543-55.
99. Vieta E, Salagre E, Grande I et al. Early intervention in bipolar disorder. Am J Psychiatry 2018 2018;175:411-26.
100. Faedda GL, Baldessarini RJ, Marangoni C et al. An International Society of Bipolar Disorders task force report: Precursors and prodromes of bipolar disorder. Bipolar Disord 2019;21:720-40.
101. Benarous X, Consoli A, Milhiet V et al. Early interventions for youths at high risk for bipolar disorder: a developmental approach. Eur Child Adolesc Psychiatry 2016;25:217-33.
102. Post RM, Goldstein BI, Birmaher B et al. Toward prevention of bipolar disorder in at-risk children: potential strategies ahead of the data. J Affect Disord 2020;272:508-20.
103. Van Meter A, Guinart D, Bashir A et al. Bipolar Prodrome Symptom Scale - abbreviated screen for patients: description and validation. J Affect Disord 2019;249:357-65.
104. Correll CU, Olvet DM, Auther AM et al. The Bipolar Prodrome Symptom Interview and Scale-Prospective (BPSS-P): description and validation in a psychiatric sample and healthy controls. Bipolar Disord 2014;16:505-22.
105. Fusar-Poli P, De Micheli A, Rocchetti M et al. Semistructured Interview for Bipolar At Risk States (SIBARS). Psychiatry Res 2018;264:302-9.
106. Miklowitz DJ, Schneck CD, Walshaw PD et al. Effects of family-focused therapy vs enhanced

usual care for symptomatic youths at high risk for bipolar disorder: a randomized clinical trial. JAMA Psychiatry 2020;77:455-63.
107. Leopold K, Bauer M, Bechdolf A et al. Efficacy of cognitive-behavioral group therapy in patients at risk for serious mental illness presenting with subthreshold bipolar symptoms: results from a prespecified interim analysis of a multicenter, randomized, controlled study. Bipolar Disord 2020;22:517-29.
108. Davey CG, McGorry PD. Early intervention for depression in young people: a blind spot in mental health care. Lancet Psychiatry 2019;6:267-72.
109. Munoz RF, Beardslee WR, Leykin Y. Major depression can be prevented. Am Psychol 2012; 67:285-95.
110. Ebert DD, Cuijpers P. It is time to invest in the prevention of depression. JAMA Netw Open 2018;1:e180335.
111. Hoare E, Callaly E, Berk M. Can depression be prevented? If so, how? JAMA Psychiatry (in press).
112. Ratheesh A, Davey C, Hetrick S et al. A systematic review and meta-analysis of prospective transition from major depression to bipolar disorder. Acta Psychiatr Scand 2017;135:273-84.
113. Breedvelt JJF, Kandola A, Kousoulis AA et al. What are the effects of preventative interventions on major depressive disorder (MDD) in young adults? A systematic review and meta-analysis of randomized controlled trials. J Affect Disord 2018;239:18-29.
114. Cuijpers P, van Straten A, Smit F et al. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. Am J Psychiatry 2008;165:1272-80.
115. Moreno-Peral P, Conejo-Ceron S, Rubio-Valera M et al. Effectiveness of psychological and/or educational interventions in the prevention of anxiety: a systematic review, meta-analysis, and meta-regression. JAMA Psychiatry 2017;74:1021-9.
116. Hoare E, Thorp A, Bartholomeusz-Raymond N et al. Opportunities in the Australian national education initiative for promoting mental health in schools. Lancet Child Adolesc Health 2020;4:11-3.
117. Cuijpers P, Koole SL, van Dijke A et al. Psychotherapy for subclinical depression: meta-analysis. Br J Psychiatry 2014;205:268-74.
118. Buntrock C, Ebert DD, Lehr D et al. Effect of a web-based guided self-help intervention for prevention of major depression in adults with subthreshold depression: a randomized clinical trial. JAMA 2016;315:1854-63.
119. Garrido S, Millington C, Cheers D et al. What works and what doesn't work? A systematic review of digital mental health interventions for depression and anxiety in young people. Front Psychiatry 2019;10:759.
120. Papola D, Purgato M, Gastaldon C et al. Psychological and social interventions for the prevention of mental disorders in people living in low- and middle-income countries affected by humanitarian crises. Cochrane Database Syst Rev 2020;9:CD012417.
121. Rasic D, Hajek T, Alda M et al. Risk of mental illness in offspring of parents with schizophrenia, bipolar disorder, and major depressive disorder: a meta-analysis of family high-risk studies. Schizophr Bull 2014;40:28-38.
122. Lawrence PJ, Murayama K, Creswell C. Systematic review and meta-analysis: anxiety and depressive disorders in offspring of parents with anxiety disorders. J Am Acad Child Adolesc Psychiatry 2019;58:46-60.
123. Pepper EJ, Pathmanathan S, McIlrae S et al. Associations between risk factors for schizophrenia and concordance in four monozygotic twin samples. Am J Med Genet B Neuropsychiatr Genet 2018;177:503-10.
124. Barnett JH, Smoller JW. The genetics of bipolar disorder. Neuroscience 2009;164:331-43.
125. Hettema JM, Neale MC, Kendler KS. A review and meta-analysis of the genetic epidemiology of anxiety disorders. Am J Psychiatry 2001; 158:1568-78.
126. McGuffin P, Katz R, Watkins S et al. A hospital-based twin register of the heritability of DSM-IV unipolar depression. Arch Gen Psychiatry 1996; 53:129-36.
127. Lo LE, Kaur R, Meiser B et al. Risk of schizophrenia in relatives of individuals affected by schizophrenia: a meta-analysis. Psychiatry Res 2020;286:112852.
128. Chen MH, Hsu JW, Huang KL et al. Risk and coaggregation of major psychiatric disorders among first-degree relatives of patients with bipolar disorder: a nationwide population-based study. Psychol Med 2019;49:2397-404.
129. Gottschalk MG, Domschke K. Genetics of generalized anxiety disorder and related traits. Dialogues Clin Neurosci 2017;19:159-68.
130. Wilde A, Chan HN, Rahman B et al. A meta-analysis of the risk of major affective disorder in relatives of individuals affected by major depressive disorder or bipolar disorder. J Affect Disord 2014;158:37-47.
131. Agerbo E, Sullivan PF, Vilhjalmsson BJ et al. Polygenic risk score, parental socioeconomic status, family history of psychiatric disorders, and the risk for schizophrenia: a Danish population-based study and meta-analysis. JAMA Psychiatry 2015;72:635-41.
132. Schneider M, Debbane M, Bassett AS et al. Psychiatric disorders from childhood to adulthood in 22q11.2 deletion syndrome: results from the International Consortium on Brain and Behavior in 22q11.2 Deletion Syndrome. Am J Psychiatry 2014;171:627-39.
133. Seidman LJ, Nordentoft M. New targets for prevention of schizophrenia: is it time for interventions in the premorbid phase? Schizophr Bull 2015;41:795-800.
134. Johnstone EC, Ebmeier KP, Miller P et al. Predicting schizophrenia: findings from the Edinburgh High-Risk Study. Br J Psychiatry 2005;186:18-25.
135. Schneider M, Armando M, Pontillo M et al. Ultra high risk status and transition to psychosis in 22q11.2 deletion syndrome. World Psychiatry 2016;15:259-65.
136. Loechner J, Starman K, Galuschka K et al. Preventing depression in the offspring of parents with depression: a systematic review and meta-analysis of randomized controlled trials. Clin Psychol Rev 2018;60:1-14.
137. Al-Sawafi A, Lovell K, Renwick L et al. Psychosocial family interventions for relatives of people living with psychotic disorders in the Arab world: systematic review. BMC Psychiatry 2020;20:413.
138. Perich T, Mitchell PB. Psychological interventions for young people at risk for bipolar disorder: a systematic review. J Affect Disord 2019;252:84-91.
139. Fusar-Poli P, Meyer-Lindenberg A. Forty years of structural imaging in psychosis: promises and truth. Acta Psychiatr Scand 2016;134:207-24.
140. Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? Mol Psychiatry 2012;17:1174-9.
141. Parnas J, Sass LA, Zahavi D. Rediscovering psychopathology: the epistemology and phenomenology of the psychiatric object. Schizophr Bull 2013;39:270-7.
142. Ioannidis JP, Munafo MR, Fusar-Poli P et al. Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention. Trends Cogn Sci 2014;18:235-41.
143. Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q 2016;94:485-514.
144. Bortolato B, Kohler CA, Evangelou E et al. Systematic assessment of environmental risk factors for bipolar disorder: an umbrella review of systematic reviews and meta-analyses. Bipolar Disord 2017;19:84-96.
145. Kohler CA, Evangelou E, Stubbs B et al. Mapping risk factors for depression across the lifespan: an umbrella review of evidence from meta-analyses and Mendelian randomization studies. J Psychiatr Res 2018;103:189-207.
146. Fullana MA, Tortella-Feliu M, Fernandez de la Cruz L et al. Risk and protective factors for anxiety and obsessive-compulsive disorders: an umbrella review of systematic reviews and meta-analyses. Psychol Med 2020;50:1300-15.
147. Carvalho AF, Solmi M, Sanches M et al. Evidence-based umbrella review of 162 peripheral biomarkers for major mental disorders. Transl Psychiatry 2020;10:152.
148. Berk M, Woods RL, Nelson MR et al. Effect of aspirin vs placebo on the prevention of depression in older people: a randomized clinical trial. JAMA Psychiatry 2020;77:1-9.
149. Sanders KM, Stuart AL, Williamson EJ et al. Annual high-dose vitamin D3 and mental well-being: randomised controlled trial. Br J Psychiatry 2011;198:357-64.
150. Davies C, Segre G, Estrade A et al. Prenatal and perinatal risk and protective factors for psychosis: a systematic review and meta-analysis. Lancet Psychiatry 2020;7:399-410.
151. Jones I, Chandra PS, Dazzan P et al. Bipolar disorder, affective psychosis, and schizophrenia in pregnancy and the post-partum period. Lancet 2014;384:1789-99.
152. Morrell CJ, Sutcliffe P, Booth A et al. A systematic review, evidence synthesis and meta-analysis of quantitative and qualitative studies evaluating the clinical effectiveness, the cost-effectiveness, safety and acceptability of interventions to prevent postnatal depression. Health Technol Assess 2016;20:1-414.
153. Keynejad RC, Hanlon C, Howard LM. Psychological interventions for common mental disorders in women experiencing intimate partner violence in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Psychiatry 2020;7:173-90.
154. O'Connor E, Thomas R, Senger CA et al. Interventions to prevent illicit and nonmedical drug use in children, adolescents, and young adults: updated evidence report and systematic review for the US Preventive Services Task Force. JAMA 2020;323:2067-79.
155. Salazar de Pablo G, De Micheli A, Solmi M et al. Universal and selective interventions to prevent poor mental health outcomes in young people:

systematic review and meta-analysis. Submitted for publication.
156. Caldwell DM, Davies SR, Hetrick SE et al. School-based interventions to prevent anxiety and depression in children and young people: a systematic review and network meta-analysis. Lancet Psychiatry 2019;6:1011-20.
157. Hu MX, Turner D, Generaal E et al. Exercise interventions for the prevention of depression: a systematic review of meta-analyses. BMC Public Health 2020;20:1255.
158. Szoke A, Kirkbride JB, Schurhoff F. Universal prevention of schizophrenia and surrogate endpoints at population level. Soc Psychiatry Psychiatr Epidemiol 2014;49:1347-51.
159. van Os J, Reininghaus U. Psychosis as a transdiagnostic and extended phenotype in the general population. World Psychiatry 2016;15:118-24.
160. Sullivan SA, Kounali D, Cannon M et al. A population-based cohort study examining the incidence and impact of psychotic experiences from childhood to adulthood, and prediction of psychotic disorder. Am J Psychiatry 2020;177:308-17.
161. Karcher NR, Barch DM, Avenevoli S et al. Assessment of the Prodromal Questionnaire-Brief Child Version for measurement of self-reported psychoticlike experiences in childhood. JAMA Psychiatry 2018;75:853-61.
162. Yung AR, Lin A. Psychotic experiences and their significance. World Psychiatry 2016;15:130-1.
163. Ferro MA. The psychometric properties of the Kessler Psychological Distress Scale (K6) in an epidemiological sample of Canadian youth. Can J Psychiatry 2019;64:647-57.
164. Ross RG, Hunter SK, McCarthy L et al. Perinatal choline effects on neonatal pathophysiology related to later schizophrenia risk. Am J Psychiatry 2013;170:290-8.
165. Ross RG, Hunter SK, Hoffman MC et al. Perinatal phosphatidylcholine supplementation and early childhood behavior problems: evidence for CHRNA7 moderation. Am J Psychiatry 2016; 173:509-16.
166. Li Y, Freedman R. Prospects for improving future mental health of children through prenatal maternal micronutrient supplementation in China. Pediatr Investig 2020;4:118-26.
167. Catena A, Munoz-Machicao JA, Torres-Espinola FJ et al. Folate and long-chain polyunsaturated fatty acid supplementation during pregnancy has long-term effects on the attention system of 8.5-y-old offspring: a randomized controlled trial. Am J Clin Nutr 2016;103:115-27.
168. McGrath J, Saari K, Hakko H et al. Vitamin D supplementation during the first year of life and risk of schizophrenia: a Finnish birth cohort study. Schizophr Res 2004;67:237-45.
169. Fujita Y, Fujita A, Ishima T et al. Dietary intake of glucoraphanin during pregnancy and lactation prevents the behavioral abnormalities in the offspring after maternal immune activation. Neuropsychopharmacol Rep 2020;40:268-274.
170. Codagnone MG, Spichak S, O'Mahony SM et al. Programming bugs: microbiota and the developmental origins of brain health and disease. Biol Psychiatry 2019;85:150-63.
171. Shinde S, Weiss HA, Varghese B et al. Promoting school climate and health outcomes with the SEHER multi-component secondary school intervention in Bihar, India: a cluster-randomised controlled trial. Lancet 2018;392:2465-77.
172. Shinde S, Weiss HA, Khandeparkar P et al. A multicomponent secondary school health promotion intervention and adolescent health: an extension of the SEHER cluster randomised controlled trial in Bihar, India. PLoS Med 2020; 17:e1003021.
173. Lund C, Brooke-Sumner C, Baingana F et al. Social determinants of mental disorders and the Sustainable Development Goals: a systematic review of reviews. Lancet Psychiatry 2018;5:357-69.
174. Schuch FB, Stubbs B, Meyer J et al. Physical activity protects from incident anxiety: a meta-analysis of prospective cohort studies. Depress Anxiety 2019;36:846-58.
175. Bueno-Antequera J, Munguia-Izquierdo D. Exercise and schizophrenia. Adv Exp Med Biol 2020; 1228:317-32.
176. Melo MC, Daher Ede F, Albuquerque SG et al. Exercise in bipolar patients: a systematic review. J Affect Disord 2016;198:32-8.
177. Fusar-Poli P, Bauer M, Borgwardt S et al. European College of Neuropsychopharmacology Network on the Prevention of Mental Disorders and Mental Health Promotion (ECNP PMD-MHP). Eur Neuropsychopharmacol 2019;29:1301-11.
178. Strunk CM, King KA, Vidourek RA et al. Effectiveness of the surviving the Teens® suicide prevention and depression awareness program: an impact evaluation utilizing a comparison group. Health Educ Behav 2014;41:605-13.
179. Witvliet M, van Lier PA, Cuijpers P et al. Testing links between childhood positive peer relations and externalizing outcomes through a randomized controlled intervention study. J Consult Clin Psychol 2009;77:905-15.
180. World Health Organization. Prevention and promotion in mental health. Geneva: World Health Organization, 2002.
181. Collins S, Woolfson L, Durkin K. Effects on coping skills and anxiety of a universal school-based mental health intervention delivered in Scottish primary schools. Sch Psychol Int 2014;35:85-100.
182. Sharp P, Caperchione C. The effects of a pedometer-based intervention on first-year university students: a randomized control trial. J Am Coll Health 2016;64:630-8.
183. Pendry P, Carr AM, Smith AN et al. Improving adolescent social competence and behavior: a randomized trial of an 11-week equine facilitated learning prevention program. J Prim Prev 2014;35:281-93.
184. Rousseau C, Benoit M, Lacroix L et al. Evaluation of a sandplay program for preschoolers in a multiethnic neighborhood. J Child Psychol Psychiatry 2009;50:743-50.
185. Salazar de Pablo G, De Micheli A, Nieman D et al. Universal and selective interventions to promote good mental health in young people: systematic review and meta-analysis. Eur Neuropsychopharmacol 2020;41:28-39.
186. Solanes A, Albajes-Eizagirre A, Fullana M et al. Can we increase the subjective well-being of the general population? An umbrella review of the evidence. Rev Psiquiatr Salud Ment 2021;14:50-64.
187. Guloksuz S, Pries LK, Ten Have M et al. Association of preceding psychosis risk states and non-psychotic mental disorders with incidence of clinical psychosis in the general population: a prospective study in the NEMESIS-2 cohort. World Psychiatry 2020;19:199-205.
188. Yung AR, Wood SJ, Malla A et al. The reality of at risk mental state services: a response to recent criticisms. Psychol Med 2021;51:212-8.
189. McHugh MJ, McGorry PD, Yuen HP et al. The Ultra-High-Risk for psychosis groups: evidence to maintain the status quo. Schizophr Res 2018;195:543-8.
190. Schultze-Lutter F, Klosterkotter J, Gaebel W et al. Psychosis-risk criteria in the general population: frequent misinterpretations and current evidence. World Psychiatry 2018;17:107-8.
191. Perez J, Jones PB. Breaking the web: life beyond the at-risk mental state for psychosis. Psychol Med (in press).
192. Raballo A, Poletti M, Carpenter WT. Rethinking the psychosis threshold in clinical high risk. Schizophr Bull 2019;45:1-2.
193. Moritz S, Gaweda L, Heinz A et al. Four reasons why early detection centers for psychosis should be renamed and their treatment targets reconsidered: we should not catastrophize a future we can neither reliably predict nor change. Psychol Med 2019;49:2134-40.
194. McGorry PD, Hartmann JA, Spooner R et al. Beyond the "at risk mental state" concept: transitioning to transdiagnostic psychiatry. World Psychiatry 2018;17:133-42.
195. Ajnakina O, David AS, Murray RM. 'At risk mental state' clinics for psychosis – an idea whose time has come – and gone! Psychol Med 2019;49:529-34.
196. Malhi GS, Morris G, Hamilton A et al. Is "early intervention" in bipolar disorder what it claims to be? Bipolar Disord 2017;19:627-36.
197. Duffy A, Grof P. Commentary on McGorry et al.'s Debate on: "Is 'early intervention' in bipolar disorder what it claims to be?". Bipolar Disord 2018;20:556-7.
198. Fusar-Poli P, Sullivan SA, Shah JL et al. Improving the detection of individuals at clinical risk for psychosis in the community, primary and secondary care: an integrated evidence-based approach. Front Psychiatry 2019;10:774.
199. Fusar-Poli P, Rutigliano G, Stahl D et al. Development and validation of a clinically based risk calculator for the transdiagnostic prediction of psychosis. JAMA Psychiatry 2017;74:493-500.
200. Fusar-Poli P, Oliver D, Spada G et al. The case for improved transdiagnostic detection of first-episode psychosis: Electronic Health Record cohort study. Schizophr Res (in press).
201. Raket L, Jaskolowski J, Kinon B et al. Dynamic ElecTronic hEalth reCord deTection (DETECT) of individuals at risk of a first episode of psychosis: a case-control development and validation study. Lancet Digital Health 2020;2:E229-39.
202. Zalsman G, Hawton K, Wasserman D et al. Suicide prevention strategies revisited: 10-year systematic review. Lancet Psychiatry 2016;3:646-59.
203. Schmidt A, Cappucciati M, Radua J et al. Improving prognostic accuracy in subjects at clinical high risk for psychosis: systematic review of predictive models and meta-analytical sequential testing simulation. Schizophr Bull 2017;43:375-88.
204. Oliver D, Radua J, Reichenberg A et al. Psychosis Polyrisk Score (PPS) for the detection of individuals at-risk and the prediction of their outcomes. Front Psychiatry 2019;10:174.
205. Oliver D, Spada G, Englund A et al. Real-world digital implementation of the Psychosis Polyrisk Score (PPS): a pilot feasibility study. Schizophr Res 2020;226:176-83.
206. Petros N, Cullen AE, Fusar-Poli P et al. Towards standardising the assessment of good outcome in people at clinical high risk for psychosis: a collaborative approach. Schizophr Res 2020;223:361-2.
207. Kas MJ, Penninx B, Sommer B et al. A quantita-

tive approach to neuropsychiatry: the why and the how. Neurosci Biobehav Rev 2019;97:3-9.
208. Maj M. Why the clinical utility of diagnostic categories in psychiatry is intrinsically limited and how we can use new approaches to complement them. World Psychiatry 2018;17:121-2.
209. Reed GM, Sharan P, Rebello TJ et al. The ICD-11 developmental field study of reliability of diagnoses of high-burden mental disorders: results among adult patients in mental health settings of 13 countries. World Psychiatry 2018;17:174-86.
210. First MB, Rebello TJ, Keeley JW et al. Do mental health professionals use diagnostic classifications the way we think they do? A global survey. World Psychiatry 2018;17:187-95.
211. Fusar-Poli P, Solmi M, Brondino N, et al. Transdiagnostic psychiatry: a systematic review. World Psychiatry 2019;18:192-207.
212. Fusar-Poli P. TRANSD recommendations: improving transdiagnostic research in psychiatry. World Psychiatry 2019;18:361-2.
213. Jauhar S, Krishnadas R, Nour MM et al. Is there a symptomatic distinction between the affective psychoses and schizophrenia? A machine learning approach. Schizophr Res 2018;202:241-7.
214. Caspi A, Houts RM, Ambler A et al. Longitudinal assessment of mental health disorders and comorbidities across 4 decades among participants in the Dunedin Birth Cohort Study. JAMA Netw Open 2020;3:e203221.
215. Mollon J, David AS, Zammit S et al. Course of cognitive development from infancy to early adulthood in the psychosis spectrum. JAMA Psychiatry 2018;75:270-9.
216. Hickie IB, Hermens DF, Naismith SL et al. Evaluating differential developmental trajectories to adolescent-onset mood and psychotic disorders. BMC Psychiatry 2013;13:303.
217. Hickie IB, Scott EM, Hermens DF et al. Applying clinical staging to young people who present for mental health care. Early Interv Psychiatry 2013;7:31-43.
218. Iorfino F, Scott EM, Carpenter JS et al. Clinical stage transitions in persons aged 12 to 25 years presenting to early intervention mental health services with anxiety, mood, and psychotic disorders. JAMA Psychiatry 2019;76:1167-75.
219. de la Fuente-Tomas L, Sanchez-Autet M, Garcia-Alvarez L et al. Clinical staging in severe mental disorders; bipolar disorder, depression and schizophrenia. Rev Psiquiatr Salud Ment 2019;12:106-15.
220. Duffy A, Malhi GS, Grof P. Do the trajectories of bipolar disorder and schizophrenia follow a universal staging model? Can J Psychiatry 2017; 62:115-22.
221. Dodd S, Berk M, Kelin K et al. Treatment response for acute depression is not associated with number of previous episodes: lack of evidence for a clinical staging model for major depressive disorder. J Affect Disord 2013;150:344-9.
222. Romer AL, Elliott ML, Knodt AR et al. Pervasively thinner neocortex as a transdiagnostic feature of general psychopathology. Am J Psychiatry 2021;178:174-82.
223. Fusar-Poli P, Palombini E, Davies C et al. Why transition risk to psychosis is not declining at the OASIS ultra high risk service: the hidden role of stable pretest risk enrichment. Schizophr Res 2018;192:385-90.
224. Parnas J, Henriksen MG. Epistemological error and the illusion of phenomenological continuity. World Psychiatry 2016;15:126-7.
225. Wigman JT, van Nierop M, Vollebergh WA et al. Evidence that psychotic symptoms are prevalent in disorders of anxiety and depression, impacting on illness onset, risk, and severity – implications for diagnosis and ultra-high risk research. Schizophr Bull 2012;38:247-57.
226. Fusar-Poli P, Raballo A, Parnas J. What is an attenuated psychotic symptom? On the importance of the context. Schizophr Bull 2017;43: 687-92.
227. Schultze-Lutter F, Michel C, Ruhrmann S et al. Prevalence and clinical significance of DSM-5-attenuated psychosis syndrome in adolescents and young adults in the general population: the Bern Epidemiological At-Risk (BEAR) study. Schizophr Bull 2014;40:1499-508.
228. Mishara AL. Klaus Conrad (1905-1961): delusional mood, psychosis, and beginning schizophrenia. Schizophr Bull 2010;36:9-13.
229. Rutigliano G, Valmaggia L, Landi P et al. Persistence or recurrence of non-psychotic comorbid mental disorders associated with 6-year poor functional outcomes in patients at ultra high risk for psychosis. J Affect Disord 2016;203:101-10.
230. Fusar-Poli P, Nelson B, Valmaggia L et al. Comorbid depressive and anxiety disorders in 509 individuals with an at-risk mental state: impact on psychopathology and transition to psychosis. Schizophr Bull 2014;40:120-31.
231. van Os J, Guloksuz S. A critique of the "ultra-high risk" and "transition" paradigm. World Psychiatry 2017;16:200-6.
232. Fusar-Poli P, Spencer T, De Micheli A et al. Outreach and support in South-London (OASIS) 2001-2020: twenty years of early detection, prognosis and preventive care for young people at risk of psychosis. Eur Neuropsychopharmacol 2020;39:111-22.
233. Anglin DM, Galea S, Bachman P. Going upstream to advance psychosis prevention and improve public health. JAMA Psychiatry 2020; 77:665-6.
234. Mennigen E, Bearden CE. Psychosis risk and development: what do we know from population-based studies? Biol Psychiatry 2020;88:315-25.
235. Firth J, Solmi M, Wootton RE et al. A meta-review of "lifestyle psychiatry": the role of exercise, smoking, diet and sleep in the prevention and treatment of mental disorders. World Psychiatry 2020;19:360-80.
236. O'Neil A, Jacka FN, Quirk SE et al. A shared framework for the common mental disorders and non-communicable disease: key considerations for disease prevention and control. BMC Psychiatry 2015;15:15.
237. Fusar-Poli P, Hijazi Z, Stahl D et al. The science of prognosis in psychiatry: a review. JAMA Psychiatry 2018;75:1289-97.
238. Sanfelici R, Dwyer DB, Antonucci LA et al. Individualized diagnostic and prognostic models for patients with psychosis risk syndromes: a meta-analytic view on the state of the art. Biol Psychiatry 2020;88:349-60.
239. Ioannidis JP, Tzoulaki I. What makes a good predictor?: the evidence applied to coronary artery calcium score. JAMA 2010;303:1646-7.
240. Salazar de Pablo G, Studerus E, Vaquerizo J et al. Implementing precision psychiatry: a systematic review of individualised prediction models for clinical practice. Schizophr Bull (in press).
241. Oliver D, Spada G, Colling C et al. Real-world implementation of precision psychiatry: transdiagnostic risk calculator for the automatic detection of individuals at-risk of psychosis. Schizophr Res 2021;227:52-60.
242. Nelson B, Amminger GP, Bechdolf A, et al. Evidence for preventive treatments in young patients at clinical high risk of psychosis: the need for context. Lancet Psychiatry 2020;7:378-80.
243. Kuharic DB, Kekin I, Hew J et al. Preventive treatments in patients at high risk of psychosis. Lancet Psychiatry 2020;7:384-5.
244. Wagenfeld MO. The primary prevention of mental illness: a sociological perspective. J Health Soc Behav 1972;13:195-203.
245. Ebert DD, Buntrock C, Reins JA et al. Efficacy and moderators of psychological interventions in treating subclinical symptoms of depression and preventing major depressive disorder onsets: protocol for an individual patient data meta-analysis of randomised controlled trials. BMJ Open 2018;8:e018582.
246. Wallach JD, Sullivan PG, Trepanowski JF et al. Sex based subgroup differences in randomized controlled trials: empirical evidence from Cochrane meta-analyses. BMJ 2016;355: i5826.
247. Wallach JD, Sullivan PG, Trepanowski JF et al. Evaluation of evidence of statistical support and corroboration of subgroup claims in randomized clinical trials. JAMA Intern Med 2017; 177:554-60.
248. Schuit E, Li AH, Ioannidis JPA. How often can meta-analyses of individual-level data individualize treatment? A meta-epidemiologic study. Int J Epidemiol 2019;48:596-608.
249. Schandelmaier S, Briel M, Varadhan R et al. Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses. CMAJ 2020;192:E901-6.
250. Kent DM, Paulus JK, van Klaveren D et al. The Predictive Approaches to Treatment effect Heterogeneity (PATH) statement. Ann Intern Med 2020;172:35-45.
251. Fusar-Poli P, Davies C, Solmi M et al. Preventive treatments for psychosis: umbrella review (just the evidence). Front Psychiatry 2019;11:764.
252. Fusar-Poli P, Rutigliano G, Stahl D et al. Deconstructing pretest risk enrichment to optimize prediction of psychosis in individuals at clinical high risk. JAMA Psychiatry 2016;73:1260-7.
253. ESHRE Capri Workshop Group. Protect us from poor-quality medical research. Hum Reprod 2018;33:770-6.
254. Berge E, Al-Shahi Salman R, van der Worp HB et al. Increasing value and reducing waste in stroke research. Lancet Neurol 2017;16:399-408.
255. Correll CU, Rubio JM, Inczedy-Farkas G et al. Efficacy of 42 pharmacologic cotreatment strategies added to antipsychotic monotherapy in schizophrenia: systematic overview and quality appraisal of the meta-analytic evidence. JAMA Psychiatry 2017;74:675-84.
256. Pallmann P, Bedding AW, Choodari-Oskooei B et al. Adaptive designs in clinical trials: why use them, and how to run and report them. BMC Med 2018;16:29.
257. Malhi GS, Bell E, Hamilton A et al. Early intervention for risk syndromes: what are the real risks? Schizophr Res (in press).
258. Solmi M, Fornaro M, Ostinelli EG et al. Safety of 80 antidepressants, antipsychotics, anti-attention-deficit/hyperactivity medications and mood stabilizers in children and adolescents with psychiatric disorders: a large scale system-

atic meta-review of 78 adverse effects. World Psychiatry 2020;19:214-32.
259. Correll CU, Rubio JM, Kane JM. What is the risk-benefit ratio of long-term antipsychotic treatment in people with schizophrenia? World Psychiatry 2018;17:149-60.
260. Catalan A, Salazar de Pablo G, Vaquerizo Serrano J et al. Annual research review: Prevention of psychosis in adolescents – systematic review and meta-analysis of advances in detection, prognosis and intervention. J Child Psychol Psychiatry (in press).
261. Berk M, Parker G. The elephant on the couch: side-effects of psychotherapy. Aust N Z J Psychiatry 2009;43:787-94.
262. Tabak AG, Herder C, Rathmann W et al. Prediabetes: a high-risk state for diabetes development. Lancet 2012;379:2279-90.
263. Beauchamp TL, Childress JF. Principles of biomedical ethics, 7th ed. Oxford: Oxford University Press, 2013.
264. Kim SW, Polari A, Melville F et al. Are current labeling terms suitable for people who are at risk of psychosis? Schizophr Res 2017;188:172-7.
265. Yang LH, Link BG, Ben-David S et al. Stigma related to labels and symptoms in individuals at clinical high-risk for psychosis. Schizophr Res 2015;168:9-15.
266. Anglin DM, Greenspoon MI, Lighty Q et al. Spontaneous labelling and stigma associated with clinical characteristics of peers 'at-risk' for psychosis. Early Interv Psychiatry 2014;8:247-52.
267. Lee EH, Hui CL, Ching EY et al. Public stigma in China associated with schizophrenia, depression, attenuated psychosis syndrome, and psychosis-like experiences. Psychiatr Serv 2016;67:766-70.
268. Lane NM, Hunter SA, Lawrie SM. The benefit of foresight? An ethical evaluation of predictive testing for psychosis in clinical practice. Neuroimage Clin 2020;26:102228.
269. Fusar-Poli P, Cappucciati M, Bonoldi I et al. Prognosis of brief psychotic episodes: a meta-analysis. JAMA Psychiatry 2016;73:211-20.
270. Blasco D, Stortz SW, Grivel MM et al. Naturalistic conceptions of genetic optimism and precision psychiatry among those at clinical high-risk for psychosis. Early Interv Psychiatry (in press).
271. Davison J, Scott J. Should we intervene at stage 0? A qualitative study of attitudes of asymptomatic youth at increased risk of developing bipolar disorders and parents with established disease. Early Interv Psychiatry 2018;12:1112-9.
272. Schultze-Lutter F, Schmidt SJ, Theodoridou A. Psychopathology – a precision tool in need of resharpening. Front Psychiatry 2018;9:446.
273. Saxena S, Jane-Llopis E, Hosman C. Prevention of mental and behavioural disorders: implications for policy and practice. World Psychiatry 2006;5:5-14.
274. Bloom DE, Cafiero ET, Jané-Llopis E et al. The global economic burden of noncommunicable diseases. Geneva: World Economic Forum, 2011.
275. Knapp M, Wong G. Economics and mental health: the current scenario. World Psychiatry 2020;19:3-14.
276. Henderson C, Dixon S, Bauer A et al. Cost-effectiveness of PoNDER health visitor training for mothers at lower risk of depression: findings on prevention of postnatal depression from a cluster-randomised controlled trial. Psychol Med 2019;49:1324-34.
277. Ho FY, Yeung WF, Ng TH et al. The efficacy and cost-effectiveness of stepped care prevention and treatment for depressive and/or anxiety disorders: a systematic review and meta-analysis. Sci Rep 2016;6:29281.
278. Aceituno D, Vera N, Prina AM et al. Cost-effectiveness of early intervention in psychosis: systematic review. Br J Psychiatry 2019;215:388-94.
279. Arango C, Diaz-Caneja CM, McGorry PD et al. Preventive strategies for mental health. Lancet Psychiatry 2018;5:591-604.
280. Patton GC, Sawyer SM, Santelli JS et al. Our future: a Lancet commission on adolescent health and wellbeing. Lancet 2016;387:2423-78.
281. Patel V, Chisholm D, Parikh R et al. Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities, 3rd edition. Lancet 2016;387:1672-85.
282. Patel V, Burns JK, Dhingra M et al. Income inequality and depression: a systematic review and meta-analysis of the association and a scoping review of mechanisms. World Psychiatry 2018;17:76-89.
283. Ridley M, Rao G, Schilbach F et al. Poverty, depression, and anxiety: causal evidence and mechanisms. Science (in press).
284. Lund C, De Silva M, Plagerson S et al. Poverty and mental disorders: breaking the cycle in low-income and middle-income countries. Lancet 2011;378:1502-14.
285. Wahlbeck K. Public mental health: the time is ripe for translation of evidence into practice. World Psychiatry 2015;14:36-42.
286. Moreno C, Wykes T, Galderisi S et al. How mental health care should change as a consequence of the COVID-19 pandemic. Lancet Psychiatry 2020;7:813-24.
287. Oliver D, Kotlicka-Antczak M, Minichino A et al. Meta-analytical prognostic accuracy of the Comprehensive Assessment of at Risk Mental States (CAARMS): the need for refined prediction. Eur Psychiatry 2018;49:62-8.
288. Salazar de Pablo G, Catalan A, Fusar-Poli P. Clinical validity of DSM-5 attenuated psychosis syndrome: advances in diagnosis, prognosis, and treatment. JAMA Psychiatry 2020;77:311-20.
289. Bechdolf A, Ratheesh A, Cotton SM et al. The predictive validity of bipolar at-risk (prodromal) criteria in help-seeking adolescents and young adults: a prospective study. Bipolar Disord 2014;16:493-504.
290. Linscott RJ, van Os J. An updated and conservative systematic review and meta-analysis of epidemiological evidence on psychotic experiences in children and adults: on the pathway from proneness to persistence to dimensional expression across mental disorders. Psychol Med 2013;43:1133-49.
291. Cannon TD, Yu CH, Addington J et al. An individualized risk calculator for research in prodromal psychosis. Am J Psychiatry 2016;173:980-8.
292. Carrion RE, Cornblatt BA, Burton CZ et al. Personalized prediction of psychosis: external validation of the NAPLS-2 psychosis risk calculator with the EDIPPP Project. Am J Psychiatry 2016;173:989-96.
293. Zhang T, Li H, Tang Y et al. Validating the predictive accuracy of the NAPLS-2 psychosis risk calculator in a clinical high-risk sample from the SHARP (Shanghai At Risk for Psychosis) Program. Am J Psychiatry 2018;175:906-8.
294. Osborne KJ, Mittal VA. External validation and extension of the NAPLS-2 and SIPS-RC personalized risk calculators in an independent clinical high-risk sample. Psychiatry Res 2019;279:9-14.
295. Zhang T, Xu L, Tang Y et al. Prediction of psychosis in prodrome: development and validation of a simple, personalized risk calculator. Psychol Med 2019;49:1990-8.
296. Fusar-Poli P, Werbeloff N, Rutigliano G et al. Transdiagnostic risk calculator for the automatic detection of individuals at risk and the prediction of psychosis: second replication in an independent national health service trust. Schizophr Bull 2019;45:562-70.
297. Puntis S, Oliver D, Fusar-Poli P. Third external replication of an individualised transdiagnostic prediction model for the automatic detection of individuals at risk of psychosis using electronic health records. Schizophr Res 2021;228:403-9.
298. Oliver D, Wong CMJ, Bøg M et al. Transdiagnostic individualized clinically-based risk calculator for the automatic detection of individuals at-risk and the prediction of psychosis: external replication in 2,430,333 US patients. Transl Psychiatry 2020;10:364.
299. Irving J, Patel R, Oliver D et al. Using natural language processing on electronic health records to enhance detection and prediction of psychosis risk. Schizophr Bull (in press).
300. King M, Walker C, Levy G et al. Development and validation of an international risk prediction algorithm for episodes of major depression in general practice attendees: the PredictD study. Arch Gen Psychiatry 2008;65:1368-76.
301. Nigatu YT, Liu Y, Wang JL. External validation of the international risk prediction algorithm for major depressive episode in the US general population: the PredictD-US study. BMC Psychiatry 2016;16:256.
302. King M, Bottomley C, Bellon-Saameno JA et al. An international risk prediction algorithm for the onset of generalized anxiety and panic syndromes in general practice attendees: predictA. Psychol Med 2011;41:1625-39.
303. Nigatu YT, Wang JL. External validation of the International Risk Prediction Algorithm for the onset of generalized anxiety and/or panic syndromes (The Predict A) in the US general population. J Anxiety Disord 2019;64:40-4.
304. Birmaher B, Merranko JA, Goldstein TR et al. A risk calculator to predict the individual risk of conversion from subthreshold bipolar symptoms to bipolar disorder I or II in youth. J Am Acad Child Adolesc Psychiatry 2018;57:755-63.e4.

DOI:10.1002/wps.20869

Public health psychiatry: an idea whose time has come

Six years ago, K. Walhbeck[1] proposed in this journal that "the evidence base for public mental health interventions is convincing and the time is now ripe to move from knowledge to action". Unfortunately, the field of public mental health has moved too slowly. Indeed, the scholarly review by Fusar-Poli et al[2] concludes that "prevention of mental disorders in young people has not yet solidified as global research or programmatic focus".

Prevention has a long history in medicine, with early successes such as the use of lemons by J. Lind in 1747 to prevent scurvy in the British Navy, and J. Snow's removal of the handle of the Broad Street water pump in 1854 to prevent the spread of cholera in London. There have also been notable advances in prevention of neuropsychiatric disorders. One hundred and fifty years ago, patients with neurosyphilis, such as F. Nietzsche, occupied thousands of beds in mental hospitals. More recently, the toxic effects of phenylketonuria were neutralized by phenylalanine-free diet, and the threatened epidemic of AIDS-related dementia was averted by the development of effective medicines for HIV.

Public health approaches are common in medicine. Mass X-ray screening for tuberculosis was highly effective, and indeed one of us (RMM) was diagnosed, while a Glasgow medical student, as having early tuberculosis by such a screening campaign. Cardiologists, faced with an epidemic of fatal myocardial infarction in the mid 20th century, realized that treatment with evermore expensive interventions was not reducing prevalence; influenced by epidemiologists such as G. Rose, they turned their attention to prevention. Tackling the risk factors for coronary artery disease (such as poor diet, high blood pressure, high cholesterol and smoking) has led to dramatic reductions in the prevalence of myocardial infarction. Similarly, oncologists have long embraced screening and prevention of lung cancer by reducing tobacco smoking in the general population, and now hepatologists are realizing that they cannot continue to treat end-stage liver disease without tackling the root cause – alcohol.

Why has psychiatry lagged so far behind other specialties in embracing a preventive approach? It has not always been like this. During the period of psychoanalytic supremacy, from the 1940s to the 1970s, psychiatrists commonly gave advice on how to improve mental health, for instance by more liberal child rearing practices. Indeed, A. Gregg told the American Psychiatric Association in 1944: "there will be applications [of psychiatry]... to the human relations of normal people – in politics, national and international, between races, between capital and labor, in government, in family life, in education, in every form of human relationship, whether between individuals or groups"[3].

With the decline of psychoanalysis, however, psychiatry retrenched to the clinic and the idea of prevention disappeared from view. The Decade of the Brain from 1990 to 1999 had a primary focus on "brain research", with ever more sophisticated neuroscience, imaging and genetic techniques. But improved knowledge of how the brain "works" did not lead to a reduction in prevalence of mental illness.

As outlined by Fusar-Poli et al, the re-emergence of interest in prevention in psychiatry came with indicated prevention, in the form of early intervention units for first episode psychosis. These have been shown to improve patient health and to be cost-effective. Subsequently, selective prevention in the form of "at risk mental state" services was proposed by McGorry and Yung in Australia, and enthusiastically adopted by academic centres in the US and Europe. The "at risk mental state" paradigm has brought a fresh way of thinking about prevention of mental illness, and, as Fusar-Poli et al note, has now expanded to subsume a transdiagnostic approach and a focus on youth mental health in general. Sadly, this approach has not resulted in the hoped-for reduction in incidence of psychotic disorders, as the service model reaches only a minority of those individuals who will ultimately develop psychosis[4].

Psychiatry needs to move "upstream" and identify possible candidates for selected prevention in childhood, such as subclinical psychotic experiences, developmental delays, psychological and behavioural problems, or family history of mental illness. Focusing on children with a combination of these risk factors, or possibly combining them with biological measures, has potential for intervention. But how to intervene? It has been suggested[5] that "fostering self-esteem, improving parent-child relationships, promoting secure attachment relationships with trusted others, increasing social and neighbourhood supports, and reducing bullying all play a part in improving outcomes". The evidence is there, but psychiatry cannot act alone to implement such broad-ranging measures, and needs "buy-in" from policy makers.

In medicine, universal primary prevention has been shown to be more cost-effective than developing "high-tech" treatments for those with established disease. Persuading the general public not to smoke tobacco has saved many more lives than operating on those with lung cancer or thrombotic coronary arteries. Do we have equivalent opportunities to prevent mental disorder by diminishing population exposure to risk increasing factors? Fusar-Poli et al raise the possibility of reducing mental illness by developing more equitable societies, and point to the high rates of mental disorder in inner cities. High population density, greater exposure to stress, pollution and crime, and lack of green space have all been suggested as responsible for the psychotoxic effect of urbanicity. Although urban planning is beyond the expertise of mental health professionals, we can convince policy makers, by presenting the evidence, that there is an urgent need to re-engineer our cities to improve public mental health.

When examining individual-level risk factors, the best-replicated risk factors in the field of psychosis are obstetric events, child abuse, migration, adverse life events, and heavy cannabis use[6]. Improved perinatal care, supporting positive parenting, and reducing poverty and income inequality can pay dividends for future generations[7]. But there is an urgent need to address one risk factor which is increasing rapidly in both

strength and prevalence – cannabis use.

The worldwide trend towards increasing use of cannabis, especially of high potency varieties, cries out for a preventive approach[8]. A trans-European study estimated that, in London and Amsterdam, 30% and 50% of new cases of psychosis, respectively, would be prevented if no one smoked high-potency cannabis. The risk of developing psychotic disorder was increased 5-fold in those with daily use of high-potency cannabis compared with those who did not use cannabis[8]. This is a similar effect size as between asbestos and lung cancer, but the outcome is much earlier in life. We cannot just wait in our units and emergency departments to treat the increasing numbers of young people with cannabis-related psychosis. There is much to learn from the public education programme implemented in Iceland over the last 20 years, with remarkable decreases in rates of alcohol consumption and tobacco and cannabis smoking among young people[9].

It is time for mental health professionals to speak up about the risks of heavy use of cannabis on rates of psychosis and other mental health problems. It is time to move out of the clinic, remove the handle from the pump, and embrace the challenge of public health psychiatry.

Robin M. Murray[1], Mary Cannon[2]
[1]Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; [2]Royal College of Surgeons of Ireland, Education & Research Centre, Beaumont Hospital, Beaumont, Ireland

1. Wahlbeck K. World Psychiatry 2015:14:36-42.
2. Fusar-Poli P, Correll CU, Arango C et al. World Psychiatry 2021;20:200-21.
3. Murray RM. Lancet 1979;1:255-8.
4. Ajnakina O, David AS, Murray RM. Psychol Med 2019;49:529-34.
5. Healy C, Cannon M. Am J Psychiatry 2020;177: 285-7.
6. Stilo SA, Murray RM. Curr Psychiatry Rep 2019; 21:100.
7. Cannon M, Healy C, Clarke MC et al. Lancet Psychiatry 2020;7:372-3.
8. Murray RM, Hall W. JAMA Psychiatry 2020;77: 777-8.
9. Kristjansson AL, Mann MJ, Sigfusson J et al. Health Promot Pract 2020;21:62-9.

DOI:10.1002/wps.20868

Full speed ahead on indicated prevention of psychosis

Fusar-Poli and eminent colleagues[1] conclude their encyclopedic review of prevention in psychiatry by calling for governments to tackle inequalities in young people's mental health and to invest in improving its social determinants: education, employment, social care, housing, criminal justice, poverty alleviation, social security/welfare benefits, community development, and immigration. We stand firmly with Fusar-Poli et al on this position and would add social justice and public safety to the list. Academics as individuals and their institutions and professional organizations should assist governments to pursue youth mental health as a top priority.

We further commend Fusar-Poli et al for their scholarly review of prevention concepts and in particular their noting that both the public health framework and the World Health Organization framework provide the possibility that some disorders carry risk for other disorders and that conceptual boundaries between preventive and treatment interventions can be porous. We often hear in academic discussions that an intervention must be either preventive or a treatment and that an entity must be defined and named either by risk or by severity, as in clinical high risk (CHR)[2] vs. attenuated psychosis, or prodromal Alzheimer's disease vs. mild cognitive impairment. Our view has long been that the same intervention can provide both treatment and prevention, and that CHR is both a disorder and an indicator of risk for future more severe disorders. In this context, the term "risk syndrome"[3] may be preferable.

We may part ways, however, with Fusar-Poli and colleagues on the relative roles of universal and indicated prevention. Notwithstanding the promise of interventions such as phosphatidylcholine and folic acid tested against surrogate biomarkers, the authors' extensive review sadly identifies few if any universal or selective interventions that meet effectiveness, cost-effectiveness, and implementation standards for reducing the incidence of any mental disorder. The authors' contention that universal public health approaches hold the greatest potential for reducing the risk profile of the whole population does not seem predicated on empirical evidence but rather on theoretical potential.

Along those lines, we take issue with the authors' conceptual Figure 1, partly the basis for their advocacy for universal prevention. This figure shows universal prevention shifting the curve between spectrum of risk and numbers of people to the left, such that there would appear to be no people remaining in the highest risk group who would require indicated prevention. Rather than a shift of a normal curve's x-intercepts to the left, under a universal approach we would expect to see a skewing of the curve such that the risk x-intercepts remain fixed, the left side becomes steeper and higher, indicating a larger number of persons at lower risk, and the right side flattens, indicating a smaller but not zero number of persons at higher risk.

In our alternate conceptualization, there would be a continued need for indicated prevention even under conditions of successful universal prevention. This situation appears to be what occurred in the authors' appropriate example of reducing tobacco use in the population, where new incident cases of non-small cell lung cancer have been reduced by anti-tobacco measures but have not been eliminated[4].

Fusar-Poli et al do advocate for combining universal and indicated prevention, and we staunchly support that advocacy. The non-small cell lung cancer example[4], where mortality has diminished faster than incidence due to the availability of effective new treatments, demonstrates the value at least of tertiary prevention and a potential role for indicated prevention even in the context of effective universal prevention.

With regard to the CHR syndrome as a

vehicle for indicated prevention of psychosis, one of the recent criticisms of the approach, echoed by Fusar-Poli et al, derives from the NEMESIS-2 cohort report that antecedent mood disorders account for more of the incidence of clinical psychosis than do psychotic-like symptoms[5]. We see three important limitations of the NEMESIS-2 data that have received little attention. First, psychotic-like experiences gauged through questionnaires or non-clinical interviews in the general population are not comparable to clinician-assessed CHR syndromes[6]. Second, the time-points in NEMESIS-2 were spaced three years apart. Partly-prospective data show that the average duration of CHR symptoms is two years or less in two-thirds of patients converting to psychosis[7], suggesting that the development of psychotic-like symptoms prior to psychosis may have been missed by the NEMESIS-2 design in as many as half the cases. Third and most crucially, the average age of cohort members at the second time point was 47.7 years, far older than the 12-to-early 30s range where CHR has been reported to predict psychosis and where the incidence of psychosis is known to be highest[8]. As a consequence of these limitations, in our view the NEMESIS-2 data are only partially relevant to the value of CHR as a vehicle for indicated prevention.

With regard to evidence for the success of preventive interventions for CHR, Fusar-Poli et al rightly point out that meta-analytic evidence so far is contradictory and that clinical trials featuring conversion to psychosis as their primary outcome require very large sample sizes. We do, however, see hope on the horizon. This past fall the US National Institute of Mental Health and the Foundation for the National Institutes of Health announced the Accelerated Medicines Partnership in Schizophrenia (AMP SCZ), a collaborative effort to advance early intervention for CHR individuals[9]. This initiative seeks to identify parameters for future clinical trials on alternate outcomes of CHR such as social functioning or attenuated positive symptoms. These alternate endpoints can also potentially serve as surrogate outcomes for reducing the incidence of psychopathology, which can then be investigated directly after entry of the new treatments into clinical practice through epidemiologic methods.

In conclusion, our view is not only that a combined universal and indicated approach is likely to be the best way to prevent psychosis in the future, but also that the CHR syndrome for psychosis continues to provide the most promising option for the indicated prevention component. We acknowledge a potential bias, working as we do in the CHR field, but we like to think we chose this field because it offers the best opportunities in psychiatry for improving public health rather than that we believe it offers the best opportunities for public health because we have chosen it.

Scott W. Woods[1,2], Jimmy Choi[3], Daniel Mamah[4]
[1]Department of Psychiatry, Yale University School of Medicine, and Child Studies Center at Yale, New Haven, CT, USA; [2]Connecticut Mental Health Center, New Haven, CT, USA; [3]Olin Neuropsychiatry Research Center, Hartford Healthcare Behavioral Health Network, Hartford, CT, USA; [4]Department of Psychiatry, Washington University School of Medicine, Washington, WA, USA

1. Fusar-Poli P, Correll CU, Arango C et al. World Psychiatry 2021;20:200-21.
2. Woods SW, Bearden CE, Sabb FW et al. Schizophr Res 2021;227:10-7.
3. Woods SW, Walsh BC, Saksa JR et al. Schizophr Res 2010;123:199-207.
4. Howlader N, Forjaz G, Mooradian MJ et al. N Engl J Med 2020;383:640-9.
5. Guloksuz S, Pries LK, ten Have M et al. World Psychiatry 2020;19:199-205.
6. Schultze-Lutter F, Klosterkötter J, Gaebel W et al. World Psychiatry 2018;17:107-8.
7. Powers AR, Addington J, Perkins DO et al. Schizophr Res 2020;216:443-9.
8. Sullivan SA, Kounali D, Cannon M et al. Am J Psychiatry 2020;177:308-17.
9. US National Institute of Mental Health. Accelerating Medicines Partnership - Schizophrenia (AMP SCZ). https://www.nimh.nih.gov.

DOI:10.1002/wps.20851

Most at-risk individuals will not develop a mental disorder: the limited predictive strength of risk factors

One major problem of preventive psychiatry is the limited predictive strength of all known risk factors for mental disorders, meaning that most of the individuals who are judged to be at risk have only a small chance of developing a mental disorder within the next period of their lives. Fusar-Poli et al[1] have produced an excellent overview of the current state of preventive psychiatry, and they refer to this problem several times. However, we think this is a key issue that deserves more exploration, because it can also give directions for how the prevention field can move forward.

The problem of the low predictive strength of risk factors is partly related to the different priorities of epidemiological research and prevention science. In epidemiological research, the relative risk (RR) or the odds ratio (OR) is often the main indicator describing the strength of the association between a risk factor and a health outcome. However, these indicators only have limited value for prevention science.

For example, if the incidence of a mental disorder in the next year is 1% of the population, and the RR of a group at risk is 4, that means that 4% of this high-risk group will develop the disorder instead of only 1% in the general population. Epidemiological researchers usually stop when they find a (significant) RR of 4, because this indicates a clear high-risk group. However, this is not enough for prevention science. A preventive intervention for a group with 4% risk (instead of 1%) still means that almost all people with this risk factor (96%) will not develop the disorder. Suppose that a preventive intervention can reduce this risk from 4% to 2%. That means that, of the 100 high-risk participants in the intervention, 96 would not develop the disorder anyway and, of the 4 who would, only two will benefit from the preventive effect. This is neither cost-effective nor ethical.

Unfortunately, even though high RRs and ORs are often found in epidemiological research, almost all risk factors in mental health suffer from a low predictive strength. Having a parent with a depressive disorder is often given as an example of a group with an exceptionally high risk. One study even indicates that 50% of these children will develop a depression by the age of 20[2], which is much larger than any other risk factor for mental disorders. But, from the perspective of preventive interventions, even such an elevated incidence rate is still problematic. Suppose that the development of depression starts at the age of 12 and is evenly divided over the subsequent 8 years. This means that every year still only about 6% of these children will develop depression. Offering a preventive intervention to a group in which 94% will not develop the disorder in the following year is still problematic.

Screening for high-risk groups has comparable problems. For example, testing positive for high risk for psychosis has been found to be associated with a 6% lifetime risk of actually developing psychosis[3]. This means that 94% of those who score positive will not develop psychosis in their lifetime, and it can be disputed whether preventive interventions should be considered in these cases[4].

So, from the perspective of preventive interventions, RRs and ORs are clearly not sufficient as indicators of risk. An absolute risk of developing a disorder within a reasonable time frame would be a better indicator. In addition, we need to take the prevalence of the risk factor in the population into account (exposure prevalence), because that indicates the size of the population that will have to be given the intervention.

For example, it is known that women have a higher chance of developing a depressive disorder, but intervening in half of the population is simply not feasible nor cost-efficient (apart from all ethical issues). On the other hand, an intervention in a small group (i.e., with a small exposure prevalence) and a high risk may be useful for the individual participants, but it will not have a large impact on the incidence of a disorder in the general population. This implies that, from the perspective of preventive interventions, we need to identify a population with a modest prevalence (because otherwise the cost of intervening is too high), but this population should be responsible for as many new cases as possible, meaning that the absolute risk is as high as possible in this group.

Finally, preventive interventions should reduce the incidence of the disorder in the population as much as possible. From this perspective, the weak predictive power of most risk indicators is also problematic, because the lower the incidence rate in the population, the larger randomized trials need to be, in order to have sufficient statistical power to be able to show a significant reduction of the incidence[5]. For example, if we were able to identify a high-risk group with 25% incidence in the next year and we had an intervention that is capable to reduce the incidence to 17%, we would need a trial of about 1,000 participants (assuming an alpha of 0.05, 80% power and 20% attrition)[5].

How can this problem of the low predictive power of most risk factors be solved? One possible solution is to focus on combinations of risk factors, that identify groups that are as small as possible but are at the same time responsible for as many incident cases as possible. For example, in one study among older adults, we found that those with sub-threshold depression, functional limitations, a small social network and female gender were 8% of the population, but they explained 24% of the new incident cases of depression[6].

A related solution is to develop prediction tools to identify individuals with a much increased risk for developing mental disorders. The PredictD method has been studied in several large European epidemiological studies[7]. A comparable method has been developed in the US[8]. Based on well-established predictors for the development of depression, these methods calculate the exact personal risk to develop a depressive disorder in the coming year. Unfortunately, these methods do not solve the problem of the low specificity of known risk factors[1]. However, the digitalization of our societies and the progress in epidemiology has resulted in large datasets which may improve such approaches with machine learning techniques.

In addition to the identification of high-risk groups with greater certainty, we also need better interventions. The impact of preventive interventions not only depends on the absolute risk in the target group, but also on their ability to reduce that risk. Some strategies may strengthen the effects of interventions. For example, by focusing on multiple disorders instead of only one, the absolute risk in the target group may be higher and the effects could be demonstrated easier in prevention trials[9]. Stepped care approaches, in which at-risk people are followed over time, may also improve outcome, although that has not been confirmed in all studies.

We conclude that the predictive strength of most risk factors for the development of mental disorders is low and the identification of populations at ultra-high risk is key to the further development of preventive psychiatry.

Pim Cuijpers[1], Filip Smit[2], Toshi A. Furukawa[3]
[1]Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; [2]Trimbos Institute, Utrecht, The Netherlands; [3]Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine, School of Public Health, Kyoto, Japan

1. Fusar-Poli P, Correll CU, Arango C. World Psychiatry 2021;20:200-21.
2. Weissman MM, Wickramaratne P, Nomura Y et al. Am J Psychiatry 2006;163:1001-8.
3. Fusar-Poli P, Cappucciati M, Rutigliano G et al. World Psychiatry 2015;14:322-32.
4. Moritz S, Gaweda L, Heinz A et al. Psychol Med 2019;49:2134-40.
5. Cuijpers P. Am J Psychiatry 2003;160:1385-91.
6. Smit F, Ederveen A, Cuijpers P et al. Arch Gen Psychiatry 2006;63:290-6.
7. King M, Walker C, Levy G et al. Arch Gen Psychiatry 2008;65:1368-76.
8. Wang JL, Sareen J, Patten S et al. J Epidemiol Community Health 2014;68:418-24.
9. Van 't Veer-Tazelaar PA, van Marwijk HWJ, van Oppen P et al. Arch Gen Psychiatry 2009;66:297-304.

DOI:10.1002/wps.20852

Prenatal prevention of psychiatric illness and childhood development population-wide

Fusar-Poli et al[1] comprehensively describe a series of developmental steps that lead to psychosis and related psychiatric illnesses. This perspective is especially relevant when considering prenatal intervention. Several pillars of evidence support the notion that the prenatal period is the first developmental step towards psychosis.

The first pillar is the epidemiological evidence derived from case-control studies that point to maternal adversities, such as prenatal starvation and infection, as risk factors for later psychosis. The second pillar is the molecular evidence that a substantial group of genes associated with psychosis is expressed more robustly in the fetal brain before birth than in the brain after birth[2].

The third pillar is given by post-mortem findings from psychotic individuals that provide information about the maturation of excitatory and inhibitory neurotransmission that normally occurs during gestation. In the brain of persons with schizophrenia, the normal pre-term maturation of glutamate receptors – from lower affinity, slower acting NMDA-type to higher affinity, fast acting kainate type – is incomplete[3]. The maturation of the chloride transporters that support GABA's inhibitory function from the less effective embryonic NKCC1 to the more effective mature KCCN2 is also incomplete[4].

These failures in gestational maturation are thus apparently irreversible over an individual's life span. Furthermore, the likely functional result, brains that process information more slowly and less efficiently, is consistent with schizophrenia patients' life-long deficits, including their well-documented deficit in processing speed. The ability of neurocognitive and pharmacological interventions to ameliorate this deficit in patients after birth is limited. Prevention after birth is important, as outlined by Fusar-Poli et al, but it is hampered by the need to compensate for failures in development before birth that are generally not reversible.

Despite the evidence supporting the potential value of prenatal intervention, its effectiveness for the prevention of psychosis is difficult to ascertain, because of its remoteness from the diagnostic emergence of psychosis in adulthood. An obvious remote aspect is temporal. The results of prenatal interventions instituted today cannot be judged until decades later, when the clinicians and investigators who designed and delivered the intervention may be long forgotten. A second remote aspect is the absence of a nosological link of the early childhood effects of prenatal intervention to psychosis. Because few babies in prenatal intervention studies are destined to develop schizophrenia, the outcome for most babies will be more general improvement in their behavior and cognition. No test in early childhood identifies babies who would have developed schizophrenia, had the intervention not occurred.

In our work, as cited in Fusar-Poli et al's review, we observed a decreased prevalence of a physiological dysfunction in newborns of mothers who received phosphatidylcholine supplements compared to placebo. Phosphatidylcholine is the dietary source of the choline needed to activate fetal alpha 7-nicotinic acetylcholine receptors. The receptor's gene *CHRNA7* is associated with schizophrenia and related psychiatric illnesses, and is expressed more robustly in fetal brain than after birth[2]. Alpha 7-nicotinic receptor activation is a critical mechanism in the maturation of both glutamate receptors and GABA-related chloride transporters[5]. The physiological dysfunction, a partial failure in inhibition of the cerebral evoked response to repeated auditory stimuli that occurs in many people with schizophrenia, indicates that these neurotransmitter mechanisms are not functioning optimally[6]. We observed normal inhibition of the evoked response more frequently among the newborns in the phosphatidylcholine-supplemented group than in the placebo group.

Fusar-Poli et al note that the positive effects of phosphatidylcholine on physiological dysfunctions found in schizophrenia and related mental illness are surrogate markers for prevention of psychosis. As childhood progresses, children whose mothers received phosphatidylcholine supplementation also have positive behavioral effects associated with their positive physiological response as newborns. These children have decreased problems with attention and social withdrawal, compared to children whose mothers received placebo[6]. Conversely, increased problems with attention and social withdrawal are rated retrospectively in children who later have developed schizophrenia, compared to those who did not[7]. Prenatal phosphatidylcholine supplementation appears to help children avoid a developmental pathway that is typical of many individuals who later develop schizophrenia as adults.

Improvement in attention and social function in early childhood is not prevention of mental illness, but neither is it merely a surrogate marker. Children with better attention and social behavior are benefitted in their future success in school, regardless of whether they were destined to become psychotic. Based on the low frequency of psychosis, these more general effects of phosphatidylcholine supplementation may be as important for population well-being as any specific effect on psychosis. If preventive efforts in psychiatry are exclusively focused on prevention of mental illness, we may overlook broader benefits.

Unique aspects of psychotic illnesses, including the psychotic break in late adolescence, certainly merit investigation. However, individuals who do not convert to psychosis, the majority of patients in prodromal or attenuated symptom status, have problems with attention and other cognitive deficits that are similar to those who do convert to psychosis. These cognitive deficits are disabling regardless of whether an individual becomes psychotic or not[8]. Neuropsychological studies in schizophrenia patients find that attention and learning, rather than psychotic symptoms, are the major contributor to most patients' adverse outcomes. Current genome-wide associa-

tion studies, which now identify hundreds of genes in association with schizophrenia as well as with developmental problems, support the thesis that much of the molecular pathology of schizophrenia resides in general brain development that underlies social behavior, attention, and other brain functions.

These clinical and genetic findings suggest a broadened reconceptualization of schizophrenia as a general alteration of neurodevelopmental processes, rather than the outcome of a psychosis-specific pathogenesis. This reconceptualization is congruent with a common characteristic of population-wide primary prevention: beneficial effects on development that extend broadly beyond a narrow disease target. Folic acid, for example, has positive effects on cognition and behavior, in addition to its targeted use to prevent spina bifida and facial clefts. Vitamin D, included in prenatal vitamins to support bone development, appears to be helpful in the prevention of autism spectrum disorder and schizophrenia. Thus, folic acid, vitamin D, and now choline, along with other primary interventions to protect the uterine environment as part of good obstetrical care, have broad beneficial effects for the offspring, in addition to the possible prevention of later psychiatric illness. An example is the significant protective effects of prenatal choline on the development of attention in offspring of women who contract respiratory viruses in gestation[9]. These findings can provide guidance for treatment of pregnant women in the COVID-19 pandemic, so that their children might not add another stone to the pillar of evidence linking prenatal infection to schizophrenia.

Most beneficial effects will appear in early childhood, long before preventive effects for psychosis and other psychiatric illnesses can be definitely ascertained. If expectant families are to see the benefit of improved childhood behavior and cognition with the eventual possible prevention of psychosis, psychiatry cannot be the only discipline to promulgate these prenatal interventions. Prenatal nutrients such as choline that have early beneficial childhood effects require widespread acceptance by obstetricians and maternal-fetal medicine specialists, family medicine physicians, midwives and pediatricians. Working relationships with obstetricians for the assessment of perinatal depression is a model for what needs to happen to allow choline and other prenatal primary preventive interventions to become truly population-wide.

Robert Freedman[1], Sharon K. Hunter[1], Amanda J. Law[1-3], M. Camille Hoffman[1,4]
[1]Department of Psychiatry, University of Colorado School of Medicine, Aurora, CO, USA; [2]Department of Cell and Developmental Biology, University of Colorado School of Medicine, Aurora, CO, USA; [3]Department of Medicine, University of Colorado School of Medicine, Aurora, CO, USA; [4]Department of Obstetrics and Gynecology, Division of Maternal and Fetal Medicine, University of Colorado School of Medicine, Aurora, CO, USA

The perspective of this commentary came from the authors' discussions with the late R.G. Ross. This work was supported by the US National Institute of Child Health and Human Development (K12HD 001271-11), the US National Center for Advancing Translational Sciences (UL1 TR001082), the Institute for Children's Mental Disorders, and the Anschutz Foundation.

1. Fusar-Poli P, Correll C, Arango C et al. World Psychiatry 2021;20:200-21.
2. Birnbaum R, Jaffe AE, Hyde TM et al. Am J Psychiatry 2014;171:758-67.
3. Kerwin R, Patel S, Meldrim N. Neurosci 1990; 39:25-32.
4. Hyde TM, Lipska BK, Ali T et al. J Neurosci 2011; 31:11088-95.
5. Lozada AF, Wang X, Gounko NV et al. J Neurosci 2012;32:7651-61.
6. Ross R, Hunter SK, Hoffman MC et al. Am J Psychiatry 2016;173:509-16.
7. Rossi A, Pollice R, Daneluzzo E et al. Schizophr Res 2000;44:121-8.
8. Addington J, Cornblatt BA, Cadenhead KS et al. Am J Psychiatry 2011;168:800-5.
9. Freedman R, Hunter SK, Law AJ et al. J Psychiatr Res 2020;128:1-4.

DOI:10.1002/wps.20853

Prevention in psychiatry: a role for epigenetics?

In their landmark paper on the current state of knowledge in the field of preventive psychiatry, Fusar-Poli et al[1] state that "robust genetic and environmental epidemiological knowledge is required to inform evidence-based preventive approaches". Indeed, in order to most effectively tailor selective and indicated preventive interventions to an individual's needs, a valid biological and biographical marker panel predictive of future disease risk is essential.

The classical vulnerability-stress model and the extended vulnerability-stress-coping model assume an intricate interplay of biological, particularly genetic, factors with both negative and positive environmental influences in shaping the spectrum of risk and resilience towards mental disorders[2]. However, as rightfully stated by the authors[1], there is currently a "lack of valid biomarkers of risk", and "the variance explained [by polygenic risk scores] is still too small for implementation in selective prevention and does not provide singular neurobiological targets". In other words, to date the field of genetic research, including gene-environment interaction studies and genome-wide approaches, has not fulfilled its initial promise to unambiguously unravel the pathogenetic mechanisms of mental disorders. Consequently, at the present stage, genetic markers are indeed not suitable as valid biomarkers that could inform targeted preventive interventions.

In recent years, however, increasing evidence has accumulated for epigenetic mechanisms such as DNA methylation and histone modifications to crucially govern gene function beyond variation of the DNA itself, and to dynamically respond to environmental influences[3]. Along these lines, epigenetic markers have been suggested to represent an adaptive (or maladaptive) mechanism in the face of environmental challenge, a "molecular embodiment of biography", a "biological archiving" of trauma, adversity, lifestyle and sociocultural context at the crossroads between biology and environment.

Thus, beyond the static genetic level, plastic epigenetic mechanisms seem to be of particular relevance in the conferral of risk or resilience towards mental disorders. Accordingly, epigenetic signatures such as alterations in DNA methylation in blood or saliva have been associated with a number

of mental disorder phenotypes[4,5]. Furthermore, there is initial evidence for peripheral epigenetic markers to be modifiable by psychotherapeutic interventions such as cognitive-behavioral therapy, in that disease-associated DNA methylation patterns have been shown to "normalize" along with treatment response[5]. Overall, these findings suggest a great potential for epigenetic signatures to represent: a) predictive disorder risk markers reflecting both biological and biographical vulnerability, and b) malleable targets for preventive interventions.

Indeed, in plants there is ample evidence for an epigenetic memory of resistance towards environmental pathogens, which has been proposed as a potential new direction in preventing disease in crops[6]. Also, oncological research has identified numerous epigenetic targets in cancer treatment, such as histone deacetylases (HDACs) or DNA methyltransferases (DNMTs), which could further inform preventive strategies for various diseases[7].

With respect to mental disorders, a study probing the effects of a randomized controlled family-centered prevention training program (Strong African American Families, SAAF) discerned parental depressive symptoms to be predictive of accelerated epigenetic aging in the offspring and, reciprocally, the preventive intervention to confer a protective effect regarding epigenetic aging[8].

Additionally, a lifestyle intervention such as physical activity, which is considered to contribute to the promotion of mental health, has been shown to impact the epigenetic machinery. Finally, the field of "nutritional psychiatry" has recently been refueled by evidence for folic acid and vitamin B12 to influence DNA methylation status. In turn, nutritional supplements or epigenetic modifiers such as the natural methyl-group donor S-adenosyl methionine have been suggested as promising adjuncts in the prevention of mental disorders[5].

Given this burgeoning evidence for a possible role of epigenetic processes as targetable risk markers in selective and indicated prevention of mental disorders, further research – ideally expanding to an epigenome-wide and environment-wide level as well as applying a longitudinal study design covering the critical time windows of mental disorder manifestation – is needed to validate and confirm the potential of epigenetic signatures to integratively reflect both a genetic and environmental risk, and thereby confer vulnerability to mental disorder onset.

Additionally, future studies are warranted to explore the malleability of epigenetic markers by preventive interventions. These might comprise classical preventive measures derived from cognitive-behavioral therapy, as well as explore psychopharmacological options, given that several psychoactive substances – such as selective serotonin reuptake inhibitors, antipsychotics, lithium and valproate – have already been reported to impact the epigenetic machinery. Along those lines, "epigenetic drugs" such as HDAC or DNMT inhibitors, if designed specifically enough, might catalyze preventive effects by enhancing learning and neuronal plasticity.

However, some caveats have to be considered when pursuing this line of research. While there is some evidence from studies in rodents and rhesus monkeys, or human positron emission tomography (PET) studies, for a certain comparability of peripheral and central epigenetic processes, some epigenetic signatures seem to be tissue- or even cell-specific, which might limit their use as reliable peripheral biomarkers of mental disorder risk. Also, a number of factors impacting epigenetic mechanisms – such as smoking, exercise, nutrition, body weight, alcohol and drug consumption, or physical diseases – might confound the validity of epigenetic processes as risk markers of mental disorders. Finally, as a general proviso in biomarker research, ethical guidelines and social as well as legal policies for clinical and scientific use of epigenetic information should be implemented alongside such research efforts.

In sum, epigenetics is to be considered a promising field in mental disorder prevention research. First, epigenetic markers – as accessible, integrated and dynamic biosensors of biological as well as biographical risk of mental disorders – might be particularly suited as both indicators and targets of preventive interventions. Second, epigenetic processes – if modifiable by selective or indicated preventive measures – could biologically and thus mechanistically confer resilience towards mental disorders. Finally, as epigenetically imprinted trauma has been reported to potentially be transmissible to future generations via the germline[9], successful preventive interventions embodied in epigenetic signatures might even promote a "transgenerational prevention" of mental disorders, by providing an epigenetic memory of the ability to adapt to a changing environment to future generations.

Katharina Domschke
Department of Psychiatry and Psychotherapy, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, and Center for Basics in NeuroModulation, Faculty of Medicine, University of Freiburg, Freiburg, Germany

This work was partly supported by the German Research Foundation (project no. 44541416) and the German Ministry of Research and Education (01EE1402F).

1. Fusar-Poli P, Correll C, Arango C et al. World Psychiatry 2021;20:200-21.
2. Schiele MA, Herzog K, Kollert L et al. Br J Psychiatry 2020;217:645-50.
3. Schuebel K, Gitik M, Domschke K et al. Int J Neuropsychopharmacol 2016;19:pyw058.
4. Schiele MA, Domschke K. Genes Brain Behav 2018;17:e12423.
5. Schiele MA, Gottschalk MG, Domschke K. Clin Psychol Rev 2020;77:101830.
6. Furci L, Jain R, Stassen J et al. Elife 2019;8:e40655.
7. Feinberg AP. N Engl J Med 2018;378:1323-34.
8. Brody GH, Yu T, Chen E et al. J Child Psychol Psychiatry 2016;57:566-74.
9. Bohacek J, Mansuy IM. Nat Rev Genet 2015;16:641-52.

DOI:10.1002/wps.20854

Primary challenges and practical solutions in preventive psychiatry

Fusar-Poli et al[1] provide a scholarly and detailed overview of the state of knowledge on preventive approaches in psychiatry. Their paper should be considered an obligatory read for anyone entering or already practicing in this emerging field.

The need for preventive approaches in psychiatry is readily transparent. According to the US National Comorbidity Survey[2], a nationally representative population-based survey of mental disorders, one in two adults in the US suffers from the symptomatic and functional challenges of one mental disorder during his/her lifetime. Almost one in three adults will suffer from two or more mental disorders. Regrettably, like much else in psychiatry, preventive approaches are lagging behind general medicine. Fusar-Poli et al make strong arguments about several crucial challenges that critically hamper the implementation of preventive strategies. Here we elaborate on some of the key challenges mentioned in the review, and introduce a set of possible solutions to those.

The primary challenge is finding those who are at risk. Despite the longstanding history of neurobiological research, the underlying causal mechanisms of mental disorders remain mostly unknown. Symptom ratings have been widely used in psychiatry to detect individuals at risk. However, outside of specialty clinics, this strategy seems prone to failure. In a population-based study of 18 to 21-year-olds[3], the presence of symptoms, while associated with subsequent hospitalization for mental disorders, had positive predictive values ranging from 0.54% to 1.99%. In other words, for every correctly identified "case", there would be between 50 and 200 "non-cases" that would be *incorrectly* identified as "cases". Such a high false-positive detection rate, often found when prodrome studies are extrapolated to the general population, questions the utility of current paradigms that aim to identify at-risk groups for large-scale preventive efforts.

Advances in genetic research have identified some syndromic cases across multiple mental disorders, yet the overwhelming majority of individuals with these disorders, and especially those with common disorders (depression, anxiety), are idiopathic, with an unknown etiology. Targetable biomarkers are unavailable to use for early detection and/or efficient early intervention. As Fusar-Poli et al[1] note, only two of 162 peripheral biomarkers were reliably associated with psychosis, bipolar disorder, or depression. Collectively, our current lack of both understanding of underlying causal mechanisms and targetable biomarkers for mental disorders that can be applied at the population level substantially limit preventive strategies.

An additional challenge is that even early intervention often comes too late. Considerable evidence from genetics, epidemiology, basic neuroscience, and neuroimaging implicates early neurodevelopment as the critical period for the risk of developing most mental disorders. Almost all mental disorders are recognizable before or during the second decade of life. Yet, atypical neurobiological development surely predates the emergence of many mental disorders. For instance, evidence suggests that the first signs of cognitive abnormalities in those who will later develop schizophrenia are detectable by the age of four – decades before the disorder is usually diagnosed[4]. Furthermore, the brain most rapidly develops in utero, and continues to do so during early childhood. Indeed, evidence in children of patients with schizophrenia implicates aberrant early, possibly prenatal, brain development[5]. Therefore, these early periods are those when preventive strategies are most likely to have an impact. Fusar-Poli et al[1] highlight this point, but it is transparent that targeting this developmental period is particularly challenging.

A final challenge underscores how we address comorbidities[2]. Comorbidity rates are high in psychiatry and conform to a 50% rule. Approximately half of all people with one psychiatric disorder meet the criteria for a second disorder concurrently; half the people with two disorders meet the criteria for a third; and so on. Evidence based on multiple studies highlights a general underlying dimension, termed the p factor, which captures the tendency to develop psychopathology. In the Dunedin Multidisciplinary Health and Development Study, conducted in an unselected longitudinal birth cohort, higher scores on the general tendency to psychopathology were associated with compromised early-life brain function, and impairments in maturation[6]. Such findings foster the debate regarding categorical versus dimensional models that are relevant to research and in the clinic. In sum, since psychiatric disorders often co-occur, the challenge to clinicians is how to target higher-order psychopathological dimensions and the p factor without loss of specificity[7].

A possible way to address these challenges is to identify those cases that will contribute disproportionally to morbidity and mortality. One source of intriguing evidence comes from another study of the Dunedin Multidisciplinary Health and Development sample, showing that 80% of the health burden is attributable to 20% of cases[8]. That study showed that early-life factors (familial socioeconomic characteristics, maltreatment, IQ, and self-control) clustered into 20% of the population, that accounted for disproportionately high levels of health care use (e.g., 78% of prescription fills and 57% of hospital nights). These findings imply that early life is a critical period for preventive measures for a select group in the population. However, there is potential to abuse this approach; population segments may suffer from stigma. Nevertheless, easing the effects of childhood disadvantage is a critical aim which, if attained in early life, may support families and children, as well as benefit all of society.

A second alternative is to implement universal psychiatric prevention. General medicine has advanced in this prevention (e.g., the efficacy of the COVID-19 vaccines). Evidence-based examples in psychiatry are few, but there are some, such as means restriction to prevent suicide, and physical activity to prevent incident anxiety and depression[9]. Selective universal prevention subtly differs by stratifying prevention to a large group in the population (e.g., nutrient use among pregnant women and the elderly). Better designed, easier to administer universal prevention strategies have the potential to reduce incident mental disorders. They may involve a significant financial investment, but also indirect benefits, including improvements in general health, unemployment, and even crime.

A third alternative is to target not the outcome but an effect modifier for intervention. While biomarkers for mental disorders are not yet available, it is well established that cognitive impairment accompanies, and most often predates by many years, the onset of the majority of mental disorders. There are also reliable ways to measure cognitive functioning and plausible intervention strat-

egies. Implementing interventions to ameliorate cognitive impairments early in life may be a means for psychiatric prevention with substantial societal benefits beyond prevention of psychiatric outcomes (e.g., increasing the cognitive reserve in midlife may be a strategy to reduce dementia).

So, there are multiple challenges to implementing preventive strategies in psychiatry. There is, however, a clear need, and the time is ripe to make the leap towards primary and secondary prevention pathways in the critical period of early life and via cognition.

Abraham Reichenberg[1], Stephen Z. Levine[2]
[1]Department of Psychiatry and Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA; [2]Department of Community Mental Health, University of Haifa, Haifa, Israel

1. Fusar-Poli P, Correll CU, Arango C et al. World Psychiatry 2021;20:200-21.
2. Kessler RC, Berglund P, Demler O et al. Arch Gen Psychiatry 2005;62:593-602.
3. Livny A, Reichenberg A, Fruchter E et al. Am J Psychiatry 2018;175:351-8.
4. Mollon J, David AS, Zammit S et al. JAMA Psychiatry 2018;75:270-9.
5. Birnbaum R, Weinberger DR. Nat Rev Neurosci 2017;18:727-40.
6. Caspi A, Houts RM, Belsky DW et al. Clin Psychol Sci 2014;2:119-37.
7. Kotov R, Jonas KG, Carpenter WT et al. World Psychiatry 2020;19:151-72.
8. Caspi A, Houts RM, Belsky DW et al. Nat Hum Behav 2016;1:0005.
9. Firth J, Solmi M, Wootton RE et al. World Psychiatry 2020;19:360-80.

DOI:10.1002/wps.20855

Prevention in the mental health field should be implemented synergically at different levels

Fusar-Poli et al[1] present a comprehensive preventive framework for improving mental health in young people. Prevention in psychiatry is not a high funding priority, which is also reflected in the relatively low number of publications in the field. The responsibility for primary prevention and mental health promotion is placed in the social and educational sectors and, most often, the evidence base for initiatives is lacking.

In spite of research showing that risk of mental illness is associated with adversities during pregnancy and birth, low socioeconomic status, poor parenting skills, lack of stimulation and support during childhood, bullying, trauma, and early exposure to alcohol and drugs, initiatives to reduce these risk factors have attracted little scientific attention. Much can be done to improve the evidence base for early and broad preventive efforts.

Prevention of psychiatric disorders requires a coherent and multifaceted strategy, including at least five levels. The first is universal primary prevention to improve well-being (e.g., initiatives at the population level focusing on a healthy childhood, such as efforts to improve mental health literacy and parenting in early childhood). The second is universal primary prevention to prevent development of mental illness (e.g., interventions such as prevention of preterm birth and perinatal depression as well as initiatives to prevent bullying and traumatic childhood experiences and to reduce risk of adolescents engaging in substance abuse). The third is selective primary prevention to reduce risk of mental illness in risk groups (e.g., children born to parents with mental illness). The fourth is indicated primary prevention for young people showing signs or symptoms foreshadowing emerging disorder (e.g., clinical high-risk groups for psychosis or children with common mental health problems). The fifth is secondary prevention in early stages of psychiatric disorders (e.g., early intervention services in psychosis or early treatment of attention-deficit/hyperactivity disorder and autism spectrum disorders in child and adolescent services).

Here we focus briefly on selective interventions for families with parental mental illness and on indicated primary prevention initiatives, on the basis of the experience in Denmark.

Children born to parents with mental illnesses constitute an important risk group with a large prevention potential. Danish register-based figures indicate that every sixth child has a parent who has been diagnosed and treated in the secondary mental health sector. The true number at risk is likely to be even higher, since this does not include treatment in primary health care, nor those who, due to lack of accessible treatment offers, fail to be helped by health services. So, this is a very large number of children, who have been shown repeatedly to have a markedly increased risk of being diagnosed with a mental disorder before age 18[2,3], are more likely to live with a single parent[4], are at higher risk of having poor school performance[5], and have more neurocognitive, social and motor problems[6,7] than controls. Due to the parental mental illness, they are also more likely to experience insufficient support and stimulation in the home environment and to be exposed to traumatic life events – all factors that hamper their healthy developmental course.

Parental mental illness is often silenced in the family, passing on stigmatization across generations. Programmes directed towards the whole family should be developed and tested in order to change this trajectory that has been known for decades. Parental training and support as part of the recovery approach, collaboration of adult and child psychiatry with the primary sector, systematic family-based psychoeducation, and social, financial and practical support may be some elements potentially improving the functioning of the entire family and building resilience in the children at risk.

Concerning indicated prevention, implementation of transdiagnostic interventions are suggested to meet the needs of youths with common and multiple mental health problems. A Danish effectiveness study[8] documented the superiority of a new scalable transdiagnostic cognitive behavioral therapy (CBT), called "Mind My Mind" (MMM), compared to management as usual (MAU), for youths aged 6-16 years with emotional and/or behav-

ioral problems below the threshold for referral to mental health care.

A stage-based screening and stratification approach[9] was set up in non-specialized school-based services, with the dual goal to identify: a) the target group of youths with common emotional and/or behavioral problems; and b) those with emerging/severe mental illnesses, e.g. psychosis, who were supported to seek specialized care. The common treatment elements were "distilled" from evidence-based single-disorder CBT programs and organized into modules, materials, video-based feedback, supervision and training of the therapists to help them tailor the treatment to the individual subject.

The flexible and modular transdiagnostic implementation of CBT outperformed MAU on multiple endpoints, including reduced impact of mental health problems on functioning in daily life at the end of treatment, corresponding to a Cohen's effect size of 0.60. Harms were low and non-differential by the end of treatment, but significantly lower with MMM versus MAU at follow-up[8].

All the above-mentioned levels of prevention should be integrated in a common strategy. Interventions at different levels should be regarded not as contradictory, but as synergistic. Therefore, it is sad to witness psychiatrists spending time discussing, for example, the discontinuation of early interventions for high-risk populations in order to prioritize efforts to reduce cannabis use[1]. Instead, we should be inspired by the synergistic approaches implemented in other areas of medicine. Would we see a similar fight in cancer (i.e., scientists attacking each other's efforts in smoking cessation initiatives or screening programs versus surgical or medical treatment for cancer)? Our approach should be that it is important to intervene at all levels depicted above, and that we need studies, and preferably controlled trials, to identify the most effective interventions.

Merete Nordentoft[1,2], Pia Jeppesen[2,3], Anne Amalie Elgaard Thorup[1-3]
[1]Copenhagen Research Center for Mental Health, Mental Health Services, Capital Region of Denmark, Copenhagen, Denmark; [2]Department of Clinical Medicine, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark; [3]Child and Adolescent Mental Health Center, Mental Health Services, Capital Region of Denmark, Copenhagen, Denmark

1. Fusar-Poli P, Correll CU, Arango C et al. World Psychiatry 2021;20:200-21.
2. Thorup AAE, Munk Laursen T, Munk-Olsen T et al. Schizophr Res 2018;197:298-304.
3. Ellersgaard D, Plessen KJ, Jepsen JR et al. World Psychiatry 2018;17:210-9.
4. Ranning A, Munk Laursen T, Thorup A et al. J Am Acad Child Adolesc Psychiatry 2016;55:953-61.
5. Ranning A, Laursen T, Agerbo E et al. Psychol Med 2018;48:1993-2000.
6. Hemager N, Plessen KJ, Thorup A et al. JAMA Psychiatry 2018;75:844-52.
7. Klee Burton B, Thorup AAE, Jepsen JR et al. Lancet Psychiatry 2017;4:400-8.
8. Jeppesen P, Wolf RT, Nielsen SM et al. JAMA Psychiatry (in press).
9. Wolf RT, Puggaard LB, Pedersen MMA et al. Eur Child Adolesc Psychiatry (in press).

DOI:10.1002/wps.20856

Characterizing transdiagnostic premorbid biotypes can help progress in selective prevention in psychiatry

Fusar-Poli et al's insightful paper[1] is a timely appraisal of the foundations of preventive psychiatry. It is a call to action for our field to mount an individual, societal and global response to improve the lives of people with and those at risk for mental disorders. The authors outline a series of ambitious next steps in preventive psychiatry. They seek to advance this goal by integrating universal and targeted frameworks and by advancing our epidemiological knowledge of the multifactorial causation of mental disorders. An additional important step is to use such data toward developing stratified and personalized approaches. However, a major challenge in tackling these ambitious goals is the enormous heterogeneity of mental disorders, at symptomatic, pathophysiologic and etiologic levels. In this light, several strategies deserve consideration toward a successful move forward with Fusar-Poli et al's suggested next steps.

Any effort at prevention should first clarify what we are planning to prevent. For this reason, an accurate and valid diagnosis is critically important. As the authors point out, caseness is difficult to determine in psychiatry, because the disorders are defined based on symptoms, not on biology. For this reason, psychiatric diagnostic systems currently lack validity[2]. A biomarker-based nosology is clearly a critical next step toward stratification of populations meaningfully separating more homogeneous entities.

In a biomarker-driven effort to address the heterogeneity of psychotic disorders, investigators in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (BSNIP) consortium recently used a K-means clustering approach to parse alterations in cognition and electrophysiology (event-related potentials and eye tracking) across the three major psychotic disorders: schizophrenia, schizoaffective disorder, and psychotic bipolar disorder.

Three distinct "biotypes" were identified which seemed orthogonal to the DSM-based categories[3]. Biotype 1 is characterized by severe cognitive impairments, reduced neural response to salient stimuli, marked gray matter reductions, social function deficits, more frequent family history of psychosis, and prominent negative symptoms. Biotype 2 is marked by moderate cognitive and social impairments and gray matter reductions, and by enhanced neural reactivity. Biotype 3 shows few neurobiological differences from healthy controls. These observations point to the possibility that biomarker-derived classifications may potentially better distinguish subtypes within the psychotic spectrum.

However, having a disease-related biomarker is not sufficient for early identification and prevention purposes, unless the biomarker is demonstrated to be present at illness onset or even before overt clinical manifestations of the disorders. This points to the potential value of identifying premorbid biotypes. Interestingly, biotype

1 appears to identify the deficit syndrome, and premorbid adjustment and cognitive profile can distinguish the schizophrenia deficit subgroup with moderate accuracy[4]. It is noteworthy that biotype 1 is associated with higher frequency of family history of psychosis compared to the other biotypes. It is also of interest that cognitive impairment and family history of psychosis[5], as well as biomarkers characterizing biotype 1 such as decreased auditory P300 amplitudes[6], are together strong predictors of risk for conversion to psychosis among individuals at clinical high risk.

A testable prediction, therefore, is whether biotype 1 psychosis may be preceded by a biotype 1-like biomarker signature in the premorbid phase of the illness that is similar to the features seen later in this subtype. Likewise, it is possible that a biotype 1-like biomarker profile may predict impaired functional outcome in early course psychosis patients. Identifying such premorbid bio-signatures requires prospective longitudinal characterization in individuals at familial and clinical high-risk, and those in the early course of a psychotic illness.

Neurobiological entities seem to cut across psychiatric diagnostic categories. Consistent with this view, biotypes of depression[7] and autism[8] have been identified in studies examining the heterogeneity of these syndromes. Interestingly, similar to psychotic disorders, cognitive impairments may serve as valuable stratification markers in these populations as well.

It is useful to consider biomarker-driven approaches in the light of the traditional (primary vs. secondary vs. tertiary) and the more recent (US Institute of Medicine and World Health Organization) models of prevention outlined by Fusar-Poli et al. The identification of transdiagnostic premorbid biomarker signatures and biotypes may be of particular relevance to the field of selective prevention, though not for universal prevention. Biomarker-driven prediction is an aspirational goal for primary selective prevention (e.g., preventing psychosis in individuals at familial high risk for psychosis), though more work is needed in this area. On the other hand, there is emerging evidence in the literature supporting the possibility of predicting psychosis for indicated secondary prevention in individuals at clinical high risk for psychosis[6], and of predicting relapse and functional outcome for the purpose of tertiary prevention in patients in the early course of psychosis[9].

The steady expansion of new knowledge of brain function, and of new approaches, such as imaging, genetics, proteomic and metabolomic technologies, offers the possibility for developing predictive biomarkers in the near future. However, the complex multifactorial determination of mental illnesses and the enormous amount of the available "omics" data make this goal challenging. As Fusar-Poli et al rightly point out, advancing stratified approaches for prevention requires a multicausal, transdiagnostic, multifinal epidemiological knowledge at an individual level. Large multi-site studies, carefully characterized populations, and sophisticated computational approaches, including machine learning, are needed to generate and harness such "big" data sets toward the development of actionable biomarkers for personalized medicine.

In summary, I agree with Fusar-Poli et al's articulation of the need to urgently develop a blueprint for preventive strategies in psychiatry. First, a transdiagnostic view may be applicable not only to psychoses as outlined here, but to all of psychiatric disorders. Second, a neuroscience-based categorization of distinct subtypes in these disorders, as opposed to symptom-based categories, may improve our ability to predict outcome and treatment response. Finally, extending such a translational approach to clinical and familial high-risk states and to early course clinical populations may help identify early predictors of illness and enable individually tailored preventive interventions.

Matcheri S. Keshavan
Beth Israel Deaconess Medical Center, Massachusetts Mental Health Center and Harvard Medical School, Boston, MA, USA

1. Fusar-Poli P, Correll CU, Arango C et al. World Psychiatry 2021;20:200-21.
2. Kapur S, Phillips AG, Insel TR. Mol Psychiatry 2012;17:1174-9.
3. Tamminga CA, Clementz BA, Pearlson G et al. Neuropsychopharmacology 2021;46:143-55.
4. Ahmed AO, Strauss GP, Buchanan RW et al. J Psychiatr Res 2018;97:8-15.
5. Seidman LJ, Giuliano AJ, Meyer EC et al. Arch Gen Psychiatry 2010;67:578-88.
6. Hamilton HK, Roach BJ, Bachman PM et al. JAMA Psychiatry 2019;76:1187-97.
7. Keller AS, Ball TM, Williams LM. Psychol Med 2020;50:2203-12.
8. Wolfers T, Floris DL, Dinga R et al. Neurosci Biobehav Rev 2019;104:240-54.
9. Suvisaari J, Mantere O, Keinänen J et al. Front Psychiatry 2018;9:580.

DOI:10.1002/wps.20857

The Horyzons project: a randomized controlled trial of a novel online social therapy to maintain treatment effects from specialist first-episode psychosis services

Mario Alvarez-Jimenez[1,2], Peter Koval[3], Lianne Schmaal[1,2], Sarah Bendall[1,2], Shaunagh O'Sullivan[1,2], Daniela Cagliarini[1,2], Simon D'Alfonso[1,4], Simon Rice[1,2], Lee Valentine[1,2], David L. Penn[5,6], Chris Miles[1,2], Penni Russon[1,2], Jessica Phillips[1,2], Carla McEnery[1,2], Reeva Lederman[4], Eoin Killackey[1,2], Cathrine Mihalopoulos[7], Cesar Gonzalez-Blanch[2,8], Tamsyn Gilbertson[1,2], Shalini Lal[9-11], Sue Maree Cotton[1,2], Helen Herrman[1,2], Patrick D. McGorry[1,2], John F.M. Gleeson[6]

[1]Orygen, Parkville, VIC, Australia; [2]Centre for Youth Mental Health, University of Melbourne, Melbourne, VIC, Australia; [3]Melbourne School of Psychological Sciences, University of Melbourne, Melbourne, VIC, Australia; [4]School of Computing and Information Systems, University of Melbourne, Melbourne, VIC, Australia; [5]Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [6]Healthy Brain and Mind Research Centre and School of Behavioural and Health Sciences, Australian Catholic University, Melbourne, VIC, Australia; [7]School of Health and Social Development, Deakin Health Economics, Deakin University, Geelong, VIC, Australia; [8]University Hospital Marques de Valdecilla-IDIVAL, Santander, Spain; [9]School of Rehabilitation, Faculty of Medicine, University of Montreal, Montreal, QC, Canada; [10]Health Innovation and Evaluation Hub, University of Montreal Hospital Research Centre, Montreal, QC, Canada; [11]Douglas Mental Health University Institute, Montreal, QC, Canada

This study aimed to determine whether, following two years of specialized support for first-episode psychosis, the addition of a new digital intervention (Horyzons) to treatment as usual (TAU) for 18 months was more effective than 18 months of TAU alone. We conducted a single-blind randomized controlled trial. Participants were people with first-episode psychosis (N=170), aged 16-27 years, in clinical remission and nearing discharge from a specialized service. They were randomly assigned (1:1) to receive Horyzons plus TAU (N=86) or TAU alone (N=84) between October 2013 and January 2017. Horyzons is a novel, comprehensive digital platform merging: peer-to-peer social networking; theory-driven and evidence-informed therapeutic interventions targeting social functioning, vocational recovery and relapse prevention; expert clinician and vocational support; and peer support and moderation. TAU involved transfer to primary or tertiary community mental health services. The primary outcome was social functioning at 18 months as measured by the Personal and Social Performance Scale (PSP). Forty-seven participants (55.5%) in the Horyzons plus TAU group logged on for at least 6 months, and 40 (47.0%) for at least 9 months. Social functioning remained high and stable in both groups from baseline to 18-month follow-up, with no evidence of significant between-group differences (PSP mean difference: –0.29, 95% CI: –4.20 to 3.63, p=0.77). Participants in the Horyzons group had a 5.5 times greater increase in their odds to find employment or enroll in education compared with those in TAU (odds ratio, OR=5.55, 95% CI: 1.09-28.23, p=0.04), with evidence of a dose-response effect. Moreover, participants in TAU were twice as likely to visit emergency services compared to those in the Horyzons group (39% vs. 19%; OR=0.31, 95% CI: 0.11-0.86, p=0.03, number needed to treat, NNT=5). There was a non-significant trend for lower hospitalizations due to psychosis in the Horyzons group vs. TAU (13% vs. 27%; OR=0.36, 95% CI: 0.11-1.08, p=0.07, NNT=7). So, although we did not find a significant effect of Horyzons on social functioning compared with TAU, the intervention was effective in improving vocational or educational attainment, a core component of social recovery, and in reducing usage of hospital emergency services, a key aim of specialized first-episode psychosis services. Horyzons holds significant promise as an engaging and sustainable intervention to provide effective vocational and relapse prevention support for young people with first-episode psychosis beyond specialist services.

Key words: Horyzons, first-episode psychosis, digital intervention, peer support, social functioning, employment, educational attainment, use of emergency services, hospitalization

(World Psychiatry 2021;20:233–243)

Psychosis can be a devastating mental health condition. It typically emerges in adolescence or early adulthood, significantly disrupting achievement of educational, occupational and social milestones and, in many cases, follows a relapsing course, leading to long-term disability[1]. Early intervention – in the form of youth-specific, recovery-focused specialized first-episode psychosis (FEP) services – is now widely seen as the most evidence-based approach to improving the long-term outcomes of psychosis[2].

There are, however, several limitations to the impact of early intervention services. First, specialist FEP services typically provide intensive support for two years, and two clinical trials indicated that some treatment benefits seen at the end of this period may not persist over time[3,4]. Second, social, educational and vocational recovery typically lags behind symptomatic remission, and many young people experience enduring social functioning deficits, and low educational completion and high unemployment rates[5]. Finally, the risk for relapse and hospital admissions remains high beyond discharge from specialized FEP services[1,3,4].

The recognition of these limitations has created an impetus for improving long-term recovery from early psychosis. Along with studies evaluating psychosocial interventions focused on preventing relapse[6] and fostering social and vocational recovery[5,7], three recent clinical trials have evaluated the effects of extending the duration of specialist support (by one[8] to three[9,10] years) compared with the typical timeframe of early intervention services (i.e., two years). These trials have yielded mixed findings, with one of them showing improved length of remission of positive and negative symptoms in the extended model of care (five years) relative to regular care[10], one failing to demonstrate additional benefits from extended specialist support[9], and one showing improved functional outcomes after three years of specialized care which were not sustained at one and two years post-specialist intervention[8].

A promising and potentially cost-effective alternative to extending the duration of specialist FEP services is to provide lower intensity, maintenance treatment following the initial two years

of specialist support[11]. Online, mobile and social media-based interventions provide a novel avenue to offer young people lower intensity, effective, sustainable and scalable support beyond discharge from specialist FEP services. Indeed, preliminary research indicates that online and mobile-based interventions are feasible, acceptable and may improve a range of important domains in early psychosis, including negative symptoms, psychotic symptoms, depression, social functioning, subjective well-being and loneliness[12,13]. Furthermore, initial evidence shows that young people with mental ill-health find online social media-based interventions easy to use, engaging and supportive[14].

Recent psychological models have proposed self-efficacy[15], intrinsic motivation and positive emotions[16] as important targets to promote social functioning in psychosis. Strengths- and mindfulness-based interventions have been put forward as key approaches to increase self-efficacy and positive emotions[17], respectively, with preliminary studies supporting their potential to improve social functioning in psychosis[18]. Similarly, self-determination theory posits that interventions addressing the basic psychological needs of competence, autonomy and relatedness will increase engagement and improve overall functioning through enhanced intrinsic motivation[19]. Recent studies support this theory by showing that increases in intrinsic motivation predict improvements of social functioning in FEP[20].

Drawing on our previous evidence-based interventions in preventing psychosis relapse[6] and improving vocational attainment[5] in FEP, combined with novel approaches to social recovery (strengths- and mindfulness-based interventions) and the principles of self-determination theory, our team developed a world-first digital intervention (Horyzons) designed to foster long-term recovery in FEP. Horyzons blends evidence-based models of social functioning, vocational recovery and relapse prevention into a therapeutic social media environment supported by peer workers as well as clinicians and vocational professionals.

The aim of this study was to examine, via a single-blind randomized controlled trial, whether extending the treatment period of a specialist FEP service through this novel digital intervention added to treatment as usual (TAU) for 18 months was more effective in improving social functioning (primary outcome variable) compared to TAU alone. Among secondary outcomes, we explored the impact of Horyzons plus TAU compared to TAU alone on vocational/educational recovery, visits to emergency services, and hospitalizations due to psychosis during the 18-month follow-up period.

METHODS

Design and participants

The Horyzons study was an 18-month, parallel-group, single-blind, phase 4 randomized controlled trial. Participants were aged 16-27 years and were receiving care at the Early Psychosis Prevention and Intervention Centre (EPPIC), a specialized program of Orygen, Melbourne (Australia). EPPIC is a publicly-funded program servicing 250-300 new FEP referrals per year. It provides 18-24 months of specialized care, after which patients are discharged and transferred to TAU[21].

The study protocol was registered (ANZCTR; ACTRN1261 4000009617) and has been described in detail elsewhere[22]. The trial was approved by the Melbourne Health Human Research Ethics Committee (HREC/12/MH/151; ref. 2013.146).

Inclusion criteria for participants were: a) a first episode of a DSM-IV psychotic disorder or mood disorder with psychotic features; b) aged 16-27 years; c) remission of positive symptoms of psychosis – defined, using the Positive and Negative Syndrome Scale (PANSS)[23], as four weeks or more of scores of 3 (mild) or below on items P2 (conceptual disorganization) and G9 (unusual thought content), and scores of 4 (moderate) or below with no functional impairment on items P3 (hallucinatory behaviour) and P1 (delusions).

Additional inclusion criteria to ensure low level of risk within the trial included: d) low aggressiveness, defined by a score of 3 or below on the poor impulse control item of the PANSS for the month prior to study entry; and e) moderate or lower suicidal risk, defined as a score of 4 or below on the suicidality subscale of the Brief Psychiatric Rating Scale (BPRS)[24] for the month preceding study entry. Finally, participants were required to nominate an emergency contact to be eligible for the study.

Exclusion criteria were: a) intellectual disability; and b) inability to converse in or read English. Additional exclusion criteria to ensure safety within the trial were: c) a DSM-IV diagnosis of either antisocial personality disorder (ASPD) or borderline personality disorder (BPD), as well as clinical evidence that BPD features caused interpersonal difficulties in the treatment environment.

The Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P)[25] was used as the standardized measure of DSM-IV diagnosis of mental illness. The BPD (13 items) and Conduct Disorder/ASPD (22 items) screening questions of the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II) were used to assess for BPD and ASPD[26].

All participants provided written informed consent, which was also obtained from a parent or legal guardian if the participant was younger than 18 years. Recruitment occurred between October 2013 and January 2017. Participants completed four assessments with research assistants at baseline, and at months 6, 12 and 18.

Randomization and patient allocation

Participants were randomly assigned (1:1), following discharge from two years of specialized treatment, to either TAU plus Horyzons or TAU alone for 18 months. An external independent statistician created a computer-generated randomization schedule, comprising randomly permuted blocks. To ensure allocation concealment, the trial coordinator was notified of each randomization via a secure online system and then informed participants of their treatment allocation.

Statisticians and study assessors were masked to treatment allocation until completion of analyses via various procedures detailed in the study protocol[22]. If a study assessor was unblinded, the corresponding participant was allocated to a different study assessor. The study assessors recorded their best guess of participants' treatment allocation at 6-, 12- and 18-month follow-up to assess the success of masking.

Experimental intervention

The Horyzons application was iteratively developed by a multidisciplinary team, in partnership with young people, with the aim of improving social functioning and vocational recovery and prevent relapse in FEP[22].

Horyzons is based on the moderated online social therapy (MOST) model[27,28], which integrates interactive online therapy ("pathways" and "steps"), peer-to-peer online social networking ("the café"), peer moderation, and expert support. Details on each of these components are given in Table 1.

Expert support was provided by registered mental health clinicians (e.g., clinical psychologists, social workers) and vocational workers (trained in Individual Placement and Support) with experience in young people with psychosis. The role of clinicians was to tailor evidence-based interventions, monitor participants' clinical status and ensure the safety of the social network. Each clinician was assigned a caseload, which was followed for the duration of the trial. After baseline assessment, the clinician contacted the participant for a brief phone meeting to review personal needs and preferences. During this initial call, the clinician collaboratively agreed with the participant on the expectations regarding frequency of logins (i.e., weekly or fortnightly). Clinicians then developed brief case formulations which were discussed during weekly supervision meetings with senior clinical psychologists from the team. Guided by the individual formulation, clinicians sent each client weekly tailored content suggestions.

The activity of moderators was informed by the self-determination theory. They supported the autonomy, self-competence and relatedness needs of participants when using Horyzons. For those young people requiring vocational assistance, the vocational moderator provided them with individualized online support, which included: assessing preferences for training, identifying suitable job openings, supporting specific job seeking activities, preparing for a job interview, support for work and study demands, and encouraging use of their personal strengths. Vocational support and online content were informed by the Individual Placement and Support model[5].

The "café" was led by trained young people with lived experience of mental illness ("peer-workers"). They facilitated social learning using Horyzons in desired ways (e.g., sharing helpful

Table 1 Description of Horyzons features

Therapy content	
Pathways	Horyzons includes a number of online "pathways" addressing distinct themes, such as understanding psychosis, identifying early warning signs and preventing relapse, fostering vocational skills, identifying and exercising personal strengths, promoting positive connections with others, fostering positive emotions, managing anxiety and dealing with depression.
Steps	To increase the usability and take-up of therapeutic content, pathways are comprised of thematically related therapy "steps". The online "steps" are evidence-based, discrete, interactive therapy modules primarily targeting: a) social functioning; for example, through fostering self-efficacy (e.g., by identifying personal strengths via an interactive card-sort game based on the strengths-based framework) and positive connections with others (e.g., by illustrating positive and negative responses and relationship dynamics with others); b) vocational recovery; for example, by providing interactive information on how to prepare for a job interview, or how to use personal strengths at work and study; c) relapse prevention; for example, by identifying early warning signs of relapse and developing a relapse prevention plan; and d) comorbid anxiety and depression symptoms; for example, by engaging in relaxation, mindfulness or behavioural activation.
Online social network (Café)	To enhance engagement and foster social support, participants are encouraged to communicate with one another through the online social network. Expert moderators (clinicians and vocational workers) are identifiable as a separate user class within the network. Posts include "icebreakers" (to encourage social interactions), user-generated threads, "reactions" (designed to facilitate social support), as well as content related to mental health or of general interest.
Step content	
Key concepts	Accessible psychoeducational descriptions of therapeutic concepts and outlines based on the purpose of the particular step for the participants.
Comics	Therapy comics, each comprising of 20 to 24 story board panels focusing on a particular therapeutic theme and target related to the treatment.
Do its	To ensure that therapeutic concepts are translated into behavioural change, the "steps" include behavioural prompts known as actions or "Do its". For example, following a step about finding jobs, the participants would find specific behavioural suggestions prompting them to "drop off their CV in the reception areas of 10 different organizations". "Do its" are also related to the participant's specific strengths (e.g., using courage when facing stressful social situations).
Talk it out	"Talk it out" is an online group function informed by the evidence-based problem-solving framework. It enables participants to propose problems (e.g., "should I discuss my mental health issues in a job interview?"), which are discussed in moderated groups through structured phases (e.g., brainstorming, pros and cons, wrap-up). Previous problems and group solutions are stored in the system, providing an easily accessible "solution wiki" for future young people.

content). Peer-workers also seeded discussion threads to promote engagement and connection and to normalize experiences.

Control intervention

Participants allocated to regular care received TAU following discharge from the EPPIC program. We chose TAU as comparator to enhance external validity because it replicates the current mainstream post-discharge treatment options available to young people with FEP. This parallels three recent randomized controlled trials examining extended interventions for FEP services[8-10].

TAU comprised various treatment options delivered by generic medical or mental health services typically available to young people. Those with complex needs were referred by the EPPIC team to adult tertiary community mental health services, whereas those who achieved a good level of recovery and clinical stability were referred to primary care services (including access to multidisciplinary youth mental health services and government-subsidized psychological and psychiatric treatment). TAU participants were also provided with a printed leaflet and a universal serial bus (USB) containing relevant information on free online youth resources (i.e., Moodgym, e-headspace, Reach-out).

Outcome measures

The primary outcome was change in social functioning, as measured by the Personal and Social Performance Scale (PSP)[29], from baseline to 18-month follow-up. Secondary outcomes (change from baseline to 18-month follow-up, or incidence within the 18-month follow-up) included visits to emergency services, hospital admissions due to mental health issues in general or specifically to psychosis, vocational/educational recovery (i.e., working in a job that paid the legislated minimum wage for a minimum of a week and/or being enrolled in education in the previous 6-month period), depression (as assessed by the Calgary Depression Scale for Schizophrenia, CDSS[30]), loneliness and social support (evaluated by the UCLA Loneliness Scale, Version 3[31], and the Medical Outcomes Study Social Support Survey, MOS-SSS[32], respectively), self-esteem and self-efficacy (assessed by the Self-Esteem Rating Scale - Short Form, SERS-SF[33], and the Mental Health Confidence Scale, MHCS[34], respectively), satisfaction with life (evaluated using the Satisfaction with Life Scale, SWLS[35]), quality of life (measured by the Assessment of Quality of Life - 8D, AQoL-8D[36]), and positive and negative psychotic symptoms (assessed by the PANSS).

Seventeen cases were selected at baseline for the purpose of checking interrater reliability on the interview rated measures – PSP, PANSS and CDSS – with an independent research assistant making simultaneous ratings. The intraclass correlation coefficients were 0.90 for PSP, 0.89 for PANSS, and 0.94 for CDSS, which indicates good interrater reliability.

To determine success of blinding, the kappa statistic was used as a measure of agreement beyond that caused by chance[37]. The guesses by the study assessors about treatment group were compared with actual treatment allocation. There was no evidence of unblinding by study assessors. The kappa statistics were 0.01, 0.08 and 0.29 at 6-, 12- and 18-month follow-up assessments, respectively. A kappa statistic of less than 0.40 indicates poor agreement[37].

Data analysis

The main analyses were done on an intention-to-treat basis, including all participants and all available data. Additional analyses were completed on *a priori* established per-protocol basis, including participants in the intervention group who received a pre-specified minimal exposure to the online intervention (i.e., >8 logins during the 18-month intervention[22]).

For continuous variables, we compared the groups using linear mixed models with a restricted maximum likelihood estimator implemented by the lme4 (version 1.1-23) and lmerTest (version 3.1.2) packages in R (version 3.6.2). The models included random intercepts for each participant, and the fixed effects of treatment, time (baseline, 6-, 12- and 18-month follow-up), and treatment-by-time interactions. Gender, age, the relevant baseline scores of the outcome variable, and covariates which were significantly different across treatment groups at baseline (i.e., duration of untreated psychosis, DUP) were also included as fixed effects (i.e., controlling for their effects).

Vocational/educational outcome (categorical) was analyzed using multilevel logistic regression including random intercepts for each participant, and the fixed effects of treatment, time, treatment-by-time interactions, gender, age and other relevant covariates as described above.

For all analyses, the primary effects of interest were the treatment-by-time interactions representing group differences in linear change from baseline to month 18 (primary end point).

The total number of hospital admissions due to psychosis or in general to mental health issues and of visits to emergency services over the 18-month follow-up period were compared between groups using logistic regression, including gender, age and DUP as covariates in the models. We used two-tailed tests with $p<0.05$ denoting statistical significance.

In addition to the planned contrast of interest for changes between baseline and 18 months, we also examined group differences at 6 and 12 months if there was a statistically significant overall treatment-by-time interaction.

RESULTS

Eighty-six participants (50.5%) were randomly assigned to the Horyzons plus TAU group and 84 (49.5%) to the TAU group. Participants had a mean age of 20.91 years (SD=2.88) (Table 2). With the exception of DUP, which was significantly longer in the Horyzons plus TAU group (median: 7.36 weeks) relative to the TAU group (median: 4.29 weeks), all socio-demographic and diagnostic covariates were

Table 2 Baseline patient characteristics

	Horyzons plus TAU (N=86)	TAU (N=84)	Total (N=170)
Age (years, mean±SD)	21.01±2.93	20.81±2.83	20.91±2.88
≤18 years, N (%)	23 (26.7)	25 (29.8)	48 (28.2)
>18 years, N (%)	63 (73.3)	59 (70.2)	122 (71.8)
Gender, N (%)			
Males	45 (52.3)	45 (53.6)	90 (52.9)
Females	41 (47.7)	39 (46.6)	80 (47.1)
Employment status, N (%)			
Unemployed	32 (39.0)	24 (29.3)	56 (34.1)
Studying only	16 (19.5)	23 (28.0)	39 (23.8)
Paid work only	20 (24.4)	17 (20.7)	37 (22.6)
Concurrent study and paid work	14 (17.1)	18 (22.0)	32 (19.5)
Educational status, N (%)			
Not currently studying	54 (62.8)	39 (46.4)	93 (54.7)
Not currently studying, but enrolled in upcoming course	2 (2.3)	4 (4.8)	6 (3.5)
Studying part-time	5 (5.8)	14 (16.7)	19 (11.2)
Studying full-time	25 (29.1)	27 (32.1)	52 (30.6)
Highest year completed at school, N (%)			
Year 8	1 (1.2)	2 (2.4)	3 (1.8)
Year 9	7 (8.2)	7 (8.3)	14 (8.3)
Year 10	16 (18.8)	19 (22.6)	35 (20.7)
Year 11	16 (18.8)	20 (23.8)	36 (21.3)
Year 12	45 (52.9)	36 (42.9)	81 (47.9)
Diagnosis, N (%)			
Affective psychosis	29 (33.7)	29 (34.5)	58 (34.1)
Non-affective psychosis	57 (66.3)	55 (65.5)	112 (65.9)
Duration of untreated psychosis (weeks, median and range)*	7.36 (1.00-52.14)	4.29 (0.64-11.93)	4.29 (0.86-19.57)

TAU – treatment as usual
*Significant difference between TAU and Horyzons plus TAU ($p<0.05$)

well balanced between groups at baseline (Table 2). There were no differences between participants who were included in the study and those who declined participation in terms of age and gender.

Seventy-two of 86 participants in the Horyzons plus TAU group (83.7%) and 75 of 84 in the TAU group (89.3%) completed at least one post-baseline (i.e., 6-, 12- and/or 18-month) assessment. Moreover, 63 participants in the Horyzons group (73.2%) and 63 in the TAU group (75.0%) completed the 18-month follow-up assessment (see Figure 1). There were no differences between those who were lost to follow-up and those who completed the 18-month assessment with respect to socio-demographic, diagnostic, clinical and functioning variables.

Data on engagement with Horyzons are provided in Table 3. Participants had an average of 106.84 logins (SD=247.05), with 69 (80.2%) participants logging on for at least 3 months, 47 (55.5%) for at least 6 months, 40 (47.0%) for at least 9 months, and 25 (29.0%) for at least 12 months.

For our primary outcome variable, changes in PSP scores at 18-month follow-up, we found no significant group-by-time interaction effect (mean difference = –0.29, 95% CI: –4.20 to 3.63, standardized effect size = –0.01, p=0.77) in the main intention-to-treat analysis. Level of functioning remained stable for both groups from baseline to 18-month follow-up (Table 4).

We found a significantly better vocational/educational outcome in the Horyzons plus TAU group compared with the TAU group (Table 5). Specifically, participants in the Horyzons group had 5.5 times greater increase in their odds of finding employment or enrolling in education from baseline to 18 months compared with those in the TAU group (odds ratio, OR=5.55, 95% CI: 1.09-28.23, p=0.04). Moreover, participants allocated to the TAU group had twice the rate of hospital admissions due to psychosis compared with their counterparts in the Horyzons plus TAU group, although this difference did not reach the level of statistical significance (27% vs. 13%, respectively; OR=0.36,

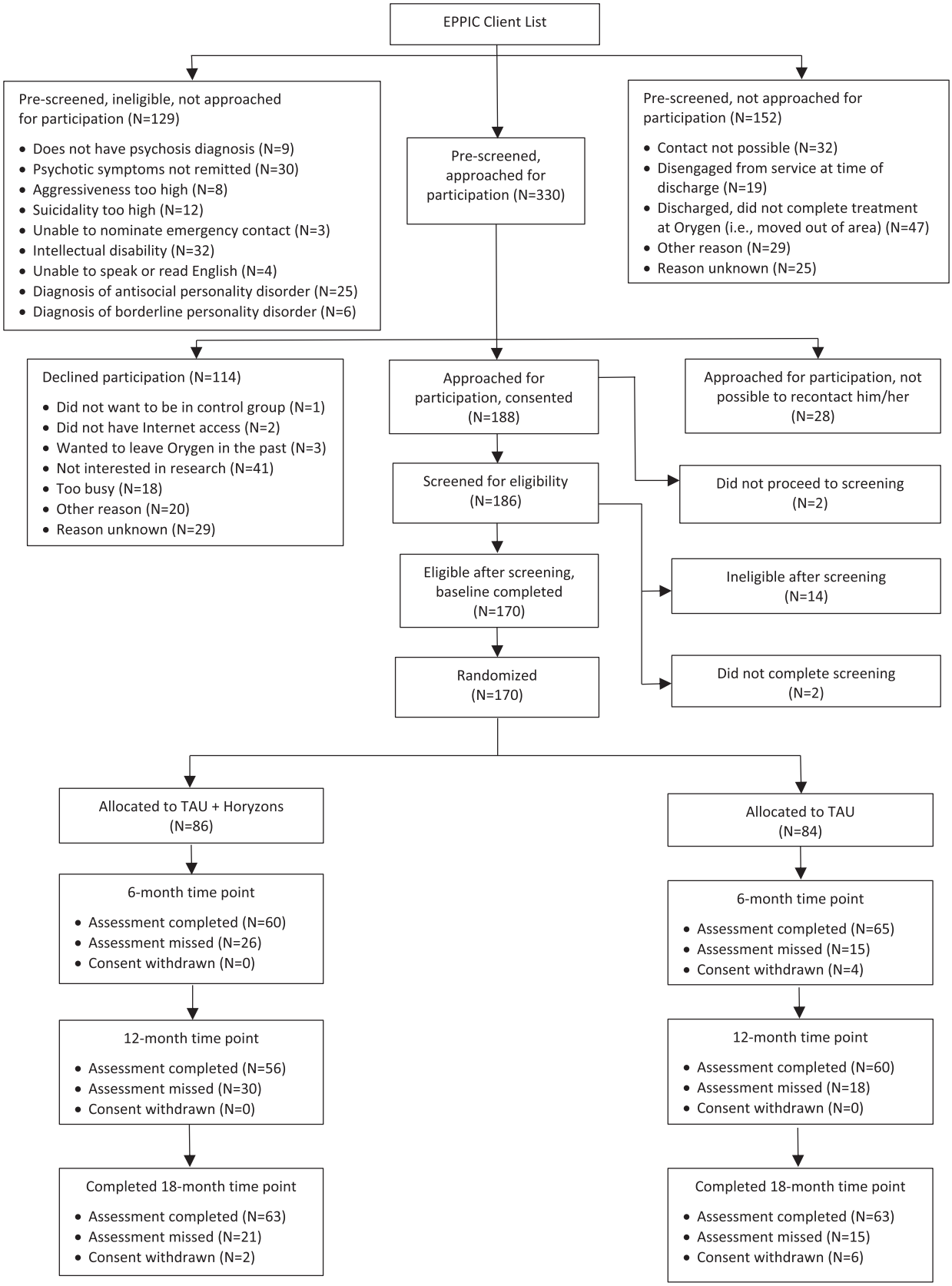

Figure 1 Trial profile. EPPIC – Early Psychosis Prevention and Intervention Centre, TAU – treatment as usual

Table 3 Engagement with Horyzons

	N (%)	Mean (SD)	Median (IQR)
Number of logins over 18 months		106.84 (247.05)	24 (8.5-84)
≤8	21 (24.7)		
9-17	15 (17.6)		
18-76	24 (28.2)		
77-1,529	25 (29.4)		
Number of steps done		16.99 (21.76)	9 (3-21)
1-5	30 (35.3)		
6-15	26 (30.6)		
16-130	29 (34.1)		
Number of actions done		5.29 (8.11)	2 (0-7)
None	27 (31.8)		
1-5	32 (37.6)		
6-47	26 (30.6)		
Number of newsfeed posts and/or comments		21.49 (41.71)	7 (1.25-21)
None	14 (16.7)		
1-5	25 (29.8)		
6-25	29 (34.5)		
26-266	16 (19.0)		
Length of engagement (months)		8.15 (5.65)	7 (3-13)
At least 1 month	76 (88.4)		
At least 3 months	69 (80.2)		
At least 6 months	47 (55.5)		
At least 9 months	40 (47.0)		
At least 12 months	25 (29.0)		
Full 18-month period	7 (8.1)		

IQR – interquartile range

95% CI: 0.11-1.08, p=0.07, number needed to treat, NNT=7) (Table 5). Consistent with this finding, those allocated to the TAU group had twice the rate of visits to emergency services compared with those in the Horyzons plus TAU group from baseline to 18 months, a statistically significant difference (39% vs. 19%, respectively; OR=0.31, 95% CI: 0.11-0.86, p=0.03, NNT=5) (Table 5).

Changes in other secondary outcome variables did not differ between the groups from baseline to 18-month follow-up (Table 4). Additional analyses to the primary contrast of interest (changes between baseline and 18 months) found a significant overall treatment-by-time interaction effect on negative symptoms (as measured by the PANSS scale). Post-hoc analyses revealed that this effect was driven by a significantly greater reduction of negative symptoms in participants allocated to the Horyzons plus TAU compared with those in the TAU group from baseline to 12-month follow-up (p<0.05); however, these effects on negative symptoms were lost from 12-month to 18-month follow-up.

Effect sizes from the per-protocol analyses were consistent with the primary intent-to-treat analyses.

DISCUSSION

Sustained social and vocational recovery is the ultimate goal of specialist FEP services as well as the most valued outcome by young people and their families[38]. Yet, follow-up studies have questioned the maintenance of treatment effects of early psychosis services[3,4]; social and vocational recovery continues to be resistant to current intervention approaches[5]; and relapse rates remain high beyond discharge from specialized services[1,3,4]. Addressing this gap, this is the first randomized controlled trial to examine whether a novel digital intervention is an effective strategy to extend the treatment benefits of early intervention and foster social and vocational recovery beyond discharge from specialist FEP services.

We did not find a significant between-group difference in social functioning (primary outcome) as measured by the PSP at 18 months. Participants in both groups showed relatively high levels of social functioning at baseline, which were maintained throughout the study. On the other hand, secondary analyses revealed that participants who received the Horyzons intervention plus TAU had a 5.5 times greater increase in their odds of finding competitive employment and/or enroll in education – a key aspect of functional recovery – compared with those receiving TAU alone from baseline to 18 months. Moreover, we found twice the incidence of hospital admissions due to psychosis in the TAU group than in the Horyzons plus TAU group. While the between-group difference did not reach the level of statistical significance (p=0.07) (event rates were low), the differential rate is notable, and this suggestive evidence is supported by the consistent finding that participants allocated to the Horyzons intervention were significantly less likely to visit emergency services over the 18-month period (p=0.03) compared with their counterparts in the TAU group.

In line with previous studies, we hypothesized that the potentially disruptive effects of transfer of care from a specialized to generic services, coupled with the sense of loss, change of clinical care and reduced multidisciplinary input would lead to a functional deterioration in the TAU group[3]. This would have been consistent with Chang et al's finding that the functional decline following termination of specialized care took place primarily in the first year following discharge[8]. By contrast, in keeping with previous research[11], we expected that, by providing an online step-down model of care, we would prevent the loss of functional gains in the Horyzons group. Contrary to our expectations, while participants allocated to the Horyzons plus TAU group maintained their level of functioning throughout the study, so did those in the TAU group.

There are a number of explanations that could account for this finding. First, baseline social functioning in our sample (at the point of discharge from a specialist FEP service) was noticeably

Table 4 Social functioning and continuous secondary outcome variables at baseline and 18 months (intent-to-treat analysis)

	Horyzons plus TAU	TAU	Mean difference (95% CI)	Standardized effect size	p
Social functioning (PSP score, mean±SE)			–0.29 (–4.20 to 3.63)	–0.01	0.77
Baseline	67.36±1.21	66.37±1.24			
18 months	67.04±1.38	66.75±1.42			
Depression (CDSS score, mean±SE)			0.31 (–0.82 to 1.44)	0.05	0.42
Baseline	3.23±0.35	3.00±0.36			
18 months	4.13±0.40	4.44±0.41			
Loneliness (UCLA score, mean±SE)			0.94 (–2.05 to 3.94)	0.06	0.54
Baseline	46.06±0.89	46.12±0.94			
18 months	44.12±1.05	45.07±1.10			
Social support (MOS-SSS score, mean±SE)			0.08 (–5.51 to 5.68)	–0.003	0.82
Baseline	71.11±1.68	70.45±1.75			
18 months	72.99±1.96	73.08±2.05			
Self-esteem (SERS-SF score, mean±SE)			1.07 (–4.89 to 7.04)	0.03	0.89
Baseline	12.24±1.79	12.84±1.88			
18 months	13.78±2.09	14.85±2.19			
Self-efficacy (MHCS score, mean±SE)			2.25 (–2.14 to 6.65)	0.09	0.30
Baseline	68.22±1.31	67.84±1.35			
18 months	68.57±1.56	70.82±1.59			
Satisfaction with life (SWLS score, mean±SE)			–0.29 (–2.13 to 1.55)	–0.03	0.67
Baseline	20.99±0.56	21.19±0.59			
18 months	22.63±0.65	22.34±0.67			
Quality of life (AQoL-8D total score, mean±SE)			0.01 (–0.04 to 0.07)	0.05	0.59
Baseline	0.60±0.02	0.60±0.01			
18 months	0.63±0.02	0.65±0.02			
Positive symptoms (PANSS Positive score, mean±SE)			–0.82 (–1.98 to 0.35)	–0.12	0.37
Baseline	10.02±0.36	9.68±0.37			
18 months	11.08±0.41	10.26±0.43			
Negative symptoms (PANSS Negative score, mean±SE)			–0.83 (–1.99 to 0.34)	–0.12	0.34
Baseline	11.21±0.36	11.05±0.37			
18 months	12.26±0.41	11.43±0.42			

The p value represents the group-by-time interaction effect from baseline to 18-month follow-up. TAU – treatment as usual, PSP – Personal and Social Performance Scale, CDSS – Calgary Depression Scale for Schizophrenia, UCLA – UCLA Loneliness Scale (Version 3), MOS-SSS – Medical Outcomes Study Social Support Survey, SERS-SF – Self-Esteem Rating Scale - Short Form, MHCS – Mental Health Confidence Scale, SWLS – Satisfaction with Life Scale, AQoL-8D – Assessment of Quality of Life - 8D, PANSS – Positive and Negative Syndrome Scale

higher compared to other similar studies. Specifically, the mean social functioning score at study entry was 66.6 in our trial (PSP), compared with 57 (Social and Occupational Functioning Assessment Scale, SOFAS) in Chang et al's study[8] and 48 (PSP) in Albert et al's trial[9]. Moreover, DUP – a marker of both long-term functioning and treatment response in extended specialist FEP services[39,40] – was also comparatively briefer in our cohort (4.3 weeks) vs. prior studies (121-164 weeks in Albert et al's study[9], 12 weeks in Malla et al's study[10], 13 weeks in Chang et al's trial[8]). These differences could reflect the intensity and quality of the background treatment in our study. In particular, unlike other specialized FEP services, EPPIC provides a comprehensive group program and Individual Placement and Support to promote social and vocational recovery as part of the service. Alternatively, the inclusion

Table 5 Binary secondary outcome variables (intent-to-treat analysis)

	Horyzons plus TAU, N (%)	TAU, N (%)	Odds ratio (95% CI)	p	Number needed to treat
Vocational or educational recovery					
Baseline	45 (62%)	56 (74%)	5.55 (1.09-28.23)	**0.04**	
18 months	47 (78%)	44 (70%)			
Hospital admissions due to mental health issues	12 (22%)	17 (31%)	0.46 (0.15-1.30)	0.15	11
Hospital admissions due to psychosis	7 (13%)	15 (27%)	0.36 (0.11-1.08)	0.07	7
Visits to emergency services	10 (19%)	21 (39%)	0.31 (0.11-0.86)	**0.03**	5

TAU – treatment as usual. Significant differences are highlighted in bold prints

and exclusion criteria employed to ensure the safety of the trial (i.e., clinical remission) could have led to a sample of higher functioning individuals at baseline[6].

Second, the sustained level of functioning in the TAU group could be accounted for by the quality and intensity of TAU following EPPIC treatment, which included follow-up treatment options such as multidisciplinary youth mental health services (e.g., headspace services) as well as government-subsidized psychological and psychiatric treatment.

Taken together, the higher baseline social functioning and shorter DUP in our cohort, coupled with the availability of publicly funded youth mental health support post-discharge from EPPIC, could have reduced the likelihood of finding group differences in social functioning over time. On the other hand, it could be that Horyzons is not effective enough in improving social functioning in this population, or that a different treatment modality, different or additional therapeutic targets, or a minimal threshold or a specific pattern of usage, are needed to demonstrate improved social functioning at follow-up.

The last above postulate is supported by our examination of the relationship between patterns of usage of Horyzons and outcomes. This analysis revealed that Horyzons users who showed consistent engagement with the social and therapeutic components of the digital platform experienced significant improvements in social functioning and negative symptoms compared with those with lower usage and those allocated to the TAU group (after controlling for potential confounders)[41].

A key finding of this study was that vocational/educational outcome improved significantly in the Horyzons plus TAU group compared with the TAU group, which deteriorated over the same period. Of note, *post-hoc* analyses provided evidence of a dose-response effect, with those participants in the top quartile of logins (i.e., logging on >77 times) showing a greater improvement on vocational and educational recovery (OR=59.71; 95% CI: 2.40-1484.37, p=0.01) compared with those in the bottom quartile of logins (i.e., <9 logins) (OR=1.40; 95% CI: 0.03-72.40, p=0.87).

This study is the first to demonstrate that extending the duration of support following specialist FEP services leads to improved vocational/educational outcome over a prolonged follow-up period. This finding has significant treatment and recovery implications. The extant evidence indicates that the positive effects of face-to-face Individual Placement and Support in FEP may wane after the intervention period[5]. Moreover, securing and maintaining employment and completing education remain a top priority for young people with psychosis, are critical aspects of mental health recovery and normative development, and constitute a protective factor against mental ill-health[42]. This study shows for the first time that a digital intervention integrating support by vocational workers and evidence-based vocational content is an effective strategy to address this critical treatment goal and potentially extend the benefits of existing evidence-based interventions in this population.

The study results provide support for the effect of Horyzons in reducing the rate of hospital admissions following discharge from specialist FEP services. While the difference with respect to the TAU group did not reach the level of statistical significance (p=0.07), the differential rate is evident (13% vs. 27%), and the low event rates significantly reduced the statistical power for this analysis. The clinical validity of this finding is strengthened by the associated finding that participants allocated to the TAU group were twice as likely to visit emergency services during the follow-up compared to those in the Horyzons group (39% vs. 19%, p=0.03). Of note, there were a total of 12 repeated visits to emergency services from seven different participants, all of which occurred in the TAU group.

It may be that Horyzons acts on distress, reducing utilization of emergency services and hospital admissions through in-the-moment access to online therapy, and peer and social support. This is in line with previous research showing that social support is associated with reduced risk of relapse in FEP[1]. The estimated NNT for Horyzons to prevent one visit to emergency services and one hospital admission were 5 and 7, respectively. This is comparable with the reported NNT for specialist FEP programs to prevent one relapse (NNT=8) and somewhat lower than the NNT with second-generation antipsychotics to prevent one relapse (NNT=10)[2].

Our exploratory analysis showed lower levels of negative symptoms from baseline to 12 months in the Horyzons group compared with the TAU group. This effect, however, was lost at 18-month follow-up. Malla et al[10] found that extending the duration of specialist FEP services was associated with improved negative symptoms at 5-year follow-up compared with TAU. In addition, similar to our findings, Chang et al[8] found a reduction in negative symptoms following one year of extended specialist FEP treatment which was

lost at 2- and 3-year follow-up. Our results suggest that Horyzons may have time-limited favorable effects on negative symptoms, corresponding with the period of higher usage of the program.

We did not find evidence for the effectiveness of Horyzons on other secondary outcome variables such as depression, social support, loneliness and quality of life. Several explanations may account for the lack of treatment effects on these variables. First, it is likely that bringing about treatment effect on specific outcome variables (e.g., depression) requires intensive, focused engagement of specific targets (e.g., rumination or behavioural activation[5,6]). Second, Horyzons is one of the first interventions harnessing social networking to promote both engagement and social support. However, we found that, whereas many young people had positive experiences of social connection on Horyzons, others experienced barriers (such as social anxiety, paranoia and confusion within the social network) that thwarted their need for connection with others[43]. Further research is required to determine the optimal features and operations of online social media-based interventions so that they support connectedness, whilst addressing barriers to meaningful engagement.

With the aim of sustaining the benefits of specialist FEP services, Horyzons was delivered for a period of 18 months. This approach is unique in the field of mental health. Typically, online interventions are provided for a median period of 10 weeks[44]. Sustained engagement has been recognized as a long-standing problem, with many patients failing to complete more than one or two sessions in self-guided online interventions, even with weekly telephone support[45]. With the aim of maximizing long-term engagement, the design of Horyzons exploited online social media technology, applied strengths-based approaches and drew on the self-determination theory. Encouragingly, our results showed that 80.2% of Horyzons users logged on for at least 3 months, 47.0% for 9 months or longer, and 29.0% for at least one year. These findings demonstrate the appeal of Horyzons in a difficult to engage cohort.

This study has several strengths. All research assessors and online therapists received regular supervision, including routine checks on interrater reliability and adherence to the therapy model. Significant efforts were made to maintain the masking of group assignment, and we confirmed that blinding was successful. The intervention was delivered in a clinical setting, increasing the clinical validity and generalizability of study methods and results.

The study also has some limitations. First, engagement with Horyzons over the 18-month intervention varied significantly amongst participants, which may moderate treatment efficacy. Moreover, the trial was by necessity single-blind, which may have had an impact on the results. Finally, we cannot rule out the possibility that the outcome of randomization influenced somewhat the discharge process, with young people allocated to the TAU group receiving a more careful discharge plan compared to those in the Horyzons group.

In conclusion, this is the first study to investigate whether a digital intervention is an effective approach to sustaining the benefits of specialist FEP services. While our results did not provide evidence to support the effectiveness of Horyzons in improving social functioning in FEP, baseline functioning was high in our cohort and, contrary to our expectations, remained high in both groups throughout the study. On the other hand, Horyzons was effective in improving vocational/educational attainment (a core aspect of social recovery), reducing visits to emergency services and reducing rates of hospital admissions due to psychosis following discharge from a specialist FEP service (a core target of specialized FEP services). Finally, our data demonstrated that Horyzons was appealing for young people with FEP, with many participants being engaged for sustained periods of time.

Horyzons has now been adapted and successfully piloted in specialized FEP services in the US[46] and Canada[47], with clinical implementation efforts underway in both countries as well as Australia. Ultimately, with specialized FEP services now being available across the US, Canada, Europe, Asia and Australasia, Horyzons holds significant promise as a novel, engaging and sustainable intervention to improve vocational recovery, reduce utilization of emergency services and provide continuous support for young people with FEP beyond specialized care.

ACKNOWLEDGEMENTS

The study was funded by the Mental Illness Research Fund (MIRF, Victorian Government, Australia), the Australian National Health and Medical Research Council (NHMRC), and the Telstra Foundation. M. Alvarez-Jimenez was supported by an investigator grant (APP1177235) from NHMRC. The authors would like to sincerely thank all the young people who participated in this study for their incredible generosity and for being an inspiration to them.

REFERENCES

1. Alvarez-Jimenez M, Priede A, Hetrick SE et al. Risk factors for relapse following treatment for first episode psychosis: a systematic review and meta-analysis of longitudinal studies. Schizophr Res 2012;139:116-28.
2. Alvarez-Jimenez M, Parker AG, Hetrick SE et al. Preventing the second episode: a systematic review and meta-analysis of psychosocial and pharmacological trials in first-episode psychosis. Schizophr Bull 2011;37:619-30.
3. Bertelsen M, Jeppesen P, Petersen L et al. Five-year follow-up of a randomized multicenter trial of intensive early intervention vs standard treatment for patients with a first episode of psychotic illness: the OPUS trial. Arch Gen Psychiatry 2008;65:762-71.
4. Gafoor R, Nitsch D, McCrone P et al. Effect of early intervention on 5-year outcome in non-affective psychosis. Br J Psychiatry 2010;196:372-6.
5. Killackey E, Allott K, Jackson HJ et al. Individual placement and support for vocational recovery in first-episode psychosis: randomised controlled trial. Br J Psychiatry 2019;214:76-82.
6. Gleeson JF, Cotton SM, Alvarez-Jimenez M et al. A randomized controlled trial of relapse prevention therapy for first-episode psychosis patients: outcome at 30-month follow-up. Schizophr Bull 2013;39:436-48.
7. Fowler D, Hodgekins J, French P et al. Social recovery therapy in combination with early intervention services for enhancement of social recovery in patients with first-episode psychosis (SUPEREDEN3): a single-blind, randomised controlled trial. Lancet Psychiatry 2018;5:41-50.
8. Chang WC, Kwong VWY, Lau ESK et al. Sustainability of treatment effect of a 3-year early intervention programme for first-episode psychosis. Br J Psychiatry 2017;211:37-44.
9. Albert N, Melau M, Jensen H et al. Five years of specialised early intervention versus two years of specialised early intervention followed by three years of standard treatment for patients with a first episode psychosis: randomised, superiority, parallel group trial in Denmark (OPUS II). BMJ 2017;356:i6681.
10. Malla A, Joober R, Iyer S et al. Comparing three-year extension of early intervention service to regular care following two years of early intervention service in first-episode psychosis: a randomized single blind clinical trial. World Psychiatry 2017;16:278-86.
11. Norman RM, Manchanda R, Malla AK et al. Symptom and functional out-

comes for a 5 year early intervention program for psychoses. Schizophr Res 2011;129:111-5.
12. Bucci S, Barrowclough C, Ainsworth J et al. Actissist: proof-of-concept trial of a theory-driven digital intervention for psychosis. Schizophr Bull 2018;44:1070-80.
13. Alvarez-Jimenez M, Gleeson J, Bendall S et al. Enhancing social functioning in young people at ultra high risk (UHR) for psychosis: a pilot study of a novel strengths and mindfulness-based online social therapy. Schizophr Res 2018;202:369-77.
14. Ridout B, Campbell A. The use of social networking sites in mental health interventions for young people: systematic review. J Med Internet Res 2018;20: e12244.
15. Vaskinn A, Ventura J, Andreassen OA et al. A social path to functioning in schizophrenia: from social self-efficacy through negative symptoms to social functional capacity. Psychiatry Res 2015;228:803-7.
16. Garland EL, Fredrickson B, Kring AM et al. Upward spirals of positive emotions counter downward spirals of negativity: insights from the broaden-and-build theory and affective neuroscience on the treatment of emotion dysfunctions and deficits in psychopathology. Clin Psychol Rev 2010;30:849-64.
17. Garland EL, Geschwind N, Peeters F et al. Mindfulness training promotes upward spirals of positive affect and cognition: multilevel and autoregressive latent trajectory modeling analyses. Front Psychol 2015;6:15.
18. Hall PL, Tarrier N. The cognitive-behavioural treatment of low self-esteem in psychotic patients: a pilot study. Behav Res Ther 2003;41:317-32.
19. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000;55:68-78.
20. Nakagami E, Hoe M, Brekke JS. The prospective relationships among intrinsic motivation, neurocognition, and psychosocial functioning in schizophrenia. Schizophr Bull 2010;36:935-48.
21. McGorry PD, Edwards J, Mihalopoulos C et al. EPPIC: an evolving system of early detection and optimal management. Schizophr Bull 1996;22:305-26.
22. Alvarez-Jimenez M, Bendall S, Koval P et al. HORYZONS trial: protocol for a randomised controlled trial of a moderated online social therapy to maintain treatment effects from first-episode psychosis services. BMJ Open 2019;9: e024104.
23. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 1987;13:261-76.
24. Overall JE, Gorham DR. The Brief Psychiatric Rating Scale. Psychol Rep 1962;10:799-812.
25. First MB, Spitzer RL, Gibbon M et al. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P). New York: Biometrics Research, New York State Psychiatric Institute, 2002.
26. First MB, Gibbon M, Spitzer RL et al. User's guide for the Structured Clinical Interview for DSM-IV Axis II Personality Disorders: SCID-II. Washington: American Psychiatric Publishing, 1997.
27. Alvarez-Jimenez M, Rice S, D'Alfonso S et al. A novel multimodal digital service (Moderated Online Social Therapy+) for help-seeking young people experiencing mental ill-health: pilot evaluation within a national youth e-mental health service. J Med Internet Res 2020;22:e17155.
28. Lederman R, Wadley G, Gleeson J et al. Moderated online social therapy: designing and evaluating technology for mental Health. ACM Transactions on Computer-Human Interaction 2014;21:5.
29. 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.
30. Addington D, Addington J, Matickatyndale E. Assessing depression in schizophrenia – the Calgary Depression Scale. Br J Psychiatry 1993;163:39-44.
31. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 1996;66:20-40.
32. Sherbourne CD, Stewart AL. The MOS social support survey. Soc Sci Med 1991;32:705-14.
33. Lecomte T, Corbiere M, Laisne F. Investigating self-esteem in individuals with schizophrenia: relevance of the Self-Esteem Rating Scale-Short Form. Psychiatry Res 2006;143:99-108.
34. 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.
35. Diener E, Emmons RA, Larsen RJ et al. The Satisfaction With Life Scale. J Pers Assess 1985;49: 71-5.
36. Richardson J, Elsworth G, Lezzi A et al. Increasing the sensitivity of the AQoL inventory for the evaluation of interventions affecting mental health. Melbourne: Centre for Health Economics, Monash University, 2011.
37. Fleiss JL, Levin B, Paik MC. Statistical methods for rates and proportions. Chichester: Wiley, 2013.
38. Santesteban-Echarri O, Paino M, Rice S et al. Predictors of functional recovery in first-episode psychosis: a systematic review and meta-analysis of longitudinal studies. Clin Psychol Rev 2017;58:59-75.
39. Albert N, Melau M, Jensen H et al. The effect of duration of untreated psychosis and treatment delay on the outcomes of prolonged early intervention in psychotic disorders. NPJ Schizophr 2017;3:34.
40. Dama M, Shah J, Norman R et al. Short duration of untreated psychosis enhances negative symptom remission in extended early intervention service for psychosis. Acta Psychiatr Scand 2019;140:65-76.
41. O'Sullivan S, Schmaal L, D'Alfonso S et al. Characterising usage of a multicomponent digital intervention to predict treatment outcomes in first-episode psychosis: a cluster analysis. Submitted for publication.
42. Iyer SN, Mangala R, Anitha J et al. An examination of patient-identified goals for treatment in a first-episode programme in Chennai, India. Early Interv Psychiatry 2011;5:360-5.
43. Valentine L, McEnery C, O'Sullivan S et al. Young people's experience of a long-term social media-based intervention for first-episode psychosis: qualitative analysis. J Med Internet Res 2020; 22:e17570.
44. Kelders SM, Kok RN, Ossebaard HC et al. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012;14:e152.
45. Gilbody S, Littlewood E, Hewitt C et al. Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): large scale pragmatic randomised controlled trial. BMJ 2015;351:h5627.
46. Ludwig KA, Browne JW, Nagendra A et al. Horyzons USA: a moderated online social intervention for first episode psychosis. Early Interv Psychiatry (in press).
47. Lal S, Gleeson J, Rivard L et al. Adaptation of a digital health innovation to prevent relapse and support recovery in youth receiving services for first-episode psychosis: results from the Horyzons-Canada Phase 1 Study. JMIR Form Res 2020;4:e19887.

DOI:10.1002/wps.20858

Efficacy and acceptability of pharmacological, psychosocial, and brain stimulation interventions in children and adolescents with mental disorders: an umbrella review

Christoph U. Correll[1-4], Samuele Cortese[5-9], Giovanni Croatto[10], Francesco Monaco[11], Damir Krinitski[12], Gonzalo Arrondo[5,13], Edoardo G. Ostinelli[14], Caroline Zangani[15], Michele Fornaro[16], Andrés Estradé[17,18], Paolo Fusar-Poli[17,19-21], Andre F. Carvalho[22,23], Marco Solmi[10,17]

[1]Department of Psychiatry, Zucker Hillside Hospital, Northwell Health, Glen Oaks, New York, NY, USA; [2]Department of Psychiatry and Molecular Medicine, Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, USA; [3]Center for Psychiatric Neuroscience, Feinstein Institute for Medical Research, Manhasset, NY, USA; [4]Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin Berlin, Berlin, Germany; [5]Center for Innovation in Mental Health, School of Psychology, Faculty of Environmental and Life Sciences, University of Southampton, Southampton, UK; [6]Clinical and Experimental Sciences (CNS and Psychiatry), Faculty of Medicine, University of Southampton, Southampton, UK; [7]Solent NHS Trust, Southampton, UK; [8]Hassenfeld Children's Hospital at NYU Langone, New York University Child Study Center, New York, NY, USA; [9]Division of Psychiatry and Applied Psychology, School of Medicine, University of Nottingham, Nottingham, UK; [10]Neurosciences Department, University of Padua, Padua, Italy; [11]Department of Mental Health, ASL Salerno, Salerno, Italy; [12]Integrated Psychiatry Winterthur, Winterthur, Switzerland; [13]Mind-Brain Group, Institute for Culture and Society, University of Navarra, Pamplona, Spain; [14]Psychiatry Department, Oxford University, Oxford, UK; [15]Department of Health Sciences, University of Milan, Milan, Italy; [16]Department of Psychiatry, Federico II University, Naples, Italy; [17]Early Psychosis: Interventions and Clinical-detection (EPIC) Lab, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; [18]Department of Clinical and Health Psychology, Catholic University, Montevideo, Uruguay; [19]OASIS Service, South London and Maudsley NHS Foundation Trust, London, UK; [20]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy; [21]Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK; [22]Department of Psychiatry, University of Toronto, Toronto, ON, Canada; [23]Centre for Addiction and Mental Health (CAMH), Toronto, ON, Canada

Top-tier evidence on the safety/tolerability of 80 medications in children/adolescents with mental disorders has recently been reviewed in this journal. To guide clinical practice, such data must be combined with evidence on efficacy and acceptability. Besides medications, psychosocial interventions and brain stimulation techniques are treatment options for children/adolescents with mental disorders. For this umbrella review, we systematically searched network meta-analyses (NMAs) and meta-analyses (MAs) of randomized controlled trials (RCTs) evaluating 48 medications, 20 psychosocial interventions, and four brain stimulation techniques in children/adolescents with 52 different mental disorders or groups of mental disorders, reporting on 20 different efficacy/acceptability outcomes. Co-primary outcomes were disease-specific symptom reduction and all-cause discontinuation ("acceptability"). We included 14 NMAs and 90 MAs, reporting on 15 mental disorders or groups of mental disorders. Overall, 21 medications outperformed placebo regarding the co-primary outcomes, and three psychosocial interventions did so (while seven outperformed waiting list/no treatment). Based on the meta-analytic evidence, the most convincing efficacy profile emerged for amphetamines, methylphenidate and, to a smaller extent, behavioral therapy in attention-deficit/hyperactivity disorder; aripiprazole, risperidone and several psychosocial interventions in autism; risperidone and behavioral interventions in disruptive behavior disorders; several antipsychotics in schizophrenia spectrum disorders; fluoxetine, the combination of fluoxetine and cognitive behavioral therapy (CBT), and interpersonal therapy in depression; aripiprazole in mania; fluoxetine and group CBT in anxiety disorders; fluoxetine/selective serotonin reuptake inhibitors, CBT, and behavioral therapy with exposure and response prevention in obsessive-compulsive disorder; CBT in post-traumatic stress disorder; imipramine and alarm behavioral intervention in enuresis; behavioral therapy in encopresis; and family therapy in anorexia nervosa. Results from this umbrella review of interventions for mental disorders in children/adolescents provide evidence-based information for clinical decision making.

Key words: Children, adolescents, pharmacotherapy, psychotherapies, psychosocial interventions, brain stimulation, ADHD, autism, disruptive behavior disorders, efficacy, acceptability

(World Psychiatry 2021;20:244–275)

Many mental disorders have an onset with clinically relevant manifestations in childhood or adolescence, followed frequently by a chronic illness course into adulthood[1,2]. Many disorders with an earlier onset are first diagnosed in adulthood, with a delay ranging for example from 6 to 8 years for mood disorders and from 9 to 23 years for anxiety disorders[3]. Due to their interference with attainment of biopsychosocial milestones, mental and neurodevelopmental disorders in children and adolescents are among the leading causes of global burden of disease and years lived with disability[4]. This situation makes the appropriate delivery of evidence-based and effective treatments for youth with mental disorders a key priority in the public health field.

Pharmacological, psychosocial and brain stimulation options are available for the management of many mental disorders in children and adolescents. However, for several of them, what should be considered the first line treatment strategy – based on efficacy, effectiveness, acceptability and tolerability/safety – remains uncertain.

A number of randomized controlled trials (RCTs) have been conducted to assess the efficacy, acceptability and tolerability of medications across different disorders in children and adolescents. The results from many of these RCTs have been pooled in pairwise meta-analyses (MAs) or network meta-analyses (NMAs)[5–8]. While most antidepressants outperform placebo to treat depression in adults[9], most antidepressants have not been shown to be superior to placebo in children and adolescents with major depressive disorder[7,10]. Similarly, yet to a lower extent, antidepressants may not be as effective in children and adolescents with anxiety disorders as in adults[11].

On the other hand, RCTs comparing psychosocial interventions with waiting list or no intervention control groups generally show a large effect size in youth with depression[10] or anxiety[12]

disorders. Yet, when compared with placebo/sham interventions, most significant findings favoring psychosocial interventions vs. placebo disappear[10,12]. Effect sizes also vary according to design, blinding, patient selection (baseline severity) and choice of the control group[13] in trials assessing combination treatments, whose superiority to monotherapies has not been consistently confirmed within and across disorders in children/adolescents.

Differences in inclusion criteria, outcomes, and a variety of features defining quality across MAs and NMAs limit the clinical value and impact of such a rich, yet complex body of evidence. Umbrella reviews may overcome these problems to some degree by taking the totality of the evidence from existing MAs and NMAs into account, and filtering top-tier meta-analytic estimates according to pre-established criteria. It is paramount to provide clinicians with structured and standardized summaries, translating the massive data into actionable clinical information.

To our knowledge, no umbrella review is available of the evidence from MAs and NMAs of RCTs on the efficacy and acceptability of pharmacological, psychosocial, and brain stimulation treatment options for the core symptoms and associated problems of the full range of mental disorders in children and adolescents. The present study aims to fill this gap, as previously done in this journal concerning the safety and tolerability of 80 pharmacological agents used for the management of child and adolescent mental disorders[14].

We focused on disease-specific symptom reduction and treatment response as efficacy measures, and on measures of acceptability that could be compared across the three different treatment modalities, namely all-cause discontinuation and intolerability-related discontinuation. Following this approach, this umbrella review intends to provide practitioners with an evidence-based atlas of therapeutic tools to inform clinical decision making, where a balance needs to be struck between efficacy, acceptability/tolerability, and safety.

METHODS

Search, inclusion and exclusion criteria

This umbrella review followed an *a priori* protocol (available upon request). We conducted a systematic search in PubMed, PsycINFO, and Cochrane database up to January 9, 2021, using an exhaustive combination of key words (full search string available upon request). We also manually searched bibliographies of included meta-analyses. Two independent authors conducted title/abstract screening, full-text assessment, and data extraction into a pre-defined excel spreadsheet. A third author triple-checked extracted data, and resolved any conflict.

Included were: a) NMAs or MAs of RCTs, b) of *a priori* defined 48 psychotropic medications, 20 psychosocial interventions, and four brain stimulation interventions, c) in children and/or adolescents, d) with any of 52 *a priori* defined mental disorders, e) reporting on 20 *a priori* defined outcomes within a specific disorder. Exclusion criteria were: a) systematic reviews without meta-analysis, b) pooling of studies other than RCTs, c) interventions for other than pre-defined disorders/outcomes.

Whenever two NMAs or MAs reported on the same combination of disorder, intervention, comparison and outcome, we considered the comparison with more RCTs, the minimum being at least one direct comparison for NMAs.

Included disorders, interventions, and comparisons

Mental disorders of interest, as grouped in the ICD-11[15], were: a) neurodevelopmental disorders (autism spectrum disorder, attention-deficit/hyperactivity disorder (ADHD), disorders of intellectual development, developmental speech or sound disorders, developmental learning disorders, developmental motor coordination disorders), b) schizophrenia and other primary psychotic disorders (schizophrenia, schizoaffective disorder, schizotypal disorder, acute and transient psychotic disorder), c) catatonia, d) mood disorders (bipolar and related disorders, depressive disorders), e) anxiety or fear-related disorders (generalized anxiety disorder, panic disorder, agoraphobia, specific phobia, social anxiety disorder, separation anxiety disorder, selective mutism), f) obsessive-compulsive and related disorders (obsessive-compulsive disorder, body dysmorphic disorder, body-focused repetitive disorders), g) movement disorders (Tourette's disorder, other tic disorder), h) disorders specifically associated with stress (post-traumatic stress disorder (PTSD), complex PTSD, prolonged grief disorder, reactive attachment disorder, disinhibited social engagement disorder), i) dissociative disorders (dissociative neurological symptom disorder, dissociative amnesia, trance disorder, dissociative identity disorder), j) feeding and eating disorders (anorexia nervosa, bulimia nervosa, binge eating disorder, avoidant-restrictive food intake disorder, pica, rumination-regurgitation disorder), k) elimination disorders (enuresis, encopresis), l) disorders of bodily distress or bodily experience (bodily distress disorder, body integrity dysphoria), m) disorders due to substance use or addictive behaviors, n) impulse control disorders (pyromania, kleptomania, compulsive sexual behavior disorder, intermittent explosive disorder), o) disruptive behavior or dissocial disorders (oppositional defiant disorder, conduct disorder).

Interventions included medications, psychosocial interventions, and brain stimulation techniques.

Medications comprised antidepressants (bupropion, mirtazapine, nefazodone, vilazodone, desvenlafaxine, duloxetine, venlafaxine, citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, sertraline, clomipramine, desipramine, imipramine, nortriptyline, amitriptyline); antipsychotics (fluphenazine, haloperidol, molindone, trifluoperazine, amisulpride, aripiprazole, asenapine, clozapine, loxapine, lurasidone, olanzapine, paliperidone, quetiapine, risperidone, thioridazine, ziprasidone); anti-ADHD medications (amphetamines, atomoxetine, clonidine, guanfacine, methylphenidate, modafinil); mood stabilizers (carbamazepine, lamotrigine, lithium, oxcarbazepine, topiramate, valproate); and others (oxybutynin, desmopressin).

Psychosocial interventions included behavioral therapy, cognitive behavioral therapy (CBT), problem solving, dialectical behavioral therapy, family-based therapy, interpersonal psychotherapy, mentalization based therapy, psychodynamic psychotherapy, supportive therapy, social skills training, acceptance and commitment therapy, mindfulness, eye movement desensitization and reprocessing, narrative exposure therapy, cognitive remediation therapy, cognitive training, parent-child interaction therapy, play therapy, art therapy, and occupational therapy.

Brain stimulation interventions included transcranial magnetic stimulation, transcranial direct current stimulation, electroconvulsive therapy, and neurofeedback.

Comparators were labeled as active drug, active psychosocial intervention, treatment as usual (TAU)/low intensity psychosocial intervention, waiting list/no treatment, or placebo/sham.

Outcomes

Co-primary outcomes were disease-specific primary symptom reduction and all-cause discontinuation ("acceptability").

Secondary continuous outcomes were measures of aggressive behavior, anxiety (other than anxiety disorders), cognition (other than ADHD), depressive symptoms (other than depressive episode/disorder), irritability, suicidal ideation, global illness severity, functioning (as defined by authors), and quality of life.

Secondary categorical outcomes were study-defined treatment response, remission, relapse, hospitalization, discontinuation due to inefficacy, discontinuation due to intolerability, suicide attempt, completed suicide, and death. When available, treatment estimates from clinicians, teachers, parents, and children/adolescents were considered separately.

Quality of evidence

The quality of MAs and NMAs was measured using A Measurement Tool for the Assessment of Multiple Systematic Reviews (AMSTAR-PLUS)[16,17] to quantify both the methodological quality of MAs and NMAs with the first 11 items (AMSTAR) and of included RCTs with six additional items (AMSTAR-Content).

Methodological quality was categorized into low (<4), medium (4-7), and high (>7). Content quality was categorized into low (<4), medium (4-6), and high (>6). The lowest score between methodological and content quality determined the overall MA or NMA quality.

Statistical analysis

We converted continuous non-standardized outcomes, such as weighted mean differences, to standardized mean differences (SMDs), and binary outcomes to odds ratio (ORs) with Comprehensive Meta-Analysis (CMA), Version 3[18]. We then calculated the mean SMD for the primary efficacy outcome across pharmacological, psychosocial, and brain stimulation interventions for each disorder against placebo/sham and waiting list/no intervention, as well as for active controlled monotherapy and combination treatment studies, prioritizing clinician rating, followed by teacher, parent, and then subject-rated estimates. For treatment response, in case no data were available for the continuous primary efficacy outcome, we converted ORs to SMDs, using CMA.

Whenever data conversion was not possible, we kept the original effect sizes as reported. Whenever we included data from meta-analyses that used fixed-effects models, we recalculated the meta-analysis using random-effects models[19]. For consistent and easy comparison, we harmonized effect sizes as follows: SMD<0 favors intervention, OR/risk ratio (RR) <1 favors intervention for discontinuation, suicide or relapse, while OR/RR>1 favors intervention for response or remission.

RESULTS

Search results and literature coverage

The search process is described in Figure 1. Out of 5,231 initial hits, we assessed 910 MAs and NMAs at full text level. Of these, we excluded 806, with specific reasons (list available upon request). The list of all included MAs and NMAs is available in Table 1, also indicating the number of included RCTs and participants, as well as the methodological quality (AMSTAR score) together with the quality of included RCTs (AMSTAR-Content median score).

We ultimately included 14 NMAs and 90 MAs, reporting on 15 disorders or groups of disorders. For ADHD, we included three NMAs[5,20,21] and 21 MAs[22-42]; for autism, one NMA[43] and 21 MAs[12,44-63] (including one focusing on comorbid anxiety disorders and autism)[12]; for depressive disorders, two NMA[7,10] and seven MAs[64-70]; for obsessive-compulsive disorder, one NMA[71] and six MAs[72-77]; for anxiety disorders, two NMAs[11,78] and five MAs[12,79-82] (plus two MAs specific on social anxiety disorder[83,84]); for enuresis, one NMA[85] and six MAs[86-91], for disruptive behavior/dissocial/conduct disorders, five MAs[92-96] (plus one focusing on youth with comorbid ADHD)[25]; for eating disorders, one NMA[97] and four MAs[98-101]; for schizophrenia spectrum disorders, three NMAs[8,102,103] and two MAs[104,105]; for bipolar disorder, four MAs[106-109]; for tic disorder, two MAs[110,111]; for Tourette's disorder, two MAs[112,113]; for encopresis, two MAs[114,115]; for developmental coordination disorder, one MA[116]; and for PTSD, one MA[117].

Overall, 85.4% of *a priori* selected medications were covered for at least one of the two co-primary outcomes, which was the case for 55% of the psychosocial interventions, and 25% of the brain stimulation interventions. Moreover, 70% of *a priori* selected outcomes were covered across monotherapy medication treatments (anti-ADHD medications: 65%; antidepressants: 55%; antipsychotics: 40%; mood stabilizers: 25%), 80% across psychosocial interventions, and 20% across brain stimulation interventions.

Among monotherapy medication treatments with data on co-primary outcomes, those most covered by the literature were atomoxetine (11 outcomes), methylphenidate (9 outcomes), amphetamines and risperidone (8 outcomes), aripiprazole, fluoxe-

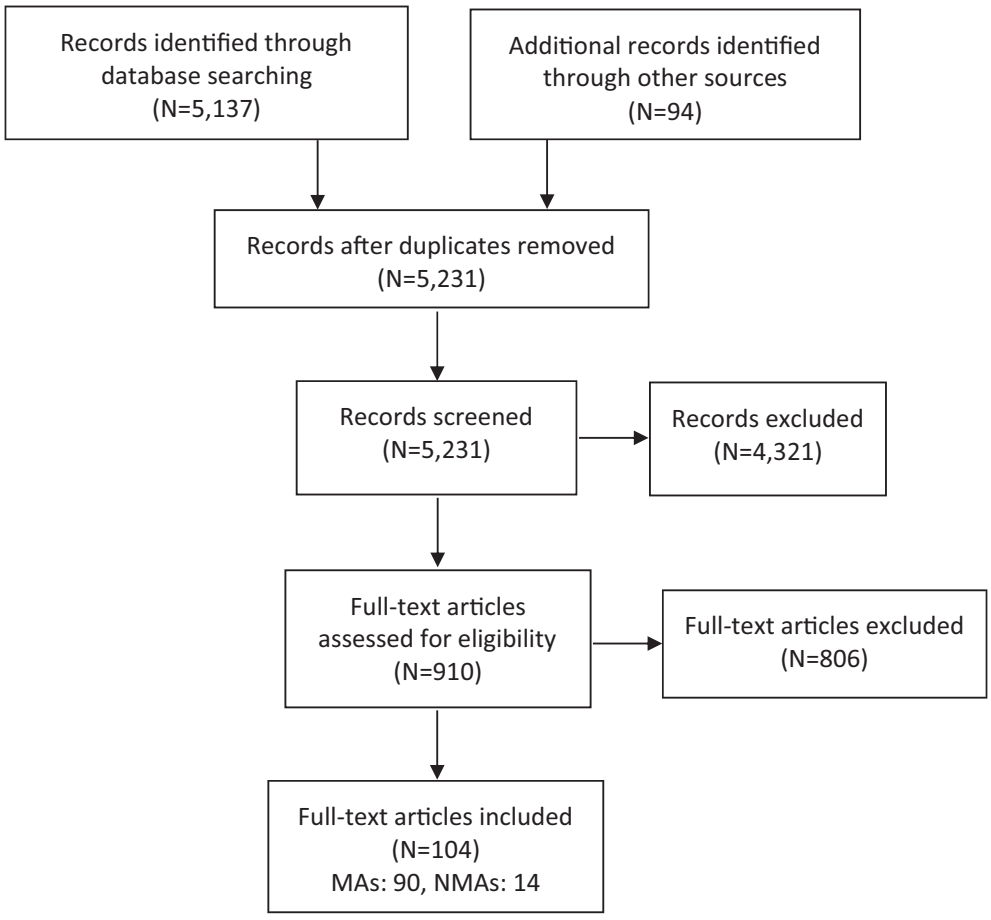

Figure 1 PRISMA flow chart, MAs – meta-analyses, NMAs – network meta-analyses

tine, guanfacine, lurasidone and quetiapine (7 outcomes), and asenapine, clonidine, olanzapine, paliperidone and sertraline (6 outcomes). Monotherapy psychosocial interventions most covered by the literature were CBT (12 outcomes), behavioral therapy (9 outcomes), parent-child interaction therapy (7 outcomes), and CBT-oriented, psychodynamic-oriented and family-based therapies (6 outcomes). Among brain stimulation interventions, neurofeedback was the only modality with data that could be included in this umbrella review (4 outcomes).

Quality of included evidence

Among 14 NMAs of RCTs, the median AMSTAR score was 9.5 (interquartile range, IQR: 7-11), and the median AMSTAR-Content score was 4 (IQR: 2.75-5). The median overall quality score across all effect sizes was low in six NMAs (42.9%), moderate in six (42.9%), high in the remaining two (14.2%).

Among 90 MAs of RCTs, the median AMSTAR score was 9 (IQR: 7-10) and the median AMSTAR-Content score was 2 (IQR: 1-3). The median overall quality score across all effect sizes was low in 71 MAs (78.9%), moderate in 19 (21.1%), and high in none.

Across NMAs and MAs of RCTs of medications, the median AMSTAR quality score was 10 (IQR: 7-11), being low in 0.8%, moderate in 24.7%, and high in 74.4% of the NMAs/MAs, while the AMSTAR-Content median quality score was 4 (IQR: 3-5), being low in 30.1%, moderate in 58.6%, and high in 11.3%.

Across NMAs and MAs of RCTs of psychosocial interventions, the median AMSTAR quality score was 11 (IQR: 10-12), being low in none of the NMAs/MAs, moderate in 8.2%, and high in 91.8%, while the median AMSTAR-Content quality score was 2 (IQR: 1-3), being low in 87.4%, moderate in 12.6%, and high in none.

Across brain stimulation interventions, the median AMSTAR quality score was 9 (IQR: 8-10), being low in none of the NMAs/MAs, medium in 16.7%, and high in 83.3%, while the median AMSTAR-Content quality score was 2 (IQR: 1-4), being low in 66.7%, moderate in 33.3%, and high in none.

Efficacy, acceptability and tolerability of pharmacological, psychosocial, and brain stimulation interventions (Tables 2-7)

ADHD

Results for ADHD are shown in Tables 2, 6 and 7. Amphetamines, methylphenidate, desipramine and modafinil had the largest effect size for the primary efficacy outcome.

Table 1 Network and pairwise meta-analyses of randomized controlled trials (RCTs) of pharmacological, psychosocial and brain stimulation interventions in children and adolescents with mental disorders included in the umbrella review

	Source	Number of RCTs/ patients	Intervention	Controls	Outcomes	A	C
Anxiety disorders							
Wang et al[79]	MA	115/7,719	AD	PBO	PE, REM	10	4
Dobson et al[11]	NMA	22/2,623	AD	PBO	RES, ACD, AED, S	7	5
Zhang et al[80]	MA	7/358	CB	WL/NT	PE	9	2
James et al[12]	MA	87/5,964	CB	PBO, WL/NT, TAU, PS	PE, REM, DEP, F, ACD	11	3
Zhou et al[78]	NMA	101/6,625	CB	PBO, WL/NT, TAU, PS	PE, QoL, ACD	11	2
Sigurvinsdóttir et al[81]	MA	81/5,913	CB	WL/NT, TAU, PS	REM	10	1
James et al[82]	MA	41/1,955	CB	TAU, PS	PE, REM	11	1.5
Anorexia nervosa							
Fisher et al[99]	MA	21/1,407	FB	TAU, PS	PE, ACD, REM	10	1
van den Berg et al[100]	MA	15/1,279	PS	TAU	PE	9	2
Zeeck et al [97]	NMA	18/1,247	FB, PSD-O	PS	PE	7	1
Social anxiety disorder							
Yang et al[83]	MA	17/1,134	CB	PBO, WL/NT	PE, REM, DEP, QoL, ACD	10	2
Kreuze et al[84]	MA	42/3,239	CB	PBO, TAU, LIP	AG, F	10	2.5
Attention-deficit/hyperactivity disorder (ADHD)							
Cortese et al[5]	NMA	133/18,199	AD, STIM, α2	PBO, AD, STIM	PE, AED, GLO	11	9
Otasowie et al[22]	MA	6/216	AD	PBO	PE, GLO	10	3
Punja et al [23]	MA	23/2,675	STIM	PBO	PE, COG, GLO	10	4
Stuhec et al[34]	MA	28/4,699	AD	PBO	PE	8	2
Luan et al[21]	NMA	73/15,025	AD, STIM, α2	PBO, PHARMA	PE, AED, ID	7	4
Catalá-López et al[20]	NMA	190/26,114	AP, AD, STIM, α2, CB, CT, NF, COMB	PBO	RES, ACD, GLO	10	4
Schachter et al[36]	MA	62/2,897	STIM	PBO	AG	9	1
Schwartz et al[37]	MA	25/3,928	AD, STIM	PBO	AG, F, QoL, S	7	5
Coghill et al[38]	MA	60/1,993	STIM	PBO	COG	8	2
Storebø et al[39]	MA	185/12,245	STIM	PBO	QoL	8	5
Bangs et al[40]	MA	32/7,248	AD, STIM	PBO	S	3	4
Hirota et al[41]	MA	12/2,276	α2+	PBO	PE, ACD, AED, ID	6	3.5
Storebø et al[42]	MA	25/2,690	SKILL, COMB	WL/NT	PE, COG, F	11	2
Sun et al[24]	MA	8/423	STIM	PBO	PE, ACD, AED	11	2
Battagliese et al[25]	MA	24/1,690	BT	MIX	PE, AG, COG, F	7	1
Faraone et al[26]	MA	4/216	STIM	STIM	AG	2	3
Van Doren et al[27]	MA	10/506	NF	PHARMA, PS	PE, RES, ACD	8	2
Cortese et al[28]	MA	16/759	CT	MIX	PE, COG	11	1
Daley et al[29]	MA	32/2,077	BT	MIX	PE, COG	9	2
Bikic et al[30]	MA	12/1,054	SKILL	MIX	PE, COG	8	2
Mulqueen et al[31]	MA	8/399	BT	MIX	PE	6	1
Cortese et al[32]	MA	13/520	NF	MIX	PE, COG	9	1.5
Bussalb et al[33]	MA	16/706	NF	MIX	PE	4	2
Faraone et al[35]	MA	7/384	STIM	PBO	AG	2	2

Table 1 Network and pairwise meta-analyses of randomized controlled trials (RCTs) of pharmacological, psychosocial and brain stimulation interventions in children and adolescents with mental disorders included in the umbrella review *(continued)*

	Source	Number of RCTs/ patients	Intervention	Controls	Outcomes	A	C
Autism spectrum disorder							
Maneeton et al[44]	MA	3/408	AP	PBO	PE, RES, GLO	7	4
Maneeton et al[52]	MA	7/372	AP	PBO	REL, RES	7	3.5
Zhou et al[53]	MA	64/3,499	STIM	PBO	PP	9	3
Murza et al[54]	MA	16/837	SKILL	WL/NT	F	8	0.5
Fletcher-Watson et al[56]	MA	22/695	SKILL	WL/NT, TAU	F	10	1
Sturman et al[55]	MA	4/113	STIM	PBO	PE	10	1
Cohen et al[57]	MA	15/995	AP	PBO	RES	5	1
Hirota et al[58]	MA	7/171	MS	PBO	RES, AG, ACD, AED, ID	6	4
Fallah et al[43]	NMA	8/878	AP	PBO, AP	AG	7	1
D'Alò et al[59]	MA	15/1,124	AP	PBO	ACD, AED	9	5
Ospina et al[60]	MA	69/2,585	BT	WL/NT, PS	PE	9	1
Reichow et al[61]	MA	5/196	SKILL	WL/NT	PE	10	1
James et al[12]	MA	87/5,964	CB	WL/NT, TAU	ANX	11	0.5
Tachibana et al [62]	MA	32/594	PS	TAU	PE	11	1
Nevill et al[63]	MA	19/1,205	PCI	TAU/LIP, MIX	PE, COG	5	1
Yu et al[45]	MA	14/555	BT	TAU	PE, F	9	0
Oono et al[46]	MA	17/919	PCI	MIX	PE, F, GLO	10	1
Parsons et al[47]	MA	21/925	SKILL	MIX	PE	9	1
Kreslins et al[48]	MA	10/470	CB	MIX	ANX	9	0
Tarver et al[49]	MA	9/521	PCI	MIX	AG	8	2
Soares et al[50]	MA	18/1,266	SKILL	MIX	F	8	2
Postorino et al[51]	MA	8/653	PCI	MIX	IR	8	1
Bipolar disorder, depressive episode							
Maneeton et al[106]	MA	3/251	AP	PBO	PE, RES, REM, GLO, ACD, AED	9	3
Bipolar disorder, manic episode							
Meduri et al[107]	MA	22/5,437	AP	PBO	PE, RES, ACD, AED, ID	10	5
Liu et al[108]	MA	46/2,666	MS	PBO	RES	7	6
Jochim et al[109]	MA	25/3,252	MS, AP	PBO, MS	ACD	10	4
Bulimia nervosa							
Linardon et al[101]	MA	79/NR	CB	PS	PE	6	0
Depressive disorders							
Zhou et al[10]	NMA	71/9,510	AD, PSD-O, FB, CB, COMB	PBO, WL/NT, TAU/LIP, PHARMA, PS	PE, ACD, S	11	5
Cipriani et al[7]	NMA	34/5,260	AD	PBO, PHARMA	RES, AED	11	5
Spielmans & Gerwig[64]	MA	8/1,756	AD	PBO	QoL	5	5
Kato et al[65]	MA	40/8,890	AD	PBO	REL	9	3
Whittington et al[66]	MA	2/376	AD	PBO	REM	9	2.5
Watanabe et al[67]	MA	27/1,744	PSD-O	WL/PBO	RES	7	2
Cox et al[68]	MA	9/882	AD, CB, COMB	PHARMA, PS	REM, S	10	3

Table 1 Network and pairwise meta-analyses of randomized controlled trials (RCTs) of pharmacological, psychosocial and brain stimulation interventions in children and adolescents with mental disorders included in the umbrella review *(continued)*

	Source	Number of RCTs/ patients	Intervention	Controls	Outcomes	A	C
Dubicka et al[69]	MA	5/1,206	COMB	PHARMA, PS	RES, F, S	7	3
Klein et al[70]	MA	11/809	CB	MIX	PE	8	4
Disruptive behavior/dissocial/conduct disorders							
Seida et al[92]	MA	62/NR	AP	PBO	PE, AG, GLO	9	3.5
Loy et al[93]	MA	10/896	AP	PBO	PE, AG	10	4
Pringsheim et al[94]	MA	18/1,195	MS	PBO	AG	10	2
Ipser & Stein[95]	MA	14/823	PHARMA	PBO	AG, ACD, GLO, RES	6	1.5
Battagliese et al[25]	MA	24/1,690	CB	WL/NT, MIX	PE	7	1.5
McQuire et al[96]	MA	14/912	AP, MS	PBO	AG	8	2
Developmental coordination disorder							
Miyahara et al[116]	MA	15/649	SKILL	WL/NT	PE	10	1
Eating disorders							
Couturier et al[98]	MA	6/369	FB	PS	REM	8	3
Encopresis							
Freeman et al[114]	MA	10/562	COMB	TAU	PE, RES	7	1
Brazzelli et al[115]	MA	21/1,371	COMB	TAU	RES	10	1
Enuresis							
Caldwell et al[86]	MA	74/5,983	BT, COMB	PHARMA, PS, WL/NT	PE, RES	11	1
Caldwell et al[87]	MA	64/4,071	AD, COMB	PBO, PHARMA, PS	PE, RES	11	1
Caldwell et al[88]	MA	16/1,643	BT	PS, WL/NT	RES	10	1
Buckley et al[89]	MA	27/1,803	SKILL, COMB	TAU, PHARMA	REM	10	1
Deshpande et al[90]	MA	40/2,440	AD, COMB	PHARMA	RES, REL	10	1
Peng et al[91]	MA	15/1,502	PHARMA	PS	ACD	9	4
Song et al[85]	NMA	18/1,649	PHARMA, COMB	PHARMA, PS	RES, REL	9	4
Obsessive-compulsive disorder							
Skapinakis et al[71]	NMA	86/15,585	AD, CB, COMB	PBO, WL/NT, PHARMA, PS	PE, ACD	10	3
Maneeton et al[72]	MA	3/188	AD	PBO	RES, GLO	9	2
McGuire et al[73]	MA	20/1,296	AD, CB	PBO, TAU/LIP, WL/NT	RES, REM	8	1
Locher et al[74]	MA	36/6,778	AD	PBO	AED	10	4
Geller[75]	MA	12/1,044	AD	PBO	GLO	8	3
Uhre et al [76]	MA	12/791	CB, AD	PBO, WL/NT, PS	REM, F, QoL	9	1
Johnco et al [77]	MA	21/1,423	CB, AD	PBO, WL/NT, TAU/LIP, PS	ACD	6	1
Post-traumatic stress disorder							
Gillies et al[117]	MA	14/758	CB	WL/NT, TAU/LIP	PE, RES, ANX, DEP, ACD	10	1
Schizophrenia spectrum disorders							
Krause et al[102]	NMA	28/3,003	AP	PBO, PHARMA	PE, RES, ACD, ID	11	3
Arango et al [103]	NMA	13/2,210	AP	PBO, PHARMA	GLO, AED	9	7
Pagsberg et al[8]	NMA	12/2,158	AP	PBO, PHARMA	GLO	8	3

Table 1 Network and pairwise meta-analyses of randomized controlled trials (RCTs) of pharmacological, psychosocial and brain stimulation interventions in children and adolescents with mental disorders included in the umbrella review *(continued)*

	Source	Number of RCTs/ patients	Intervention	Controls	Outcomes	A	C
Sarkar & Grover[104]	MA	15/995	AP	PHARMA	PE	5	1
Kumar et al[105]	MA	13/1,112	AP	PHARMA	AED	8	1
Tic disorder							
Bloch et al[110]	MA	9/477	STIM, AD	PBO	PE	4	1
Yu et al[111]	MA	15/1,070	MS	PHARMA	RES	7	3
Tourette's disorder							
Hollis et al[112]	MA	40/2,422	AP, α2, STIM, BT	PBO, MIX	PE	8	1
Zheng et al[113]	MA	6/528	AP	PHARMA	PE	10	2

MA – meta-analysis, NMA – network meta-analysis, A – AMSTAR, C – AMSTAR-Content (median), AD – antidepressants, CB – cognitive-based, FB – family-based, PS – active psychosocial, PSD-O – psychodynamic-oriented, STIM – stimulants, α2 – α2-agonists (+=augmentation with), AP – antipsychotics, CT – cognition-targeted, NF – neurofeedback, COMB – combination of more than one treatment, SKILL – skills training, BT – behavioral treatment, MS – mood stabilizers, PCI – parent-child interaction, PHARMA – mixed medications, PBO – placebo, WL – waiting list, NT – no treatment, TAU – treatment as usual, LIP – low-intensity psychosocial intervention, MIX – mixed active/inactive control group, PE – primary efficacy outcome, REM – remission, REL – relapse, RES – response, S – suicidality, ACD – all-cause discontinuation, AED – discontinuation due to adverse events, ID – discontinuation due to inefficacy, DEP – depressive symptoms, ANX – anxiety symptoms, AG – aggressivity, QoL – quality of life, GLO – global illness severity, COG – cognition, F – functioning, NR – not reported

Focusing on the two best interventions, amphetamines had the highest effect size based on the clinician-rated primary efficacy outcome vs. placebo (large effect size), and were superior to placebo also regarding response (large effect size), aggressive behavior (large effect size), academic functioning (medium effect size), global illness severity (large effect size), and less discontinuation due to inefficacy (large effect size), without significant differences regarding all-cause discontinuation ("acceptability") or discontinuation due to intolerability (see Table 2).

Methylphenidate had medium to large effect sizes regarding the primary efficacy outcome vs. placebo across different raters, and was superior to placebo regarding other-than-attention cognition broadly (small to medium effect size), global illness improvement (large effect size), quality of life (medium effect size), acceptability (small effect size), and less discontinuation due to inefficacy (medium effect size), without significant differences concerning discontinuation due to intolerability. The efficacy of methylphenidate was also confirmed in youth with comorbid intellectual disability (see Table 2).

Clonidine, guanfacine and atomoxetine were also effective regarding the primary efficacy outcome, but with less consistent results across raters. Among psychosocial interventions, social skills training improved the primary efficacy outcome and functioning (small to medium effect size); however, the control group was waiting list/no treatment. Only behavioral therapy outperformed placebo for response (small effect size), impact on global illness severity (small effect size), and acceptability (small effect size). Neurofeedback did not show any significant efficacy outcome, nor any difference emerged on acceptability (see Table 2).

Alpha-2 agonists were an effective augmentation strategy when added to stimulants vs. placebo (small effect size). Importantly, combined interventions, and specifically methylphenidate with parent training or with clonidine, and atomoxetine with parent training, showed large effect sizes regarding response vs. placebo (see Table 2). Additionally, behavioral therapy plus stimulants was superior both to behavioral therapy alone and to stimulants alone regarding response (large effect size), without any differences in acceptability (see Table 6).

In head-to-head comparisons, amphetamines outperformed methylphenidate, which outperformed bupropion (large effect sizes) and atomoxetine (small effect size) on the primary efficacy outcome. Amphetamines were superior to atomoxetine in reducing discontinuation due to inefficacy, and better than methylphenidate for aggressive behavior (small effect size), while methylphenidate was superior to atomoxetine regarding acceptability (medium effect size), and to guanfacine regarding less discontinuation due to intolerability (medium effect size). Stimulants were superior to neurofeedback regarding cognition, and neurofeedback outperformed cognitive training on acceptability (see Table 6).

Autism spectrum disorder

Results for autism spectrum disorder are shown in Tables 2, 5, 6 and 7.

Aripiprazole was superior to placebo regarding the primary efficacy outcome, as well as response, aggressive behavior, global illness severity, and acceptability (all small effect sizes). Risperidone showed the same profile, yet with a large effect size regarding response. Both aripiprazole and risperidone were not different from placebo concerning discontinuation due to intolerability (see Table 2).

Table 2 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with neurodevelopmental and disruptive behavior/dissocial/conduct disorders

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/patients	Q
Attention-deficit/hyperactivity disorder (ADHD)					
Pharmacological interventions					
Efficacy (clinician-rated)	Amphetamines	**SMD=–1.02 (–1.19 to –0.85)**	PBO/Sham	46/9,926	H
	Methylphenidate	**SMD=–0.78 (–0.93 to –0.62)**	PBO/Sham	46/9,926	H
	Clonidine	**SMD=–0.71 (–1.17 to –0.24)**	PBO/Sham	46/9,926	H
	Guanfacine	**SMD=–0.67 (–0.85 to –0.50)**	PBO/Sham	46/9,926	H
	Modafinil	**SMD=–0.62 (–0.84 to –0.41)**	PBO/Sham	46/9,926	H
	Atomoxetine	**SMD=–0.56 (–0.66 to –0.45)**	PBO/Sham	46/9,926	H
Efficacy (teacher-rated)	Desipramine	**SMD=–0.97 (–1.66 to –0.28)**	PBO/Sham	2/89	L
	Methylphenidate	**SMD=–0.82 (–1.16 to –0.48)**	PBO/Sham	16/1,843	H
	Modafinil	**SMD=–0.76 (–1.15 to –0.37)**	PBO/Sham	16/1,843	H
	Amphetamines	**SMD=–0.55 (–0.83 to –0.27)**	PBO/Sham	5/745	M
	Guanfacine	SMD=–0.63 (–1.62 to 0.35)	PBO/Sham	16/1,843	H
	Atomoxetine	SMD=–0.32 (–0.82 to 0.18)	PBO/Sham	16/1,843	H
Efficacy (parent-rated)	Desipramine	**SMD=–1.42 (–1.99 to –0.85)**	PBO/Sham	2/99	L
	Amphetamines	**SMD=–1.07 (–1.36 to –0.79)**	PBO/Sham	23/3,796	H
	Methylphenidate	**SMD=–0.84 (–0.95 to –0.72)**	PBO/Sham	23/3,796	H
	Atomoxetine	**SMD=–0.60 (–0.71 to –0.50)**	PBO/Sham	23/3,796	H
	Modafinil	**SMD=–0.46 (–0.61 to –0.31)**	PBO/Sham	23/3,796	H
	Bupropion	SMD=–0.32 (–0.69 to 0.05)	PBO/Sham	2/124	L
	Guanfacine	SMD=–0.23 (–0.90 to 0.45)	PBO/Sham	23/3,796	H
Efficacy (mixed-rated)	Atomoxetine	**SMD=–0.17 (–0.23 to –0.11)**	PBO/Sham	36/7,579	M
	Amphetamines	**SMD=–0.18 (–0.28 to –0.09)**	PBO/Sham	36/7,579	M
	Methylphenidate	**SMD=–0.14 (–0.21 to –0.08)**	PBO/Sham	36/7,579	M
	Guanfacine	**SMD=–0.16 (–0.26 to –0.05)**	PBO/Sham	36/7,579	M
	Clonidine	SMD=–0.10 (–0.23 to 0.03)	PBO/Sham	36/7,579	M
Response	Desipramine	**OR=36.76 (9.17-214)**	PBO/Sham	113/19,398	M
	Amphetamines	**OR=7.45 (5.1-11.09)**	PBO/Sham	113/19,398	M
	Modafinil	**OR=5.51 (3.04-10.32)**	PBO/Sham	113/19,398	M
	Methylphenidate	**OR=5.26 (4.09-6.82)**	PBO/Sham	113/19,398	M
	Clonidine	**OR=3.96 (1.89-8.41)**	PBO/Sham	113/19,398	M
	Atomoxetine	**OR=3.63 (2.81-4.73)**	PBO/Sham	113/19,398	M
	Guanfacine	**OR=3.29 (2.27-4.82)**	PBO/Sham	113/19,398	M
Aggressive behavior	Amphetamines	**SMD=–1.15 (–1.38 to –0.93)**	PBO/Sham	3/84	L
	Methylphenidate	SMD=–0.26 (–1.10 to 0.68)	PBO/Sham	2/181	L
	Atomoxetine	RR=1.34 (0.91 to 1.97)	PBO/Sham	15/2,067	M
Cognition: executive memory	Methylphenidate	**SMD=–0.26 (–0.39 to –0.13)**	PBO/Sham	7/468	L
Cognition: non-executive memory	Methylphenidate	**SMD=–0.60 (–0.79 to –0.41)**	PBO/Sham	8/635	L
Cognition: reaction time	Methylphenidate	**SMD=–0.21 (–0.30 to –0.12)**	PBO/Sham	21/1,095	L
Cognition: response inhibition	Methylphenidate	**SMD=–0.41 (–0.55 to –0.27)**	PBO/Sham	16/846	L

Table 2 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with neurodevelopmental and disruptive behavior/dissocial/conduct disorders *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/patients	Q
Acceptability	Clonidine	**OR=0.40 (0.20-0.78)**	PBO/Sham	171/22,961	M
	Methylphenidate	**OR=0.59 (0.46-0.75)**	PBO/Sham	171/22,961	M
	Aripiprazole	OR=0.61 (0.02-25.34)	PBO/Sham	171/22,961	M
	Modafinil	OR=0.67 (0.37-1.24)	PBO/Sham	171/22,961	M
	Desipramine	OR=0.70 (0.17-2.89)	PBO/Sham	171/22,961	M
	Amphetamines	OR=0.78 (0.52-1.18)	PBO/Sham	171/22,961	M
	Guanfacine	OR=0.79 (0.54-1.14)	PBO/Sham	171/22,961	M
	Atomoxetine	OR=0.85 (0.68-1.07)	PBO/Sham	171/22,961	M
	Bupropion	OR=1.54 (0.39-6.76)	PBO/Sham	171/22,961	M
Tolerability	Methylphenidate	OR=1.31 (0.79-2.25)	PBO/Sham	60/12,188	M
	Modafinil	OR=1.34 (0.57-3.18)	PBO/Sham	60/12,188	M
	Amphetamines	OR=1.38 (0.64-3.00)	PBO/Sham	60/12,188	M
	Clonidine	OR=2.32 (0.63-8.94)	PBO/Sham	58/NR	H
	Bupropion	OR=3.60 (0.34-130)	PBO/Sham	60/12,188	M
	Atomoxetine	**OR=1.48 (1.01-2.18)**	PBO/Sham	60/12,188	M
	Guanfacine	**OR=3.39 (1.93-6.3)**	PBO/Sham	60/12,188	M
Discontinuation due to inefficacy	Amphetamine	**OR=0.11 (0.05-0.20)**	PBO/Sham	45/9,087	M
	Clonidine	**OR=0.29 (0.13-0.56)**	PBO/Sham	45/9,087	M
	Methylphenidate	**OR=0.31 (0.18-0.53)**	PBO/Sham	45/9,087	M
	Guanfacine	**OR=0.37 (0.26-0.54)**	PBO/Sham	45/9,087	M
	Atomoxetine	**OR=0.47 (0.33-0.67)**	PBO/Sham	45/9,087	M
	Bupropion	OR=1.97 (0.19-57.4)	PBO/Sham	45/9,087	M
Functioning	Atomoxetine	**SMD=–0.48 (–0.62 to –0.33)**	PBO/Sham	8/1,308	M
Functioning: academic	Amphetamines	**SMD=–0.56 (–0.73 to –0.39)**	PBO/Sham	8/826	M
Global illness improvement	Amphetamines	**OR=7.71 (5.52-10.77)**	PBO/Sham	40/NR	H
	Atomoxetine	**OR=2.28 (1.38-3.76)**	PBO/Sham	40/NR	H
	Guanfacine	**OR=3.63 (2.36-5.57)**	PBO/Sham	40/NR	H
	Methylphenidate	**OR=5.57 (3.99-7.79)**	PBO/Sham	40/NR	H
	Modafinil	**OR=3.22 (1.91-5.43)**	PBO/Sham	40/NR	H
	Clonidine	OR=2.78 (0.91-8.53)	PBO/Sham	40/NR	H
Global illness severity	Amphetamines	**SMD=–0.86 (–1.72 to –0.01)**	PBO/Sham	2/86	M
	Desipramine	**OR=26.41 (7.41-94.18)**	PBO/Sham	2/103	L
Quality of life	Methylphenidate	**SMD=–0.61 (–0.80 to –0.42)**	PBO/Sham	3/514	M
	Atomoxetine	**SMD=–0.39 (–0.50 to –0.28)**	PBO/Sham	16/2,361	M
Suicide attempt	Atomoxetine	RR=0.84 (0.03-20.00)	PBO/Sham	23/3,883	L
Suicidal ideation	Atomoxetine	RR=1.67 (0.83-3.36)	PBO/Sham	15/2,517	M
Pharmacological augmentation					
Efficacy	α2-agonists + stimulants	**SMD=–0.36 (–0.51 to –0.21)**	PBO/Sham	3/719	M
Acceptability	α2-agonists + stimulants	RR=0.74 (0.37-1.48)	PBO/Sham	3/726	L
Tolerability	α2-agonists + stimulants	RR=0.77 (0.05-12.50)	PBO/Sham	3/726	L
Discontinuation due to inefficacy	α2-agonists + stimulants	RR=0.49 (0.21-1.13)	PBO/Sham	3/726	M

Table 2 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with neurodevelopmental and disruptive behavior/dissocial/conduct disorders *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/patients	Q
Psychosocial interventions					
Efficacy (mixed-rated)	Social skills training	**SMD=–0.39 (–0.63 to –0.15)**	WL/NT	15/2,857	L
Efficacy (teacher-rated)	Social skills training	**SMD=–0.26 (–0.47 to –0.05)**	WL/NT	14/1,379	M
Efficacy (parent-rated)	Social skills training	**SMD=–0.54 (–0.81 to –0.26)**	WL/NT	11/1,206	L
Efficacy (clinician-rated)	Social skills training	SMD=–3.15 (–9.88 to 3.57)	WL/NT	2/107	L
Response	Behavioral therapy	**OR=2.97 (1.53-5.88)**	PBO/Sham	113/19,398	M
	Cognitive training	OR=0.70 (0.12-3.87)	PBO/Sham	113/19,398	M
Acceptability	Behavioral therapy	**OR=0.58 (0.33-0.99)**	PBO/Sham	171/22,961	M
	Cognitive training	OR=1.32 (0.71-2.52)	PBO/Sham	171/22,961	M
Functioning: academic	Social skills training	SMD=–0.15 (–0.31 to 0.01)	WL/NT	5/642	M
Global illness severity	Behavioral therapy	**OR=2.99 (1.21-7.31)**	PBO/Sham	113/19,398	M
	Cognitive training	OR=0.39 (0.01-5.80)	PBO/Sham	113/19,398	M
Functioning: social skills (mixed-rated)	Social skills training	**SMD=–0.29 (–0.47 to –0.11)**	WL/NT	19/2,649	L
Functioning: social skills (parent-rated)	Social skills training + parental involvement	**SMD=–0.43 (–0.70 to –0.15)**	WL/NT	4/337	L
	Social skills training	**SMD=–0.19 (–0.32 to –0.06)**	WL/NT	15/1,609	M
Functioning: social skills (teacher-rated)	Social skills training + parental involvement	SMD=–0.15 (–0.41 to 0.12)	WL/NT	4/632	M
	Social skills training	SMD=–0.11 (–0.22 to 0.00)	WL/NT	11/1,271	M
Functioning: emotional (mixed-rated)	Social skills training	SMD=0.20 (–0.01 to 0.41)	WL/NT	5/353	L
Functioning: emotional (parent-rated)	Social skills training	SMD=0.27 (–0.05 to 0.59)	WL/NT	3/173	L
Functioning: emotional (teacher-rated)	Social skills training	SMD=0.02 (–0.68 to 0.72)	WL/NT	2/129	L
Brain stimulation interventions					
Response	Neurofeedback	OR=1.96 (0.52-8.26)	PBO/Sham	113/19,398	M
Acceptability	Neurofeedback	OR=0.59 (0.31-1.14)	PBO/Sham	171/22,961	M
Combined interventions					
Response	Methylphenidate + parent training	**OR=55.63 (3.18-29.52x10^2)**	PBO/Sham	113/19,398	M
	Methylphenidate + clonidine	**OR=21.91 (5.52-105.40)**	PBO/Sham	113/19,398	M
	Atomoxetine + parent training	**OR=2.48 (0.51-11.79)**	PBO/Sham	113/19,398	M
Acceptability	Methylphenidate + clonidine	**OR=0.32 (0.13-0.77)**	PBO/Sham	171/22,961	M
ADHD and disorders of intellectual development					
Efficacy	Methylphenidate	**SMD=–0.88 (–1.14 to –0.61)**	PBO/Sham	8/424	L
Acceptability	Methylphenidate	OR=1.68 (0.68-4.14)	PBO/Sham	4/215	L
Tolerability	Methylphenidate	OR=4.82 (0.98-23.63)	PBO/Sham	4/215	L
Autism spectrum disorder					
Pharmacological interventions					
Efficacy: inappropriate speech (mixed-rated)	Aripiprazole	**SMD=–0.30 (–0.50 to –0.09)**	PBO/Sham	3/400	L
Efficacy: stereotypic (mixed-rated)	Aripiprazole	**SMD=–0.32 (–0.53 to–0.12)**	PBO/Sham	3/400	M
	Methylphenidate	SMD=–0.18 (–0.46 to 0.11)	PBO/Sham	5/127	M
	Atomoxetine	SMD=–0.16 (–0.50 to 0.18)	PBO/Sham	4/281	L

Table 2 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with neurodevelopmental and disruptive behavior/dissocial/conduct disorders *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/patients	Q
Efficacy: overall (teacher-rated)	Methylphenidate	SMD=–0.53 (–1.26 to 0.19)	PBO/Sham	2/37	L
Efficacy: social interaction (parent-rated)	Methylphenidate	SMD=–0.21 (–0.6 to 0.18)	PBO/Sham	2/90	L
Efficacy: social interaction (teacher-rated)	Methylphenidate	SMD=–0.51 (–1.07 to 0.05)	PBO/Sham	3/103	L
Efficacy: stereotypic (parent-rated)	Methylphenidate	SMD=–0.34 (–0.84 to 0.17)	PBO/Sham	3/NR	L
Efficacy: social withdrawal (mixed-rated)	Aripiprazole	SMD=–0.13 (–0.33 to 0.08)	PBO/Sham	3/400	M
Response	Risperidone	**OR=2.57 (1.35-4.86)**	PBO/Sham	3/241	L
	Aripiprazole	**RR=2.08 (1.24-3.46)**	PBO/Sham	3/400	L
Aggressive behavior	Risperidone	**SMD=–0.29 (–0.48 to –0.11)**	PBO/Sham	8/878	L
	Aripiprazole	**SMD=–0.24 (–0.40 to –0.08)**	PBO/Sham	8/878	L
	Valproate	SMD=–0.18 (–0.71 to 0.35)	PBO/Sham	2/57	M
	Lurasidone	SMD=–0.05 (–0.27 to 0.18)	PBO/Sham	8/878	L
Acceptability	Risperidone	**RR=0.52 (0.32-0.86)**	PBO/sham	6/379	M
	Antipsychotics	**RR=0.61 (0.48-0.78)**	PBO/Sham	15/1,124	M
	Aripiprazole	**RR=0.67 (0.49-0.90)**	PBO/Sham	5/526	M
	Haloperidol	RR=0.80 (0.24-2.62)	PBO/Sham	2/60	M
	Mood stabilizers	RR=1.27 (0.53-3.06)	PBO/Sham	5/125	M
Tolerability	Risperidone	RR=0.71 (0.17-2.92)	PBO/Sham	5/339	M
	Antipsychotics	RR=0.99 (0.55-1.79)	PBO/Sham	12/1,010	M
	Mood stabilizers	RR=1.13 (0.36-3.53)	PBO/Sham	4/112	M
	Aripiprazole	RR=1.24 (0.57-2.71)	PBO/Sham	4/493	M
Discontinuation due to inefficacy	Mood stabilizers	RR=2.11 (0.36-12.42)	PBO/Sham	3/60	M
Global illness severity	Aripiprazole	**SMD=–0.54 (–0.77 to –0.32)**	PBO/Sham	3/400	M
	Risperidone	**OR=10.5 (4.80-22.60)**	PBO/Sham	6/446	L
	Mood stabilizers	RR=1.55 (0.39-6.21)	PBO/Sham	3/77	L
Relapse	Risperidone	**RR=0.30 (0.13-0.68)**	PBO/Sham	2/56	M
Psychosocial interventions					
Efficacy: emotion recognition (mixed-rated)	Computer-assisted interaction	SMD=–0.53 (–1.12 to 0.05)	WL/NT	2/48	L
	Social skills training	SMD=–0.34 (–0.88 to 0.20)	WL/NT	2/54	L
Efficacy: social competence (mixed-rated)	Social skills training	**SMD=–0.47 (–0.78 to –0.16)**	WL/NT	4/178	L
Anxiety (subject-rated)	Cognitive behavioral therapy	SMD=–0.61 (–1.54 to 0.33)	WL/NT	5/181	L
Anxiety (parent-rated)	Cognitive behavioral therapy	**SMD=–1.12 (–1.91 to –0.34)**	WL/NT	7/244	L
Functioning: joint attention	Skills training-joint attention	**SMD=–0.66 (–0.93 to –0.40)**	WL/NT	9/417	L
Disruptive behavior/dissocial/conduct disorders (with or without ADHD)					
Pharmacological interventions					
Efficacy (clinician-rated)	Risperidone	**SMD=–0.48 (–0.71 to –0.24)**	PBO/Sham	4/293	L
Efficacy (parent-rated)	Risperidone	**SMD=–0.79 (–1.06 to –0.52)**	PBO/Sham	2/225	M
Efficacy (mixed-rated)	Risperidone	**SMD=–0.32 (–0.49 to –0.16)**	PBO/Sham	4/590	M
Response: aggressive behavior	Valproate	**OR=15.6 (1.91-128.1)**	PBO/Sham	2/47	L
	Lithium	**RR=4.56 (1.97-10.56)**	PBO/Sham	3/116	L
Aggressive behavior (clinician-rated)	Mixed (risperidone, quetiapine)	SMD=–0.24 (–0.76 to 0.29)	PBO/Sham	2/57	L
Aggressive behavior (parent-rated)	Risperidone	**SMD=–0.72 (–0.99 to –0.46)**	PBO/Sham	3/238	M

Table 2 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with neurodevelopmental and disruptive behavior/dissocial/conduct disorders *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/patients	Q
Aggressive behavior (mixed-rated)	Risperidone	**SMD=–0.60 (–0.89 to –0.31)**	PBO/Sham	2/188	L
	Mixed (risperidone, lithium, methylphenidate)	SMD=–1.93 (–3.88 to 0.02)	PBO/Sham	4/172	L
Acceptability	Mixed (risperidone, lithium, methylphenidate)	RR=0.97 (0.60-1.55)	PBO/Sham	8/631	L
Global illness severity	Risperidone	**SMD=–1.31 (–1.88 to –0.74)**	PBO/Sham	2/58	L
	Mixed (risperidone, quetiapine)	**SMD=–0.30 (–0.49 to –0.12)**	PBO/Sham	5/435	M
	Mixed (carbamazepine, lithium, amphetamines)	**RR=2.39 (1.10-5.21)**	PBO/Sham	4/136	L
Psychosocial interventions					
Efficacy (parent-rated)	Parental + child behavioral interventions	**SMD=–1.00 (–1.68 to –0.32)**	WL/NT	3/207	L
Intellectual disabilities and disruptive behavior/dissocial disorders (with or without ADHD)					
Aggressive behavior (clinician-rated)	Risperidone	**SMD=–1.09 (–1.39 to –0.79)**	PBO/Sham	4/257	L
	Aripiprazole	**SMD=–0.64 (–0.91 to –0.36)**	PBO/Sham	2/308	L
	Valproate	SMD=–0.06 (–0.75 to 0.63)	PBO/Sham	2/57	L
Aggressive behavior (mixed-rated)	Risperidone	**SMD=–0.70 (–1.01 to –0.39)**	PBO/Sham	3/266	L
Developmental coordination disorders					
Efficacy	Skills training	SMD=–0.27 (–0.85 to 0.31)	WL/NT	2/51	L
Tic disorder					
Efficacy: tics (clinician-rated)	Desipramine	SMD=–0.44 (–0.91 to 0.02)	PBO/Sham	2/75	L
	Methylphenidate	SMD=–0.28 (–0.58 to 0.03)	PBO/Sham	4/191	L
Tourette's disorder					
Efficacy (clinician-rated)	Antipsychotics (haloperidol, pimozide, risperidone, ziprasidone)	**SMD=–0.74 (–1.08 to –0.41)**	PBO/Sham	4/75	L
	Guanfacine	**SMD=–0.73 (–1.26 to –0.20)**	PBO/Sham	2/58	L
	Methylphenidate	SMD=–0.17 (–0.46 to 0.11)	PBO/Sham	4/161	L

RCTs – randomized controlled trials, SMD – standardized mean difference, OR – odds ratio, RR – risk ratio, PBO – placebo, WL – waiting list, NT – no treatment, NR – not reported, Q – quality (H – high, M – medium, L – low). Bold prints indicate significant values. SMDs<0 indicate that intervention is more effective than control. For discontinuation outcomes (acceptability, tolerability, inefficacy) and relapse, OR/RR<1 favors the intervention. For response and remission, OR/RR>1 favors the intervention.

Among psychosocial interventions, social skills training had a small to large effect size regarding the primary efficacy outcome and functioning, and CBT had a large effect concerning anxiety across different control groups (see Table 2). Parent-child interaction therapy and other mixed psychosocial interventions had a small to medium effect size for the primary efficacy outcome vs. TAU, as well as a small effect regarding cognition. Parent-child interaction therapy also improved aggression (medium effect size), irritability (medium effect size), and functioning (large effect size). Finally, behavioral therapy with an imitative component had a large effect size for the primary efficacy outcome against other active psychosocial interventions without the imitative component (see Tables 5, 6 and 7).

Depressive disorders

Results for depressive disorders are shown in Tables 3, 5, 6 and 7.

Fluoxetine was the only pharmacological intervention that

was superior to placebo on the primary efficacy outcome (medium effect size), as well as on response and remission (both small effect size). Nortriptyline worsened the primary efficacy outcome (large effect size), imipramine increased all-cause drop-out (small effect size), and imipramine, venlafaxine and duloxetine increased discontinuation due to intolerability (small to medium effect size). Venlafaxine increased suicidality (large effect size) (see Table 3).

Among psychosocial interventions, a large effect size on the primary efficacy outcome was apparent for interpersonal therapy, problem-solving therapy, family therapy, and CBT vs. waiting list/no treatment. However, these results were not confirmed vs. placebo or vs. TAU, except for interpersonal therapy, that remained superior when compared to placebo and TAU (medium effect size) (see Tables 3 and 5).

CBT was also superior to mixed interventions regarding the primary efficacy outcome (medium effect size), and to selective serotonin reuptake inhibitors (SSRIs) regarding suicidality (small effect size) (see Tables 3 and 6). Psychodynamically-oriented psychotherapy had a small effect size advantage regarding response, but no significant effect on the primary efficacy outcome vs. placebo (see Table 3).

As a combination treatment, CBT plus fluoxetine had a medium effect size advantage regarding the primary efficacy outcome vs. placebo (see Table 3), and CBT plus SSRI was superior concerning remission vs. CBT monotherapy, and functioning vs. antidepressant monotherapy (small effect size) (see Table 6).

Enuresis

Results for enuresis are shown in Tables 4 and 6.

Among pharmacological interventions, imipramine outperformed placebo regarding the primary efficacy outcome and response (small effect size), and amitriptyline was superior to placebo with respect to response (small effect size) (see Table 4).

Behavioral therapy with alarm outperformed waiting list on the primary efficacy outcome (small effect size) and response (large effect size), and maintained a small effect size regarding response vs. placebo (see Table 4).

No clear superior treatment emerged in monotherapy head-to-head comparisons. Combination of desmopressin plus behavioral therapy with alarm was superior to desmopressin alone regarding the primary efficacy outcome (medium effect size) and response (small effect size), while combination of oxybutynin plus imipramine was superior to either imipramine or oxybutynin monotherapy (small effect size) (see Table 6).

Obsessive-compulsive disorder

Results for obsessive-compulsive disorder are shown in Tables 4 and 5.

Fluoxetine was the pharmacological intervention with the broadest efficacy, including primary efficacy outcome, response, and global illness severity vs. placebo (small effect sizes). SSRIs as a class also improved response, remission and global illness severity, yet had a higher discontinuation rate due to intolerability than placebo (see Table 4).

Among monotherapy psychosocial interventions, CBT was superior to waiting list regarding the primary efficacy outcome (medium effect size), response (small effect size), remission (small effect size), quality of life (small effect size) and functioning (large effect size), and also to placebo concerning remission (small effect size) (see Table 4). Behavioral therapy with exposure and response prevention outperformed TAU for both response and acceptability (small effect size) (see Table 5).

As a combination treatment, CBT and sertraline outperformed placebo (medium effect size) (see Table 4). No significant differences emerged in head-to-head comparisons.

Anxiety disorders

Results for anxiety disorders are shown in Tables 4, 5 and 6.

SSRIs (fluoxetine, fluvoxamine, paroxetine) outperformed placebo regarding the primary efficacy outcome, and response (small to medium effect). Fluoxetine also outperformed placebo with respect to remission (small effect size) (see Table 4). Sertraline reduced suicidality compared with placebo, but paroxetine increased it.

CBT was superior to waiting list in different formats (i.e., individual, Internet, group) regarding the primary efficacy outcome (small to large effect size), depressive symptoms (small effect size), remission (small to large effect size) and quality of life (large effect size). CBT was also superior to placebo with respect to quality of life (large effect size) and to TAU regarding the primary efficacy outcome, remission and functioning (large effect size). Group CBT was superior to individual CBT in head-to-head comparisons (small effect size) (see Tables 4, 5 and 6).

No meta-analysis compared pharmacological vs. psychosocial interventions or combined treatment strategies.

Disruptive behavior/dissocial/conduct disorders

Results for disruptive behavior/dissocial/conduct disorders are shown in Tables 2 and 7.

Among pharmacological interventions, risperidone outperformed placebo across different raters regarding the primary efficacy outcome (medium effect size), aggressive behavior (medium effect size, also in people with intellectual disability), and global illness severity (medium effect size). Aggressive behavior was also improved by lithium and valproate (see Table 2).

Among psychosocial interventions, a combination of parental and child behavioral interventions had a large effect size vs. waiting list concerning the primary efficacy outcome, and a medium effect size vs. a mixed control group (see Tables 2 and 7).

Table 3 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with schizophrenia spectrum, depressive, and bipolar disorders

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/ patients	Q
Schizophrenia spectrum disorders					
Efficacy (clinician-rated)	Olanzapine	**SMD=–0.74 (–1.05 to –0.44)**	PBO/Sham	28/3,003	L
	Risperidone	**SMD=–0.62 (–0.89 to –0.34)**	PBO/Sham	28/3,003	L
	Lurasidone	**SMD=–0.48 (–0.71 to –0.25)**	PBO/Sham	28/3,003	M
	Aripiprazole	**SMD=–0.43 (–0.63 to –0.24)**	PBO/Sham	28/3,003	M
	Quetiapine	**SMD=–0.42 (–0.65 to –0.19)**	PBO/Sham	28/3,003	M
	Paliperidone	**SMD=–0.42 (–0.66 to –0.18)**	PBO/Sham	28/3,003	L
	Asenapine	**SMD=–0.38 (–0.66 to –0.11)**	PBO/Sham	28/3,003	M
	Ziprasidone	SMD=–0.14 (–0.40 to 0.11)	PBO/Sham	28/3,003	L
Response	Risperidone	**OR=3.46 (1.92-6.23)**	PBO/Sham	28/3,003	L
	Olanzapine	**OR=2.64 (1.07-4.18)**	PBO/Sham	28/3,003	L
	Lurasidone	**OR=2.56 (1.45-4.48)**	PBO/Sham	28/3,003	M
	Paliperidone	**OR=2.12 (1.07-4.18)**	PBO/Sham	28/3,003	L
	Quetiapine	**OR=1.86 (1.03-3.32)**	PBO/Sham	28/3,003	M
	Asenapine	OR=1.73 (0.96-3.10)	PBO/Sham	28/3,003	M
Global illness severity	Olanzapine	**SMD=–0.6 (–1.18 to –0.02)**	PBO/Sham	13/2,210	M
	Risperidone	**SMD=–0.50 (–0.73 to –0.27)**	PBO/Sham	12/2,158	L
	Paliperidone	**SMD=–0.44 (–0.67 to –0.22)**	PBO/Sham	12/2,158	L
	Lurasidone	**SMD=–0.41 (–0.77 to –0.05)**	PBO/Sham	13/2,210	M
	Quetiapine	**SMD=–0.41 (–0.77 to –0.05)**	PBO/Sham	13/2,210	M
	Ziprasidone	**SMD=–0.40 (–0.68 to –0.12)**	PBO/Sham	13/2,210	M
	Aripiprazole	**SMD=–0.35 (–0.59 to –0.11)**	PBO/Sham	13/2,210	M
	Asenapine	**SMD=–0.29 (–0.53 to –0.06)**	PBO/Sham	13/2,210	M
Acceptability	Paliperidone	**OR=0.26 (0.08-0.80)**	PBO/Sham	28/3,003	L
	Risperidone	**OR=0.31 (0.14-0.72)**	PBO/Sham	28/3,003	L
	Olanzapine	**OR=0.36 (0.15-0.85)**	PBO/Sham	28/3,003	L
	Lurasidone	OR=0.53 (0.18-1.55)	PBO/Sham	28/3,003	M
	Ziprasidone	OR=0.59 (0.22-1.58)	PBO/Sham	28/3,003	L
	Quetiapine	OR=0.63 (0.27-1.43)	PBO/Sham	28/3,003	M
	Asenapine	OR=0.91 (0.33-2.56)	PBO/Sham	28/3,003	M
	Aripiprazole	OR=1.48 (0.60-3.67)	PBO/Sham	28/3,003	M

Table 3 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with schizophrenia spectrum, depressive, and bipolar disorders *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/ patients	Q
Tolerability	Lurasidone	OR=0.45 (0.16-1.22)	PBO/Sham	13/2,210	M
	Ziprasidone	OR=0.99 (0.45-2.30)	PBO/Sham	13/2,210	M
	Risperidone	OR=2.38 (0.57-13.56)	PBO/Sham	13/2,210	M
	Aripiprazole	OR=2.54 (0.70-14.48)	PBO/Sham	13/2,210	M
	Asenapine	OR=2.67 (0.82-12.47)	PBO/Sham	13/2,210	M
	Quetiapine	OR=3.29 (0.92-16.75)	PBO/Sham	13/2,210	M
	Olanzapine	**OR=7.76 (1.23-87.44)**	PBO/Sham	13/2,210	M
	Paliperidone	**OR=23.12 (2.38-778.70)**	PBO/Sham	13/2,210	M
Discontinuation due to inefficacy	Paliperidone	**OR=0.10 (0.04-0.28)**	PBO/Sham	28/3,003	L
	Olanzapine	**OR=0.14 (0.06-0.31)**	PBO/Sham	28/3,003	L
	Risperidone	**OR=0.17 (0.07-0.42)**	PBO/Sham	28/3,003	L
	Ziprasidone	**OR=0.41 (0.20-0.84)**	PBO/Sham	28/3,003	L
	Lurasidone	OR=0.39 (0.09-1.77)	PBO/Sham	28/3,003	M
	Asenapine	OR=0.63 (0.23-1.73)	PBO/Sham	28/3,003	M
Depressive disorders					
Pharmacological interventions					
Efficacy (clinician-rated)	Fluoxetine	**SMD=–0.51 (–0.84 to –0.18)**	PBO/Sham	70/8,906	M
	Desipramine	SMD=–0.43 (–1.26 to 0.39)	PBO/Sham	70/8,906	M
	Duloxetine	SMD = –0.22 (–0.85 to 0.42)	PBO/Sham	70/8,906	M
	Venlafaxine	SMD = –0.25 (–0.87 to 0.36)	PBO/Sham	70/8,906	M
	Mirtazapine	SMD = –0.23 (–0.97 to 0.51)	PBO/Sham	70/8,906	M
	Citalopram	SMD=–0.18 (–0.89 to 0.55)	PBO/Sham	70/8,906	M
	Escitalopram	SMD=–0.17 (–0.88 to 0.54)	PBO/Sham	70/8,906	M
	Paroxetine	SMD=–0.16 (–0.67 to 0.35)	PBO/Sham	70/8,906	M
	Nefazodone	SMD=–0.14 (–0.85 to 0.57)	PBO/Sham	70/8,906	M
	Desvenlafaxine	SMD=–0.12 (–0.79 to 0.54)	PBO/Sham	70/8,906	M
	Sertraline	SMD=–0.11 (–0.71 to 0.49)	PBO/Sham	70/8,906	M
	Imipramine	SMD=–0.03 (–0.75 to 0.68)	PBO/Sham	70/8,906	M
	Vilazodone	SMD=–0.09 (–1.09 to 0.90)	PBO/Sham	70/8,906	M
	Amitriptyline	SMD=0.08 (–1.11 to 1.27)	PBO/Sham	70/8,906	M
	Nortriptyline	**SMD= 1.14 (0.46-1.81)**	PBO/Sham	70/8,906	M
Response	Nefazodone	**OR=2.1 (1.06-4.89)**	PBO/Sham	34/5,260	M
	Duloxetine	**OR=1.74 (1.12-2.84)**	PBO/Sham	34/5,260	M
	Fluoxetine	**OR=1.70 (1.25-2.39)**	PBO/Sham	34/5,260	M
	Desipramine	OR=1.59 (0.67-4.84)	PBO/Sham	34/5,260	M
	Escitalopram	OR=1.53 (0.96-2.58)	PBO/Sham	34/5,260	M
	Sertraline	OR=1.44 (0.79-2.97)	PBO/Sham	34/5,260	M
	Paroxetine	OR=1.3 (0.89-1.99)	PBO/Sham	34/5,260	M
	Venlafaxine	OR=1.16 (0.72-2.03)	PBO/Sham	34/5,260	M
	Citalopram	OR=1.02 (0.62-1.82)	PBO/Sham	34/5,260	M
	Imipramine	OR=0.83 (0.48-1.54)	PBO/Sham	34/5,260	M
	Nortriptyline	OR=0.57 (0.24-1.64)	PBO/Sham	34/5,260	M
	Amitriptyline	OR=0.22 (0.05-2.78)	PBO/Sham	34/5,260	M

Table 3 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with schizophrenia spectrum, depressive, and bipolar disorders *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/ patients	Q
Acceptability	Nefazodone	OR=0.49 (0.21-1.39)	PBO/Sham	66/9,075	M
	Vilazodone	OR=0.59 (0.27-1.54)	PBO/Sham	66/9,075	M
	Nortriptyline	OR=0.76 (0.28-3.41)	PBO/Sham	66/9,075	M
	Fluoxetine	OR=0.78 (0.56-1.15)	PBO/Sham	66/9,075	M
	Mirtazapine	OR=0.83 (0.40-2.08)	PBO/Sham	66/9,075	M
	Desvenlafaxine	OR=0.85 (0.47-1.74)	PBO/Sham	66/9,075	M
	Citalopram	OR=0.96 (0.52-1.97)	PBO/Sham	66/9,075	M
	Duloxetine	OR=1.04 (0.62-1.96)	PBO/Sham	66/9,075	M
	Venlafaxine	OR=1.12 (0.53-2.70)	PBO/Sham	66/9,075	M
	Amitriptyline	OR=1.16 (0.29-12.13)	PBO/Sham	66/9,075	M
	Paroxetine	OR=1.3 (0.81-2.27)	PBO/Sham	66/9,075	M
	Escitalopram	OR=1.4 (0.77-2.86)	PBO/Sham	66/9,075	M
	Sertraline	OR=162 (0.83-3.22)	PBO/Sham	66/9,075	M
	Desipramine	OR=2.21 (0.88-7.67)	PBO/Sham	66/9,075	M
	Imipramine	**OR=2.51 (1.26-6.25)**	PBO/Sham	66/9,075	M
Tolerability	Amitriptyline	OR=0.10 (0.02-32.16)	PBO/Sham	34/5,260	M
	Fluoxetine	OR=1.03 (0.5-2.7)	PBO/Sham	34/5,260	M
	Citalopram	OR=1.13 (0.45-3.66)	PBO/Sham	34/5,260	M
	Nefazodone	OR=1.29 (0.3-21.89)	PBO/Sham	34/5,260	M
	Mirtazapine	OR=1.36 (0.41-10.99)	PBO/Sham	34/5,260	M
	Paroxetine	OR=1.59 (0.77-3.95)	PBO/Sham	34/5,260	M
	Escitalopram	OR=1.64 (0.46-13.49)	PBO/Sham	34/5,260	M
	Desipramine	OR=2.85 (0.83-21.8)	PBO/Sham	34/5,260	M
	Sertraline	OR=2.94 (0.94-17.19)	PBO/Sham	34/5,260	M
	Duloxetine	**OR=2.80 (1.20-9.42)**	PBO/Sham	34/5,260	M
	Venlafaxine	**OR=3.19 (1.01-18.7)**	PBO/Sham	34/5,260	M
	Imipramine	**OR=5.49 (1.96-20.86)**	PBO/Sham	34/5,260	M
Quality of life	Mixed (fluoxetine, paroxetine, sertraline)	SMD=–0.11 (–0.26 to 0.03)	PBO/Sham	3/765	M
Relapse	SSRIs	**OR=0.34 (0.18-0.64)**	PBO/Sham	3/164	L
Remission	Fluoxetine	**RR=1.82 (1.25-2.63)**	PBO/Sham	2/315	M
	Sertraline	RR=1.09 (0.72-1.61)	PBO/Sham	2/376	M
Suicide attempt/ideation	Nefazodone	OR=0.29 (0.06-6.31)	PBO/Sham	34/NR	M
	Mirtazapine	OR=0.53 (0.10-40.83)	PBO/Sham	34/NR	M
	Imipramine	OR=0.59 (0.19-3.07)	PBO/Sham	34/NR	M
	Desvenlafaxine	OR=0.74 (0.41-1.49)	PBO/Sham	34/NR	M
	Escitalopram	OR=0.94 (0.44-2.55)	PBO/Sham	34/NR	M
	Duloxetine	OR=0.93 (0.55-1.71)	PBO/Sham	34/NR	M
	Fluoxetine	OR=1.11 (0.74-1.75)	PBO/Sham	34/NR	M
	Paroxetine	OR=1.71 (0.81-5.05)	PBO/Sham	34/NR	M
	Citalopram	OR=1.18 (0.46-4.43)	PBO/Sham	34/NR	M
	Vilazodone	OR=1.96 (0.45-100.00)	PBO/Sham	34/NR	M
	Sertraline	OR=2.22 (0.75-12.5)	PBO/Sham	34/NR	M
	Venlafaxine	**OR=8.33 (1.92-NC)**	PBO/Sham	34/NR	M

Table 3 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with schizophrenia spectrum, depressive, and bipolar disorders *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/ patients	Q
Psychosocial interventions					
Efficacy (clinician-rated)	IPT	**SMD=–1.37 (–2.04 to –0.7)**	WL/NT	70/8,906	L
	PSOLV	**SMD=–1.26 (–2.48 to –0.03)**	WL/NT	70/8,906	L
	FT	**SMD=–1.03 (–1.66 to –0.4)**	WL/NT	70/8,906	L
	CBT	**SMD=–0.94 (–1.40 to –0.48)**	WL/NT	70/8,906	L
	IPT	**SMD=–0.70 (–1.29 to –0.12)**	PBO/Sham	70/8,906	L
	FT	SMD=–0.36 (–0.95 to 0.24)	PBO/Sham	70/8,906	L
	CBT	SMD=–0.27 (–0.72 to 0.18)	PBO/Sham	70/8,906	L
	PSD-O	SMD=0.08 (–0.67 to 0.84)	PBO/Sham	70/8,906	L
Response	PSD-O	**RR=1.68 (1.08-2.63)**	WL/PBO/ Sham	2/83	L
Acceptability	IPT	OR=0.53 (0.20-1.15)	PBO/Sham	66/9,075	M
	IPT	OR=0.65 (0.19-1.62)	WL/NT	66/9,075	M
	CBT	OR=0.65 (0.32-1.16)	PBO/Sham	66/9,075	M
	PSOLV	OR=0.77 (0.01-4.40)	WL/NT	66/9,075	M
	CBT	OR=0.77 (0.34-1.48)	WL/NT	66/9,075	M
	FT	OR=0.84 (0.35-1.72)	PBO/Sham	66/9,075	M
	PSD-O	OR=0.96 (0.37-1.93)	PBO/Sham	66/9,075	M
	BT	OR=1.27 (0.19-4.32)	PBO/Sham	66/9,075	M
Suicide attempt/ideation	IPT	OR=0.64 (0.04-2.59)	PBO/Sham	34/NR	M
	CBT	OR=11.31 (0.01-46.11)	PBO/Sham	34/NR	M
	PSD-O	OR=8.64 (0.01-40.05)	PBO/Sham	34/NR	M
Combination interventions					
Efficacy (clinician-rated)	Fluoxetine+CBT	**SMD=–0.73 (–1.39 to –0.07)**	PBO/Sham	70/8,906	M
Acceptability	Fluoxetine+CBT	OR=0.75 (0.39-1.65)	PBO/Sham	66/9,075	M
Suicide attempt/ideation	Fluoxetine+CBT	OR=0.88 (0.41-2.35)	PBO/Sham	34/NR	M
Bipolar disorder, depressive episode					
Efficacy (clinician-rated)	Quetiapine	SMD=–0.10 (–0.32 to 0.13)	PBO/Sham	2/224	M
Response	Quetiapine	RR=1.1 (0.89-1.35)	PBO/Sham	3/250	L
Acceptability	Quetiapine	RR=0.73 (0.36-1.49)	PBO/Sham	2/225	L
Global illness severity	Quetiapine	**SMD=–0.20 (–0.46 to –0.06)**	PBO/Sham	2/224	M
Remission	Quetiapine	RR=1.23 (0.90-1.68)	PBO/Sham	3/250	L
Tolerability	Quetiapine	RR=0.31 (0.11-1.01)	PBO/Sham	2/225	L
Bipolar disorder, manic episode					
Efficacy (clinician-rated)	Aripiprazole	**SMD=–1.08 (–1.32 to –0.85)**	PBO/Sham	2/339	M

Table 3 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with schizophrenia spectrum, depressive, and bipolar disorders *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/ patients	Q
Response	Mixed (mood stabilizers and antipsychotics)	**OR=2.24 (z=8.12, p<0.001)**	PBO/Sham	9/1,362	M
	Aripiprazole	**RR=1.86 (1.43-2.43)**	PBO/Sham	2/332	M
	SGAs	**z=10.34, p<0.001**	PBO/Sham	6/1,190	H
	Mood stabilizers	**z=2.06, p=0.04**	PBO/Sham	2/172	M
Acceptability	Aripiprazole	RR=0.80 (0.51-1.27)	PBO/Sham	2/339	M
	Valproate	OR=1.77 (0.83-3.78)	PBO/Sham	2/179	M
Tolerability	Aripiprazole	RR=5.19 (0.92-29.25)	PBO/Sham	2/339	M
Discontinuation due to inefficacy	Aripiprazole	**RR=0.27 (0.09-0.82)**	PBO/Sham	2/339	M

RCTs – randomized controlled trials, SMD – standardized mean difference, OR – odds ratio, RR – risk ratio, PBO – placebo, WL – waiting list, NT – no treatment, NR – not reported, NC – not calculable, Q – quality (H – high, M – medium, L – low), BT – behavioral therapy, CBT – cognitive behavioral therapy, FT – family therapy, IPT – interpersonal therapy, PSD-O – psychodynamic-oriented, PSOLV – problem solving, SSRIs – selective serotonin reuptake inhibitors, SGAs – second-generation antipsychotics. Bold prints indicate significant values. SMDs<0 indicate that intervention is more effective than control. For discontinuation outcomes (acceptability, tolerability, inefficacy) and relapse, OR/RR<1 favors the intervention. For response and remission, OR/RR>1 favors the intervention.

Eating disorders

Results for eating disorders are shown in Table 6.

No meta-analysis on pharmacological intervention met the inclusion criteria of this umbrella review. Among psychosocial interventions, family therapy outperformed other interventions in anorexia nervosa regarding the primary efficacy outcome (body weight, small effect size).

Schizophrenia spectrum disorders

Results for schizophrenia spectrum disorders are shown in Tables 3 and 6.

For schizophrenia, only pharmacological interventions were covered. All investigated antipsychotics but ziprasidone outperformed placebo, with a small effect size, except for olanzapine and risperidone, which had a large effect size. Small effect sizes emerged regarding response (except for asenapine), and all antipsychotics improved global illness severity. Acceptability was superior vs. placebo for paliperidone, risperidone and olanzapine, without differences for the other antipsychotics. Paliperidone and olanzapine were associated with more discontinuation due to intolerability than placebo, while discontinuation due to inefficacy favored paliperidone, olanzapine, risperidone and ziprasidone (see Table 3).

In head-to-head comparisons, risperidone and second-generation antipsychotics outperformed first-generation antipsychotics (large effect size), and clozapine outperformed olanzapine on the primary efficacy outcome (large effect size) (see Table 6).

Bipolar disorder

Results for bipolar disorder are shown in Tables 3 and 6.

Regarding bipolar depression, quetiapine was not superior to placebo regarding the primary efficacy outcome, separating only on global illness severity (small effect size). Regarding mania, aripiprazole was more effective than placebo regarding the primary efficacy outcome (large effect size) and response (small effect size), without differences vs. placebo regarding acceptability, while being superior regarding less discontinuations for inefficacy (see Table 3).

Other disorders

Results for tic disorder are shown in Tables 2 and 6. Desipramine and methylphenidate were similar to placebo, but topiramate was superior to haloperidol regarding the primary outcome.

Results for Tourette's disorder are shown in Tables 2 and 7. Antipsychotics (including haloperidol, pimozide, risperidone and ziprasidone) and guanfacine were superior to placebo regarding the primary efficacy outcome (both moderate effect size). No significant difference vs. placebo emerged for methylphenidate (see Table 2). Among psychosocial interventions, behavioral therapy outperformed waiting list or low intensity psychosocial intervention (medium effect size) regarding the primary efficacy outcome (see Table 7).

Results for encopresis are shown in Table 5. No pharmacological intervention was eligible. Behavioral therapy outperformed TAU regarding the primary efficacy outcome and response (small effect size).

Results for developmental coordination disorders are shown in Table 2. In the single meta-analysis meeting inclusion criteria, skills training had no significant effect vs. waiting list on motor coordination.

Results for PTSD are shown in Table 4. No pharmacological intervention met inclusion criteria. CBT was superior regarding the primary efficacy outcome, response and depressive symptoms vs. waiting list (large effect sizes).

Table 4 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with anxiety, obsessive-compulsive, stress-related, and mixed disorders

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/patients	Q
Anxiety disorders					
Pharmacological interventions					
Efficacy (clinician-rated)	Paroxetine	**SMD=–0.43 (–0.75 to –0.10)**	PBO/Sham	14/2,502	M
	Fluvoxamine	**SMD=–0.36 (–0.61 to –0.10)**	PBO/Sham	14/2,502	M
	Imipramine	SMD=–0.27 (–0.92 to 0.39)	PBO/Sham	14/2,502	M
	Guanfacine	SMD=–0.13 (–0.39 to 0.12)	PBO/Sham	14/2,502	M
	Fluoxetine	SMD=–0.11 (–0.33 to 0.12)	PBO/Sham	14/2,502	M
	Atomoxetine	SMD=–0.11 (–0.38 to 0.16)	PBO/Sham	14/2,502	M
	Duloxetine	SMD=–0.09 (–0.27 to 0.09)	PBO/Sham	14/2,502	M
	Sertraline	SMD=–0.08 (–0.25 to 0.09)	PBO/Sham	14/2,502	M
	Venlafaxine	SMD=–0.06 (–0.22 to 0.04)	PBO/Sham	14/2,502	M
Efficacy (subject-rated)	Fluoxetine	**SMD=–0.51 (–0.85 to –0.18)**	PBO/Sham	2/154	M
	SNRIs	SMD=–2.14 (–9.75 to 5.48)	PBO/Sham	3/622	M
	Venlafaxine	SMD=–1.71 (–3.93 to 0.51)	PBO/Sham	2/443	M
	SSRIs	SMD=–0.42 (–0.96 to 0.12)	PBO/Sham	4/197	M
	Atomoxetine	SMD=–0.29 (–0.51 to 0.08)	PBO/Sham	2/331	M
	TCAs	SMD= 0.36 (–0.27 to 0.99)	PBO/Sham	2/41	M
Efficacy (parent-rated)	SSRIs	**SMD=–0.82 (–1.38 to –0.27)**	PBO/Sham	2/96	L
Response	Fluvoxamine	**OR=8.17 (1.35-49.40)**	PBO/Sham	19/2,656	M
	Sertraline	**OR=6.05 (2.23-49.40)**	PBO/Sham	19/2,656	M
	Fluoxetine	**OR=4.06 (1.49-18.17)**	PBO/Sham	19/2,656	M
	Guanfacine	OR=5.47 (0.74-49.40)	PBO/Sham	19/2,656	M
	Atomoxetine	OR=4.06 (0.67-24.53)	PBO/Sham	19/2,656	M
	Paroxetine	OR=3.67 (0.67-20.09)	PBO/Sham	19/2,656	M
	Imipramine	OR=3.00 (0.61-14.88)	PBO/Sham	19/2,656	M
	Venlafaxine	OR=2.46 (0.90-6.69)	PBO/Sham	19/2,656	M
	Duloxetine	OR=2.01 (0.37-11.02)	PBO/Sham	19/2,656	M
	Clomipramine	OR=1.22 (0.22-6.69)	PBO/Sham	19/2,656	M
Acceptability	Clomipramine	OR=0.55 (0.02-7.39)	PBO/Sham	20/2,679	M
	Paroxetine	OR=0.61 (0.12-3.32)	PBO/Sham	20/2,679	M
	Fluvoxamine	OR=0.67 (0.11-4.06)	PBO/Sham	20/2,679	M
	Sertraline	OR=0.67 (0.14-2.72)	PBO/Sham	20/2,679	M
	Guanfacine	OR=0.67 (0.10-4.95)	PBO/Sham	20/2,679	M
	Atomoxetine	OR=0.82 (0.15-4.95)	PBO/Sham	20/2,679	M
	Duloxetine	OR=1.00 (0.18-5.47)	PBO/Sham	20/2,679	M
	Venlafaxine	OR=1.11 (0.33-3.67)	PBO/Sham	20/2,679	M
	Fluoxetine	OR=1.65 (0.50-6.69)	PBO/Sham	20/2,679	M
	Imipramine	OR=2.01 (0.37-9.97)	PBO/Sham	20/2,679	M
Remission	Fluoxetine	**RR=2.52 (1.19-5.32)**	PBO/Sham	2/95	L
Suicide attempt/ ideation	Sertraline	**LogOR=–19.8 (–61.7 to 0.7)**	PBO/Sham	9/1,648	M
	Duloxetine	LogOR=0.2 (–2.5 to 2.8)	PBO/Sham	9/1,648	M
	Venlafaxine	LogOR=1.4 (–1.4 to 5.24)	PBO/Sham	9/1,648	M
	Atomoxetine	LogOR=6.6 (–31.6 to 22.7)	PBO/Sham	9/1,648	M
	Guanfacine	LogOR=16.1 (–1.0 to 58.3)	PBO/Sham	9/1,648	M
	Imipramine	LogOR=17.3 (–0.1 to 54.8)	PBO/Sham	9/1,648	M
	Paroxetine	**LogOR=20.0 (1.7 to 60.47)**	PBO/Sham	9/1,648	M

Table 4 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with anxiety, obsessive-compulsive, stress-related, and mixed disorders *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/patients	Q
Tolerability	Venlafaxine	LogOR=–0.8 (–3.8 to 2.1)	PBO/Sham	15/2,516	M
	Atomoxetine	LogOR=0.0 (–5.3 to 5.3)	PBO/Sham	15/2,516	M
	Duloxetine	LogOR=0.2 (–3.9 to 4.3)	PBO/Sham	15/2,516	M
	Sertraline	LogOR=1.7 (–2.8 to 6.6)	PBO/Sham	15/2,516	M
	Paroxetine	LogOR=1.7 (–2.5 to 6.0)	PBO/Sham	15/2,516	M
	Fluovoxamine	LogOR=2.1 (–2.4 to 7.0)	PBO/Sham	15/2,516	M
	Fluoxetine	LogOR=2.5 (–1.8 to 7.9)	PBO/Sham	15/2,516	M
	Imipramine	LogOR=16.6 (–37.5 to 83.7)	PBO/Sham	15/2,516	M
	Guanfacine	**LogOR=29.2 (2.2-94.3)**	PBO/Sham	15/2,516	M
Psychosocial interventions					
Efficacy (clinician-rated)	CBT/BT	**SMD=–0.85 (–1.12 to –0.57)**	WL/NT	7/358	L
Efficacy (subject-rated)	CBT-Child only	**SMD=–1.04 (–1.41 to –0.67)**	WL/NT	24/1,239	L
	CBT-Group	**SMD=–0.91 (–1.22 to –0.60)**	WL/NT	27/1,268	L
	CBT	**SMD=–0.67 (–0.88 to –0.47)**	WL/NT	45/2,831	L
	CBT-Child+P	**SMD=–0.45 (–0.67 to –0.23)**	WL/NT	20/1,285	L
	CBT-Individual	**SMD=–0.39 (–0.64 to –0.15)**	WL/NT	21/1,203	L
	CBT	**SMD=–0.31 (–0.51 to –0.11)**	PBO/Sham	15/978	L
	CBT-Parent only	SMD=0.04 (–0.38 to 0.46)	WL/NT	5/307	L
Efficacy (parent-rated)	CBT-Group	**SMD=–0.92 (–1.21 to –0.62)**	WL/NT	21/1,279	L
	CBT-Child only	**SMD=–0.87 (–1.21 to –0.53)**	WL/NT	13/734	L
	CBT	**SMD=–0.70 (–0.90 to –0.51)**	WL/NT	35/2137	L
	CBT-Child+P	**SMD=–0.69 (–0.98 to –0.39)**	WL/NT	17/1,031	L
	CBT-Individual	**SMD=–0.43 (–0.65 to –0.21)**	WL/NT	17/858	L
	CBT-Parent only	SMD=–0.37 (–0.77 to 0.04)	WL/NT	5/372	L
	CBT	SMD=–0.25 (–0.61 to 0.11)	PBO/Sham	8/638	L
Efficacy (mixed-rated)	BT-Group	**SMD=–1.43 (–2.36 to –0.51)**	WL/NT	101/6,625	L
	CBT-Group	**SMD=–1.43 (–1.76 to –1.09)**	WL/NT	101/6,625	L
	BT-Individual+P	**SMD=–1.09 (–1.93 to –0.25)**	WL/NT	101/6,625	L
	CBT-Group+P	**SMD=–0.99 (–1.31 to –0.68)**	WL/NT	101/6,625	L
	CBT-Individual	**SMD=–0.99 (–1.30 to –0.68)**	WL/NT	101/6,625	L
	CBT-Individual+P	**SMD=–0.84 (–1.16 to –0.53)**	WL/NT	101/6,625	L
	CBT-Group	**SMD=–0.76 (–1.16 to –0.36)**	PBO/Sham	101/6,625	L
	CBT-Parent only	**SMD=–0.70 (–1.22 to –0.19)**	WL/NT	101/6,625	L
	CBT-Internet	**SMD=–0.61 (–1.02 to –0.20)**	WL/NT	101/6,625	L
	BT-Individual+Group	SMD=–0.73 (–1.59 to 0.13)	WL/NT	101/6,625	L
	CBT-Individual+Group	SMD=–0.64 (–1.69 to 0.41)	WL/NT	101/6,625	L
	BT-Individual+P	SMD=–0.42 (–1.29 to 0.44)	PBO/Sham	101/6,625	L
	CBT-Group+P	SMD=–0.33 (–0.78 to 0.13)	PBO/Sham	101/6,625	L
	CBT-Individual	SMD=–0.32 (–0.72 to 0.07)	PBO/Sham	101/6,625	L
	CBT-Individual+P	SMD=–0.18 (–0.61 to 0.25)	PBO/Sham	101/6,625	L
	BT-Individual+Group	SMD=–0.06 (–0.94 to 0.82)	PBO/Sham	101/6,625	L
	CBT-Internet	SMD=0.06 (–0.48 to 0.60)	PBO/Sham	101/6,625	L

Table 4 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with anxiety, obsessive-compulsive, stress-related, and mixed disorders *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/patients	Q
Acceptability	CBT-Individual+Group	OR=0.26 (0.05-5.73)	WL/NT	101/6,625	L
	BT-Individual+P	OR=0.64 (0.22-2.72)	WL/NT	101/6,625	L
	BT-Individual+P	OR=0.81 (0.19-2.27)	PBO/Sham	101/6,625	L
	CBT-Group+P	OR=0.90 (0.46-1.60)	PBO/Sham	101/6,625	L
	CBT-Group	OR=0.85 (0.46-1.44)	PBO/Sham	101/6,625	L
	BT	OR=0.90 (0.32-3.95)	WL/NT	101/6,625	M
	CBT-Individual	OR=0.92 (0.52-1.52)	PBO/Sham	101/6,625	L
	CBT-Group	OR=0.93 (0.57-1.63)	WL/NT	101/6,625	L
	CBT	OR=1.09 (0.85-1.41)	WL/NT	45/3,158	L
	CBT-Group+P	OR=0.99 (0.67-1.55)	WL/NT	101/6,625	M
	CBT	OR=1.00 (0.68-1.49)	PBO/Sham	12/797	L
	CBT-Internet	OR=1.02 (0.42-2.08)	PBO/Sham	101/6,625	L
	CBT-Individual	OR=1.02 (0.67-1.67)	WL/NT	101/6,625	L
	CBT-Internet	OR=1.05 (0.59-2.05)	WL/NT	101/6,625	L
	CBT-Individual+P	OR=1.11 (0.60-1.90)	PBO/Sham	101/6,625	L
	BT-Individual+Group	OR=1.13 (0.28-3.19)	PBO/Sham	101/6,625	L
	BT-Group	OR=1.21 (0.27-22.51)	WL/NT	101/6,625	L
	CBT-Individual+P	OR=1.23 (0.80-2.02)	WL/NT	101/6,625	L
	CBT-Parent only	OR=1.43 (0.75-3.15)	WL/NT	101/6,625	L
Depressive symptoms	CBT	**SMD=–0.34 (–0.51 to –0.17)**	WL/NT	17/1,157	L
	CBT	SMD=–0.18 (–0.45 to 0.09)	PBO/Sham	10/613	L
Functioning	CBT	SMD=–1.03 (–1.38 to –0.68)	WL/NT	11/557	L
Quality of life	CBT-Parent only	**SMD=–1.87 (–3.04 to –0.71)**	WL/NT	101/6,625	L
	CBT-Individual	**SMD=–1.13 (–1.82 to –0.45)**	PBO/Sham	101/6,625	L
	CBT-Individual	**SMD=–1.01 (–1.55 to –0.48)**	WL/NT	101/6,625	L
	CBT-Internet	**SMD=–0.86 (–1.57 to –0.15)**	PBO/Sham	101/6,625	L
	CBT-Group	**SMD=–0.85 (–1.45 to –0.26)**	PBO/Sham	101/6,625	L
	CBT-Individual+P	**SMD=–0.80 (–1.33 to –0.27)**	WL/NT	101/6,625	L
	CBT-Group+P	**SMD=–0.75 (–1.34 to –0.17)**	WL/NT	101/6,625	L
	CBT-Group	**SMD=–0.73 (–1.34 to –0.11)**	WL/NT	101/6,625	L
	CBT-Internet	**SMD=–0.73 (–1.14 to –0.33)**	PBO/Sham	101/6,625	L
	BT-Individual+Group	SMD=–0.79 (–1.68 to 0.09)	WL/NT	101/6,625	L
	BT-Individual+Group	SMD=–0.67 (–1.56 to 0.21)	WL/NT	101/6,625	L
	CBT-Individual+Group	SMD=–0.55 (–1.78 to 0.69)	WL/NT	101/6,625	L
Remission	CBT-Child only	**OR=10.42 (5.84-7.60)**	WL/NT	19/1,184	M
	CBT-Group	**OR=6.25 (4.45-8.78)**	WL/NT	25/1,532	M
	CBT-Remote	**OR=6.14 (2.97-12.71)**	WL/NT	10/591	L
	CBT	**OR=5.45 (3.90-7.60)**	WL/NT	39/2,697	L
	CBT-Individual	**OR=4.53 (2.55-8.03)**	WL/NT	17/1,165	L
	CBT-Individual+P	**OR=4.08 (2.72-6.11)**	WL/NT	19/1,142	M
	CBT-Child only	**OR=3.58 (1.92-6.65)**	PBO/Sham	7/509	L
	CBT-Group	**OR=3.10 (1.14-8.45)**	PBO/Sham	5/353	L
	CBT-Parent only	**OR=2.83 (1.12-7.16)**	WL/NT	4/371	L
	CBT	**OR=2.28 (1.33-3.89)**	PBO/Sham	10/822	L
	CBT-Individual	**OR=2.04 (1.06-3.91)**	PBO/Sham	5/469	L
	CBT-Individual+P	OR=1.12 (0.65-1.92)	PBO/Sham	4/313	L

Table 4 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with anxiety, obsessive-compulsive, stress-related, and mixed disorders *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/patients	Q
Social anxiety disorder					
Efficacy (subject-rated)	CBT	**SMD=–1.59 (–2.33 to –0.86)**	WL/NT	11/603	L
	BT	**SMD=–1.22 (–2.06 to –0.38)**	WL/NT/PBO/Sham	4/169	L
	CBT	**SMD=–1.19 (–1.72 to –0.67)**	WL/NT/PBO/Sham	14/872	L
	CBT-Group	**SMD=–1.19 (–1.93 to –0.45)**	WL/NT/PBO/Sham	11/670	L
	CBT/BT	**SMD=–1.13 (–1.59 to –0.68)**	WL/NT/PBO/Sham	17/1,016	L
	CBT+P	**SMD=–1.13 (–1.59 to –0.67)**	WL/NT/PBO/Sham	17/983	L
	CBT-Individual	**SMD=–1.10 (–1.91 to –0.29)**	WL/NT/PBO/Sham	3/127	L
	CBT-Individual+Group	**SMD=–0.80 (–1.19 to –0.41)**	WL/NT/PBO/Sham	3/115	L
	CBT-Child only	**SMD=–0.75 (–1.24 to –0.26)**	WL/NT/PBO/Sham	2/70	L
	CBT-Internet	**SMD=–0.52 (–1.01 to –0.03)**	WL/NT/PBO/Sham	2/143	L
Acceptability	CBT	RR=1.00 (0.72-1.41)	WL/NT/PBO/Sham	16/1,052	M
Depressive symptoms	CBT/BT	**SMD=–0.39 (–0.63 to –0.16)**	WL/NT/PBO/Sham	8/299	L
Quality of life	CBT/BT	**SMD=–0.79 (–1.17 to –0.41)**	WL/NT/PBO/Sham	9/552	L
Remission	CBT/BT	**RR=8.99 (5.27-15.33)**	WL/NT/PBO/Sham	13/832	L
Obsessive-compulsive disorder					
Pharmacological interventions					
Efficacy (clinician-rated)	Sertraline	**SMD=–0.24 (–0.46 to –0.03)**	PBO/Sham	17/991	L
	Fluoxetine	**SMD=–0.24 (–0.47 to –0.01)**	PBO/Sham	17/991	L
	Clomipramine	SMD=–0.31 (–0.64 to 0.02)	PBO/Sham	17/991	L
	Fluvoxamine	SMD=–0.21 (–0.49 to 0.06)	PBO/Sham	17/991	L
Response	Fluoxetine	**RR=1.49 (1.15-1.96)**	PBO/Sham	2/146	L
	SSRI/TCAs	**RR=1.80 (1.43-2.26)**	PBO/Sham	7/692	L
Acceptability	Fluoxetine	MOR=0.74 (0.25-1.68)	PBO/Sham	18/1,143	L
	Fluvoxamine	MOR=0.79 (0.24-2.07)	PBO/Sham	18/1,143	L
	Sertraline	MOR=0.89 (0.32-2.07)	PBO/Sham	18/1,143	L
	Paroxetine	MOR=1.12 (0.37-3.42)	PBO/Sham	18/1,143	L
	Clomipramine	MOR=3.06 (0.54-21.69)	PBO/Sham	18/1,143	L
Tolerability	SSRIs	**RR=3.59 (1.89-6.84)**	PBO/Sham	7/807	L
Global illness severity	Fluoxetine	**SMD=–0.52 (–0.86 to –0.18)**	PBO/Sham	2/146	L
	SSRIs	**SMD=–0.42 (–0.61 to –0.23)**	PBO/Sham	5/556	M
Remission	SSRIs	**RR=2.06 (1.03-4.13)**	PBO/Sham	3/302	L
Pharmacological augmentation (in SSRI-refractory cases)					
Response	Risperidone	**OR=6.35 (1.48-27.3)**	PBO/Sham	3/72	M
	Quetiapine	OR=2.33 (0.88-6.20)	PBO/Sham	3/102	M
	Olanzapine	OR=2.74 (0.34-21.9)	PBO/Sham	2/70	L
Psychosocial interventions					
Efficacy (clinician-rated)	CBT	**SMD=–0.78 (–1.05 to –0.51)**	WL/NT	17/991	L
	BT	**SMD=–0.72 (–1.20 to –0.24)**	WL/NT	17/991	L
	CBT	SMD=–0.23 (–0.56 to 0.11)	PBO/Sham	17/991	L
Response	CBT/BT-ERP	**RR=3.93 (2.52-6.14)**	WL/NT/PBO/Sham	6/236	L

Table 4 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. inactive control in children/adolescents with anxiety, obsessive-compulsive, stress-related, and mixed disorders *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/patients	Q
Acceptability	CBT	MOR=0.49 (0.09-2.40)	PBO/Sham	18/1,143	L
	BT-ERP	RR=0.80 (0.35-1.84)	PBO/WL	6/301	L
	CBT	MOR=0.86 (0.23-3.24)	PBO/Sham	18/1,143	L
	CBT	MOR=0.94 (0.21-4.79)	WL/NT	18/1,143	L
	BT	MOR=14.28 (0.87-785.20)	WL/NT	18/1,143	L
Functioning (subject-rated)	CBT	**SMD=–1.15 (–2.11 to –0.19)**	WL/NT	3/194	L
Functioning (parent-rated)	CBT	**SMD=–0.95 (–1.61 to –0.28)**	WL/NT	3/194	L
	CBT	SMD=–0.31 (–0.63 to 0.01)	PBO/Sham	2/183	L
Remission	CBT	**RR=2.33 (1.33-4.00)**	WL/NT	4/271	L
	CBT	**RR=1.59 (1.28-1.96)**	PBO/Sham	3/153	L
Quality of life	CBT	**SMD=–0.39 (–0.77 to –0.02)**	WL/PBO/Sham	2/223	L
Combined interventions					
Efficacy	CBT+sertraline	**SMD=–0.58 (–0.91 to –0.25)**	PBO/Sham	17/991	L
Acceptability	CBT+sertraline	MOR=0.54 (0.08-3.15)	PBO/Sham	18/1,143	L
Post-traumatic stress disorder					
Efficacy	CBT	**SMD=–1.34 (–1.79 to –0.89)**	WL/NT	3/98	L
	EMDR	SMD=–0.61 (–1.96 to 0.74)	WL/NT	2/65	L
	NET	SMD=–0.57 (–1.23 to 0.09)	WL/NT	2/79	L
Response	CBT	**OR=8.64 (2.01-37.14)**	WL/NT	2/49	L
	NET	OR=3.82 (0.67-21.8)	WL/NT	2/78	L
Acceptability	NET	OR=5.13 (0.56-47.28)	WL/NT	2/83	L
Anxiety symptoms	NET	SMD=–0.66 (–1.33 to 0.01)	WL/NT	2/59	L
Depressive symptoms	CBT	**SMD=–0.8 (–1.47 to –0.131)**	WL/NT	3/98	L
Enuresis					
Pharmacological interventions					
Efficacy	Imipramine	**SMD=–0.46 (–0.67 to –0.24)**	PBO/Sham	4/347	M
Response	Amitriptyline	**RR=1.22 (1.02-1.45)**	PBO/Sham	2/98	L
	Imipramine	**RR=1.35 (1.11-1.64)**	PBO/Sham	12/831	L
Psychosocial interventions					
Efficacy	BT-Alarm	**SMD=–1.30 (–2.16 to –0.44)**	WL/NT	4/127	L
Response	BT-Alarm	**RR=7.23 (1.40-37.77)**	WL/NT	18/827	L
	BT-Alarm	**RR=1.59 (1.16-2.17)**	PBO/Sham	2/181	L
	BT-Reward	**RR=1.22 (1.03-1.45)**	WL/NT	2/325	L

RCTs – randomized controlled trials, SMD – standardized mean difference, OR – odds ratio, MOR – median odds ratio, RR – risk ratio, PBO – placebo, WL – waiting list, NT – no treatment, Q – quality (H – high, M – medium, L – low), BT – behavioral therapy, BT-ERP – behavioral therapy with exposure and response prevention, CBT – cognitive behavioral therapy, EMDR – eye movement desensitization and reprocessing, NET – narrative exposure therapy, P – parental involvement, SSRIs – selective serotonin reuptake inhibitors, SNRIs – serotonin-norepinephrine reuptake inhibitors, TCAs – tricyclic antidepressants. Bold prints indicate significant values. SMDs<0 indicate that intervention is more effective than control. For discontinuation outcomes (acceptability, tolerability, inefficacy) and relapse, OR/RR<1 favors the intervention. For response and remission, OR/RR>1 favors the intervention.

Table 5 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. treatment as usual (TAU) or low intensity psychosocial intervention (LIP) in children/adolescents (only significant differences are reported)

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/ patients	Q
Anxiety disorders					
Efficacy (mixed-rated)	CBT-Group	**SMD=–0.84 (–1.47 to –0.21)**	TAU	101/6,625	L
Functioning	CBT	**SMD=–1.06 (–1.57 to –0.55)**	TAU/LIP/PBO/Sham	5/467	L
Remission	CBT-Individual+P	**OR=8.56 (3.10-23.66)**	TAU	5/172	L
Autism spectrum disorder					
Efficacy: overall (mixed-rated)	PCIT	**SMD=–0.22 (–0.41 to –0.03)**	TAU/LIP	6/420	L
Efficacy: reciprocity (clinician-rated)	Mixed psychosocial interventions	**SMD=–0.53 (–0.78 to –0.29)**	TAU	8/380	L
Cognition: developmental quotient	Mixed psychosocial interventions	**SMD=–0.36 (–0.66 to –0.05)**	TAU	5/232	L
Cognition	PCIT	**SMD=–0.24 (–0.46 to –0.03)**	TAU/LIP	6/334	L
Anxiety disorder remission	CBT	**OR=11.25 (3.11-40.79)**	TAU	4/142	L
Depressive disorders					
Efficacy (clinician-rated)	IPT	**SMD=–0.66 (–1.22 to –0.09)**	TAU	70/8,906	L
Encopresis					
Efficacy: soiling	BT+TAU	**SMD=–0.35 (–0.63 to –0.07)**	TAU	4/209	L
Response	BT+TAU	**RR=1.78 (1.25-2.55)**	TAU	4/216	L
Obsessive-compulsive disorder					
Response	BT-ERP	**RR=1.71 (1.29-2.25)**	TAU/LIP	4/271	L
Acceptability	BT-ERP	**RR=0.60 (0.39-0.93)**	TAU/LIP	4/251	L

RCTs – randomized controlled trials, SMD – standardized mean difference, OR – odds ratio, RR – risk ratio, PBO – placebo, Q – quality (H – high, M – medium, L – low), BT – behavioral therapy, BT-ERP – behavioral therapy with exposure and response prevention, CBT – cognitive behavioral therapy, IPT – interpersonal therapy, PCIT – parent-child interaction therapy, P – parental involvement. SMDs<0 indicate that intervention is more effective than control. For discontinuation outcomes (acceptability, tolerability, inefficacy) and relapse, OR/RR<1 favors the intervention. For response and remission, OR/RR>1 favors the intervention.

DISCUSSION

Pooling top-tier evidence from 104 MAs/NMAs of RCTs reporting on the effects of pharmacological, psychosocial and brain stimulation interventions, targeting 20 different outcomes in 15 mental disorders or groups of mental disorders, this umbrella review provides a comprehensive meta-analytic view of the evidence base regarding the efficacy, acceptability and other relevant outcomes of psychiatric treatments in children and adolescents (see supplementary information for further details).

Considered together with a complementary umbrella review published in this journal[14], focusing on the detailed evaluation of tolerability and safety of pharmacological interventions, the current review can inform clinicians, youth and their families, as well as other stakeholders, in making evidence-based decisions regarding the choice and use of pharmacological, psychosocial and brain stimulation interventions in children/adolescents, in monotherapy and in combination. On the basis of these reviews, some evidence-based recommendation can be made.

For ADHD, amphetamines and methylphenidate are the most effective interventions on a broad set of outcomes. Whilst amphetamines outperform methylphenidate on the primary efficacy outcome, methylphenidate is the medication least different from placebo concerning safety[14]. Some evidence is available regarding behavioral therapy, covering a narrow set of efficacy outcomes, and with small effect sizes compared with those for medications. Importantly, whilst social skills training shows promising results against waiting list, no evidence is available comparing this intervention with placebo. Hence, amphetamines or methylphenidate can be considered the first-line treatment, augmented with alpha-2 agonists if needed, and ideally in combination with behavioral therapy as an optimal treatment regimen. Behavioral therapy could be considered if medications are contraindicated.

For autism, aripiprazole and risperidone are the pharmacological treatment options of choice. However, various psychosocial interventions have proven efficacy on a broad set of outcomes, ranging from anxiety (CBT), to irritability, aggressive behavior and functioning (parent-child interaction therapy), to

Table 6 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. active psychological intervention or drug condition in children/adolescents (only significant differences are reported)

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/patients	Q
Anorexia nervosa					
Efficacy: weight gain	FT	**SMD=–0.44 (–0.74 to –0.14)**	Other than FT	4/178	L
Anxiety disorders					
Efficacy (mixed-rated)	CBT-Group	**SMD=–0.44 (–0.82 to –0.06)**	CBT-Individual	101/6,625	L
Attention-deficit/hyperactivity disorder (ADHD)					
Efficacy (clinician-rated)	Amphetamines	**SMD=–0.24 (–0.44 to –0.05)**	Methylphenidate	46/NR	H
	Methylphenidate	**SMD=–0.22 (–0.39 to –0.05)**	Atomoxetine	46/NR	H
Efficacy (parent-rated)	Methylphenidate	**SMD=–1.07 (–1.74 to –0.40)**	Bupropion	23/NR	H
	Methylphenidate	**SMD=–0.23 (–0.37 to –0.10)**	Atomoxetine	23/NR	H
Response	Methylphenidate	**OR=1.44 (1.08-1.92)**	Atomoxetine	113/19,398	M
Aggressive behavior	Amphetamines	**SMD=–0.35 (–0.56 to –0.13)**	Methylphenidate	2/132	L
Acceptability	Methylphenidate	**OR=0.68 (0.52-0.91)**	Atomoxetine	171/22,961	M
Tolerability	Methylphenidate	**OR=0.39 (0.18-0.83)**	Guanfacine	60/12,188	M
Discontinuation due to inefficacy	Amphetamines	**OR=0.23 (0.10-0.44)**	Atomoxetine	45/9,087	M
Global illness severity	Amphetamines	**OR=3.39 (1.95-5.88)**	Atomoxetine	40/NR	H
Efficacy: inattention (mixed-rated)	Neurofeedback	**SMD=0.44 (0.02 to 0.86)**	Stimulants	4/161	L
Acceptability	Neurofeedback	**OR=0.45 (0.21-0.95)**	COG TR	171/22,961	M
Response	BT+stimulants	**OR=4.76 (2.50-9.09)**	BT	113/19,398	M
	BT+stimulants	**OR=4.58 (2.49-8.75)**	Stimulants	113/19,398	M
Autism spectrum disorder					
Efficacy: stereotypic (clinician-rated)	BT-IT	**SMD=–0.78 (–1.42 to –0.13)**	BT-CI	2/40	L
Efficacy: distal social behavior (clinician-rated)	BT-IT	**SMD=–0.98 (–1.64 to –0.32)**	BT-CI	2/40	L
Bipolar disorder, manic episode					
Efficacy (clinician-rated)	Risperidone	**SMD=–1.01 (–1.29 to –0.74)**	Valproate	2/228	M
Enuresis					
Acceptability	Desmopressin	**OR=0.45 (0.29-0.71)**	BT-Alarm	15/1,502	M
Efficacy	BT-Alarm	**SMD= –0.43 (–0.77 to –0.08)**	Desmopressin	4/285	L
Relapse	BT-Alarm	**OR=0.15 (0.03-0.53)**	Desmopressin	12/1,381	M
Efficacy	Desmopressin+ BT-Alarm	**SMD= –0.58 (–0.89 to –0.26)**	Desmopressin	2/156	L
Response	Desmopressin+anticholinergics	**OR=2.80 (1.50-5.40)**	Desmopressin	15/1,350	M
	Imipramine+oxybutynin	**RR=1.47 (1.09-2.00)**	Imipramine	2/101	L
	Imipramine+oxybutynin	**RR=1.46 (1.06-2.01)**	Oxybutynin	2/100	L
	Desmopressin+BT-Alarm	**RR=1.32 (1.08-1.62)**	Desmopressin	5/359	L
Relapse	Oxybutynin+ imipramine	**RR=0.50 (0.30-0.81)**	Oxybutynin	2/81	L
	Oxybutynin+ imipramine	**RR=0.48 (0.31-0.74)**	Imipramine	2/85	L
Depressive disorders					
Efficacy (clinician-rated)	Fluoxetine	**SMD=–1.65 (–2.34 to –0.95)**	Nortriptyline	70/8,906	M
Response	Fluoxetine	**OR=3.02 (1.04-7.22)**	Nortriptyline	34/5,260	M

Table 6 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. active psychological intervention or drug condition in children/adolescents (only significant differences are reported) *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/patients	Q
Tolerability	Paroxetine	**OR=0.22 (0.08-0.87)**	Imipramine	34/5,260	M
	Fluoxetine	**OR=0.31 (0.13-0.95)**	Duloxetine	34/5,260	M
Suicidal ideation	CBT	**SMD=–0.27 (–0.51 to –0.03)**	SSRIs	2/268	L
Remission	CBT+SSRI	**OR=2.15 (1.15-4.02)**	CBT+PBO	2/173	M
Functioning	CBT+SSRI	**SMD=–0.20 (–0.33 to –0.08)**	Standalone AD	4/850	L
Schizophrenia spectrum disorders					
Efficacy (clinician-rated)	Haloperidol	**SMD=–1.35 (–2.16 to –0.55)**	Fluphenazine	28/3,003	L
	Clozapine	**SMD=–0.86 (–1.54 to –0.17)**	Olanzapine	28/3,003	L
	SGAs	**SMD=–0.36 (–0.56 to –0.16)**	FGAs	4/243	L
Response	Risperidone	**OR=5.53 (2.01-15.18)**	Haloperidol	28/3,003	L
Tic disorder					
Response	Topiramate	**RR=1.10 (1.02-1.18)**	Haloperidol/ tiapride	14/1,017	M
	Topiramate	**RR=1.09 (1.01-1.19)**	Haloperidol	10/727	L

RCTs – randomized controlled trials, SMD – standardized mean difference, OR – odds ratio, RR – risk ratio, PBO – placebo, Q – quality (H – high, M – medium, L – low), BT – behavioral therapy, BT-IT– behavioral therapy imitative interaction, BT-CI – behavioral therapy contingency interaction, CBT – cognitive behavioral therapy, FT – family therapy, COG TR - cognitive training, AD – antidepressant, SSRI – selective serotonin reuptake inhibitor, SGAs – second-generation antipsychotics, FGAs – first-generation antipsychotics, NR – not reported. SMDs<0 indicate that intervention is more effective than control. For discontinuation outcomes (acceptability, tolerability, inefficacy) and relapse, OR/RR<1 favors the intervention. For response and remission, OR/RR>1 favors the intervention.

the primary efficacy outcome and functioning (social skills training, and behavioral therapy with imitative component). These benefits are not only observed vs. waiting list, but also against other active interventions. Given the different outcomes that these treatment modalities target, a variety of therapeutic tools can be considered, according to the patient's and family's resources, needs and choice, as well as the disease course and the presence of environmental stressors.

For depressive disorders in youth, fluoxetine is the only evidence-based pharmacological option. All other medications do not improve depression vs. placebo, but placebo effects are considerable. Imipramine, nortriptyline, and likely also venlafaxine should be avoided, given poor acceptability, tolerability and safety. As an alternative to medications, interpersonal therapy is the only psychosocial intervention outperforming placebo. The combination of CBT with fluoxetine also outperformed placebo on the primary efficacy outcome, and was superior to either monotherapy.

For enuresis, imipramine is the most effective pharmacological intervention. It can be combined with oxybutynin to maximize efficacy. However, due to the potential problems with tolerability of this medication in youth, psychosocial interventions should be tried first, including especially alarm behavioral therapy, that is supported by the largest body of evidence. No difference emerges among different types of alarms, and alarm maintains its efficacy after stopping the intervention[86].

For obsessive-compulsive disorder, fluoxetine and SSRIs as a class should be considered the first-line pharmacological treatment. Among psychosocial interventions, CBT and behavioral therapy with exposure and response prevention are effective options. If fluoxetine/SSRIs are ineffective, a switch to psychosocial interventions should be performed, and vice versa[71].

For anxiety disorders, fluoxetine and fluvoxamine are evidence-based pharmacological treatment strategies. Among psychosocial interventions, CBT – and in particular group CBT – should be offered as first-line treatment, likely before medications, given the large effect size and broad beneficial effect even vs. placebo in children and adolescents.

For disruptive behavior/dissocial/conduct disorders, risperidone emerges as the most effective pharmacological agent, but different types of behavioral treatment (including parent training) should be regarded as the first-line treatment options[118,119].

For anorexia nervosa in children and adolescents, family therapy is the intervention supported by the most significant evidence.

For schizophrenia spectrum disorders, antipsychotic treatment is the cornerstone of treatment. All tested antipsychotics, except for ziprasidone, have broadly similar superior efficacy vs. placebo, with olanzapine and risperidone being the most effective, and lurasidone/aripiprazole a more tolerable treatment option[102]. Ideally, starting with safer medications minimizing the risk of adverse events and maximizing adherence is a recommended strategy[14].

For bipolar disorder, little meta-analytic evidence is available overall. For mania, the only positive data are available for aripipra-

Table 7 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. mixed control conditions in children/adolescents (only significant differences are reported)

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/ patients	Q
Attention-deficit/hyperactivity disorder (ADHD)					
Efficacy (mixed-rated)	BI	**SMD=–0.55 (–0.77 to –0.32)**	WL/AC/LIP	6/333	L
Efficacy (probably blinded rater)	COG TR	**SMD=–0.20 (–0.40 to –0.01)**	Mixed	11/566	L
Efficacy (most proximal rater)	COG TR	**SMD=–0.37 (–0.66 to –0.09)**	Mixed	14/727	L
	BT	**SMD=–0.35 (–0.50 to –0.19)**	Mixed	19/1,430	L
Efficacy (teacher-rated)	ST	**SMD=–0.26 (–0.52 to –0.01)**	Mixed	6/615	L
Efficacy (parent-rated)	BT-Parental	**SMD=–0.65 (–1.05 to –0.25)**	TAU/WL/LIP	8/399	L
	ST	**SMD=–0.56 (–0.74 to –0.38)**	Mixed	10/934	L
Aggressive behavior	BI	**SMD=–0.40 (–0.71 to –0.10)**	Mixed	5/350	L
Functioning: academic	ST	**SMD=–0.33 (–0.51 to –0.14)**	Mixed	7/695	L
	BT	**SMD=–0.28 (–0.59 to –0.06)**	Mixed	9/817	L
Efficacy (most proximal rater)	Neurofeedback	**SMD=–0.35 (–0.59 to –0.11)**	Mixed	13/540	M
Efficacy (parent-rated)	Neurofeedback	**SMD=–0.32 (p=0.013)**	Mixed	16/706	L
Autism spectrum disorder					
Efficacy: socialization (mixed-rated)	PCIT	**SMD=–0.22 (–0.36 to –0.09)**	Mixed	13/846	L
Efficacy: language (mixed-rated)	PCIT	**SMD=–0.16 (–0.31 to –0.02)**	Mixed	13/785	L
Efficacy: language comprehension (parent-rated)	PCIT	**SMD=–0.29 (–0.56 to –0.01)**	Mixed	3/204	L
Anxiety (clinician-rated)	CBT	**SMD=–1.05 (–1.65 to –0.45)**	TAU/WL	6/208	L
Anxiety (parent-rated)	CBT	**SMD=–1.00 (–1.80 to –0.21)**	TAU/WL	7/283	L
Aggressive behavior	PCIT	**SMD = –0.67 (–0.85 to –0.49)**	Mixed	9/521	L
Functioning: shared/joint attention	ST-ToM	**SMD=–0.55 (–0.99 to –0.11)**	TAU/WL	2/88	L
	PCIT	**SMD=–0.41 (–0.68 to –0.14)**	Mixed	3/215	L
Functioning: social skills	SST-Computer	**SMD=–0.93 (–1.29 to –0.57)**	TAU/WL	5/138	L
	SST	**SMD=–0.83 (–1.07 to –0.60)**	TAU/WL	18/1,266	L
	SST-Face to face	**SMD=–0.81 (–1.08 to –0.53)**	TAU/WL	14/1,128	L
Functioning: parent synchrony	PCIT	**SMD=–0.90 (–1.23 to –0.56)**	Mixed	3/244	L
Global illness severity	PCIT	**SMD=–0.30 (–0.52 to –0.08)**	Mixed	6/316	L
Irritability	PCIT	**SMD=–0.59 (–0.88 to –0.30)**	Mixed	8/653	L
Depressive disorders					
Efficacy (mixed- rated)	CBT	**SMD=–0.53 (–0.82 to –0.24)**	Mixed	11/809	M
Oppositional defiant disorder (ODD)					
Efficacy (mixed-rated)	BI	**SMD=–0.79 (–0.93 to –0.64)**	WL/AC	17/NR	L
Tourette's disorder					
Efficacy (clinician-rated)	BT	**SMD=–0.64 (–0.99 to –0.29)**	WL/LIP	2/133	L
Disruptive behavior/dissocial/conduct disorders (with or without ADHD)					
Efficacy: ADHD symptoms (mixed- rated)	BI	**SMD=–0.34 (–0.64 to –0.05)**	WL/AC	11/518	L
Efficacy: ADHD symptoms (parent-rated)	BI	**SMD=–0.68 (–0.91 to –0.44)**	WL/AC	5/322	L
Efficacy: externalizing (mixed-rated)	BI	**SMD=–0.52 (–0.68 to –0.36)**	WL/AC	10/881	L

Table 7 Efficacy and effectiveness of pharmacological, psychosocial and brain stimulation interventions vs. mixed control conditions in children/adolescents (only significant differences are reported) *(continued)*

Outcome	Intervention	Effect size (95% CI)	Control	Number of RCTs/ patients	Q
Efficacy: ODD symptoms (mixed- rated)	BI	**SMD=–0.88 (–1.24 to –0.51)**	WL/AC	10/335	L
Efficacy: ODD symptoms (parent-rated)	BI	**SMD=–0.81 (–1.20 to –0.42)**	WL/AC	4/199	L
Aggressive behavior	BI	**SMD=–0.28 (–0.46 to –0.10)**	WL/AC	18/794	L
Cognition: attention	BI	**SMD=–0.38 (–0.52 to –0.23)**	WL/AC	15/588	L
Functioning	BI	**SMD=–0.39 (–0.52 to –0.26)**	WL/AC	22/1,027	L

RCTs – randomized controlled trials, SMD – standardized mean difference, WL – waiting list, AC – active control, TAU – treatment as usual, LIP – low intensity psychosocial intervention, Q – quality (H – high, M – medium, L – low), BT – behavioral therapy, CBT – cognitive behavioral therapy, COG TR – cognitive training, BI – combination of parental and child behavioral interventions, ST – skills training, PCIT – parent-child interaction therapy, SST – social skills training, ST-ToM – skills training: precursors of Theory of Mind, NR – not reported. SMDs<0 indicate that intervention is more effective than control.

zole, yet lithium is also an evidence-based treatment based on RCT evidence[120]. For bipolar depression, only quetiapine is superior to placebo, and only on a single outcome, namely global illness severity, but not on the primary symptom outcome. This finding is different from adults[121], and at least partially due to the larger placebo effects in youth. Our umbrella review did not include lurasidone and olanzapine/fluoxetine combination, as no meta-analysis has been conducted on them, but these are evidence-based options to treat bipolar depression in youth based on single RCTs[122,123], which led to their approval by the US Food and Drug Administration for bipolar depression in children and adolescents.

The available evidence presented in this umbrella review is not equally large across individual disorders, and also across monotherapies with pharmacological or psychosocial interventions. Even less meta-analytic data are available for head-to-head studies, within and across treatment modalities, and regarding combination treatments. Furthermore, little meta-analytic evidence exists on treatment-resistant youth with a given mental disorder. This is concerning, as early illness onset and disruption of healthy development may portend poorer response and outcomes, requiring information on non-responding conditions after first- and second-line treatments have been tried.

Among the 104 included meta-analyses, virtually none reported data on long-term treatment or relapse prevention. This is problematic, as most of these disorders are chronic and require long-term treatment.

This umbrella review clearly shows that large effect sizes emerge for psychosocial interventions when they are compared with waiting list or no treatment, where no placebo or expectation of study effect diminishes the treatment effect size. However, when those treatments are compared against psychological placebo or minimally active controls, significant effects either diminish in magnitude or disappear. This finding is relevant for indirect comparisons with pharmacological trials, in which the use of placebo makes the effect size appear smaller. The much greater difficulty of blinding treatment assignment in psychosocial trials is also to be taken into account. The risk of inflated effect sizes due to weak and methodologically flawed comparators (e.g., waiting list, no intervention) is that such interventions might be preferred to other superior treatments, delaying response and remission[121].

The results from this umbrella review should be considered within its limitations. First, we only considered evidence that was evaluated quantitatively via MAs/NMAs. This approach has excluded data from RCTs that have not (yet) been meta-analyzed. In particular, Internet-based psychosocial interventions, whose development has been recent and which may be particularly favored by youth[125,126], have not been sufficiently covered.

Second, we focused mainly on efficacy outcomes, while choices need to be made considering both efficacy and tolerability/safety. However, we included all-cause discontinuation as a global acceptability measure, as well as discontinuation due to intolerability as a core tolerability outcome, because these two events are typically measured and reported across both pharmacological and non-pharmacological treatment modalities. Detailed tolerability outcomes of pharmacological interventions in youth with mental disorders, that can be used to complement the present work on efficacy, have been recently published in this journal[14]. Such detailed data are not generally reported for psychosocial interventions, which is currently a major unmet need[127].

Third, as mentioned above, most meta-analytic evidence concerns the acute and short-term treatment effects, and much more data are required regarding the efficacy and safety of long-term and relapse prevention interventions for mental disorders in youth. Fourth, most evidence is available for monotherapy and vs. placebo/no treatment, although combination and augmentation treatments across and within pharmacological and psychosocial treatment modalities are commonly used in clinical practice, in youth as well as in adults[128]. Fifth, although 14 of the 104 included meta-analyses were NMAs that allow for direct and indirect head-to-head comparisons, most data were not derived from direct comparisons of active treatments, limiting the confidence with which comparative treatment choices can be made.

Sixth, since design, population and illness characteristics, as well as choice of control groups and blinding methods influence effect sizes, and these characteristics often differ substantially between pharmacological and non-pharmacological trials, indirect

comparisons of effect sizes across these treatment modalities need to be interpreted with caution. To overcome this limitation, more head-to-head comparisons and combination trials need to be conducted both within and across treatment modalities. Finally, we focused on those disorders that are most common and studied in youth, maximizing the chance of finding meta-analytic evidence, but other mental conditions could also be of interest.

Despite these limitations, inherent in the umbrella review methodology and available RCT data, this study provides the most comprehensive account of the available RCT evidence concerning pharmacological, psychosocial and brain stimulation interventions for the main psychiatric disorders in childhood and adolescents. The large body of literature reviewed here can inform future research aimed at addressing identified gaps, as well as current clinical care and guidelines regarding the choice of interventions for mental health conditions in youth, merging state-of-the-art efficacy and acceptability data with information on tolerability and safety.

ACKNOWLEDGEMENTS

E.G. Ostinelli is supported by the National Institute for Health Research (NIHR) Oxford Cognitive Health Clinical Research Facility and the NIHR Oxford Health Biomedical Research Centre (grant BRC-1215-20005). Supplementary information on this study is available at https://osf.io/2awu4/.

REFERENCES

1. Kessler RC, Berglund P, Demler O et al. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005;62:593-602.
2. Caspi A, Houts RM, Ambler A et al. Longitudinal assessment of mental health disorders and comorbidities across 4 decades among participants in the Dunedin birth cohort study. JAMA Netw Open 2020;3:e203221.
3. Wang PS, Berglund P, Olfson M et al. Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005;62:603-13.
4. GBD 2017 Child and Adolescent Health Collaborators. Diseases, injuries, and risk factors in child and adolescent health, 1990 to 2017: findings from the global burden of diseases, injuries, and risk factors 2017 Study. JAMA Pediatr 2019;173:e190337.
5. 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.
6. Pillay J, Boylan K, Carrey N et al. First- and second-generation antipsychotics in children and young adults: systematic review update. Rockville: US Agency for Healthcare Research and Quality, 2017.
7. 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 meta-analysis. Lancet 2016;388:881-90.
8. Pagsberg AK, Tarp S, Glintborg D et al. Acute antipsychotic treatment of children and adolescents with schizophrenia-spectrum disorders: a systematic review and network meta-analysis. J Am Acad Child Adolesc Psychiatry 2017;56:191-202.
9. 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.
10. Zhou X, Teng T, Zhang Y et al. Comparative efficacy and acceptability of antidepressants, psychotherapies, and their combination for acute treatment of children and adolescents with depressive disorder: a systematic review and network meta-analysis. Lancet Psychiatry 2020;7:581-601.
11. Dobson E, Bloch M, Strawn J. Efficacy and tolerability of pharmacotherapy for anxiety disorders. J Clin Psychiatry 2019;80:17r12064.
12. James AC, Reardon T, Soler A et al. Cognitive behavioural therapy for anxiety disorders in children and adolescents. Cochrane Database Syst Rev 2020;11:CD013162.
13. Huhn M, Tardy M, Spineli LM et al. Efficacy of pharmacotherapy and psychotherapy for adult psychiatric disorders: a systematic overview of meta-analyses. JAMA Psychiatry 2014;71:706-15.
14. Solmi M, Fornaro M, Ostinelli EG et al. Safety of 80 antidepressants, antipsychotics, anti-attention-deficit/hyperactivity medications and mood stabilizers in children and adolescents with psychiatric disorders: a large scale systematic meta-review of 78 adverse effects. World Psychiatry 2020;19:214-32.
15. Reed GM, First MB, Kogan CS et al. Innovations and changes in the ICD-11 classification of mental, behavioural and neurodevelopmental disorders. World Psychiatry 2019;18:3-19.
16. Correll CU, Rubio JM, Inczedy-Farkas G et al. Efficacy of 42 pharmacologic cotreatment strategies added to antipsychotic monotherapy in schizophrenia: systematic overview and quality appraisal of the meta-analytic evidence. JAMA Psychiatry 2017;74:675-84.
17. Shea BJ, Grimshaw JM, Wells GA et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.
18. Borenstein M, Hedges L, Higgins JPT et al. Comprehensive meta-analysis (Version 2.2.027). www.meta-analysis.com/.
19. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177-88.
20. Catalá-López F, Hutton B, Núñez-Beltrán A et al. The pharmacological and non-pharmacological treatment of attention deficit hyperactivity disorder in children and adolescents: a systematic review with network meta-analyses of randomised trials. PLoS One 2017;12:e0180355.
21. Luan R, Mu Z, Yue F et al. Efficacy and tolerability of different interventions in children and adolescents with attention deficit hyperactivity disorder. Front Psychiatry 2017;8:229.
22. Otasowie J, Castells X, Ehimare UP et al. Tricyclic antidepressants for attention deficit hyperactivity disorder (ADHD) in children and adolescents. Cochrane Database Syst Rev 2014;19:CD006997.
23. Punja S, Shamseer L, Hartling L et al. Amphetamines for attention deficit hyperactivity disorder (ADHD) in children and adolescents (Review). Cochrane Database Syst Rev 2016;2:CD009996.
24. Sun C-K, Tseng P-T, Wu C-K et al. Therapeutic effects of methylphenidate for attention-deficit/hyperactivity disorder in children with borderline intellectual functioning or intellectual disability: a systematic review and meta-analysis. Sci Rep 2019;9:15908.
25. Battagliese G, Caccetta M, Ines O et al. Behaviour research and therapy cognitive-behavioral therapy for externalizing disorders: a meta-analysis of treatment effectiveness. Behav Res Ther 2015;75:60-71.
26. Faraone SV, Biederman J, Roe C. Comparative efficacy of Adderall and methylphenidate in attention-deficit/hyperactivity disorder: a meta-analysis. J Clin Psychopharmacol 2002;22:468-73.
27. Van Doren J, Arns M, Heinrich H et al. Sustained effects of neurofeedback in ADHD: a systematic review and meta-analysis. Eur Child Adolesc Psychiatry 2019;28:293-305.
28. Cortese S, Ferrin M, Brandeis D et al. Cognitive training for attention-deficit/hyperactivity disorder: meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. J Am Acad Child Adolesc Psychiatry 2015;54:164-74.
29. Daley D, Van Der Oord S, Ferrin M et al. Behavioral interventions in attention-deficit/hyperactivity disorder: a meta-analysis of randomized controlled trials across multiple outcome domains. J Am Acad Child Adolesc Psychiatry 2014;53:835-47.
30. Bikic A, Reichow B, McCauley SA et al. Meta-analysis of organizational skills interventions for children and adolescents with attention-deficit/hyperactivity disorder. Clin Psychol Rev 2017;52:108-23.
31. Mulqueen JM, Bartley CA, Bloch MH. Meta-analysis: parental interventions for preschool ADHD. J Atten Disord 2015;19:118-24.
32. Cortese S, Ferrin M, Brandeis D et al. Neurofeedback for attention-deficit/hyperactivity disorder: meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. J Am Acad Child Adolesc Psychiatry 2016;55:444-55.
33. Bussalb A, Congedo M, Barthélemy Q et al. Clinical and experimental factors influencing the efficacy of neurofeedback in ADHD: a meta-analysis. Front Psychiatry 2019;10:35.
34. Stuhec M, Munda B, Svab V et al. Comparative efficacy and acceptability of atomoxetine, lisdexamfetamine, bupropion and methylphenidate in treatment of attention deficit hyperactivity disorder in children and adolescents: a meta-analysis with focus on bupropion. J Affect Disord 2015;178:149-59.

35. Faraone SV, Biederman J. Efficacy of Adderall® for attention-deficit/hyperactivity disorder: a meta-analysis. J Atten Disord 2002;6:69-75.
36. Schachter HM, Pham B, King J et al. How efficacious and safe is short-acting methylphenidate for the treatment of attention-deficit disorder in children and adolescents? A meta-analysis. CMAJ 2001;165:1475-88.
37. Schwartz S, Correll CU. Efficacy and safety of atomoxetine in children and adolescents with attention-deficit/hyperactivity disorder: results from a comprehensive meta-analysis and metaregression. J Am Acad Child Adolesc Psychiatry 2014;53:174-87.
38. Coghill DR, Seth S, Pedroso S et al. Effects of methylphenidate on cognitive functions in children and adolescents with attention-deficit/hyperactivity disorder: evidence from a systematic review and a meta-analysis. Biol Psychiatry 2014;76:603-15.
39. Storebø OJ, Ramstad E, Krogh HB et al. Methylphenidate for children and adolescents with attention deficit hyperactivity disorder (ADHD). Cochrane Database Syst Rev 2015;11:CD009885.
40. Bangs ME, Wietecha LA, Wang S et al. Meta-analysis of suicide-related behavior or ideation in child, adolescent, and adult patients treated with atomoxetine. J Child Adolesc Psychopharmacol 2014;24:426-34
41. Hirota T, Schwartz S, Correll C. Alpha-2 agonists for attention-deficit/hyperactivity disorder in youth: a systematic review and meta-analysis of monotherapy and add-on trials to stimulant therapy. J Am Acad Child Adolesc Psychiatry 2014;53:153-73.
42. Storebø OJ, Andersen ME, Skoog M et al. Social skills training for attention deficit hyperactivity disorder (ADHD) in children aged 5 to 18 years. Cochrane Database Syst Rev 2019;6:CD008223.
43. Fallah MS, Shaikh MR, Neupane B et al. Atypical antipsychotics for irritability in pediatric autism: a systematic review and network meta-analysis. J Child Adolesc Psychopharmacol 2019;29:168-80.
44. Maneeton N, Maneeton B, Putthisri S et al. Aripiprazole in acute treatment of children and adolescents with autism spectrum disorder: a systematic review and meta-analysis. Neuropsychiatr Dis Treat 2018;14:3063-72.
45. Yu Q, Li E, Li L et al. Efficacy of interventions based on applied behavior analysis for autism spectrum disorder: a meta-analysis. Psychiatry Investig 2020;17:432-43.
46. Oono IP, Honey EJ, McConachie H. Parent-mediated early intervention for young children with autism spectrum disorders (ASD). BJPsych Adv 2016;22:146.
47. Parsons L, Cordier R, Munro N et al. A systematic review of pragmatic language interventions for children with autism spectrum disorder. PLoS One 2017;12:e0172242.
48. Kreslins A, Robertson AE, Melville C. The effectiveness of psychosocial interventions for anxiety in children and adolescents with autism spectrum disorder: a systematic review and meta-analysis. Child Adolesc Psychiatry Ment Health 2015;9:22.
49. Tarver J, Palmer M, Webb S et al. Child and parent outcomes following parent interventions for child emotional and behavioral problems in autism spectrum disorders: a systematic review and meta-analysis. Autism 2019;23:1630-44.
50. Soares EE, Bausback K, Beard CL et al. Social skills training for autism spectrum disorder: a meta-analysis of in-person and technological interventions. J Technol Behav Sci (in press).
51. Postorino V, Sharp WG, McCracken CE et al. A systematic review and meta-analysis of parent training for disruptive behavior in children with autism spectrum disorder. Clin Child Fam Psychol Rev 2017;20:391-402.
52. Maneeton N, Maneeton B, Putthisri S et al. Risperidone for children and adolescents with autism spectrum disorder: a systematic review. Neuropsychiatr Dis Treat 2018;14:1811-20.
53. Zhou MS, Nasir M, Farhat LC et al. Meta-analysis: pharmacologic treatment of restricted and repetitive behaviors in autism spectrum disorders. J Am Acad Child Adolesc Psychiatry 2020;60:35-45.
54. Murza KA, Schwartz JB, Hahs-Vaughn DL et al. Joint attention interventions for children with autism spectrum disorder: a systematic review and meta-analysis. Int J Lang Commun Disord 2016;51:236-51.
55. Sturman N, Deckx L, van Driel ML. Methylphenidate for children and adolescents with autism spectrum disorder. Cochrane Database Syst Rev 2017; 11:CD011144.
56. Fletcher-Watson S, Mcconnell F, Manola E et al. Interventions based on the Theory of Mind cognitive model for autism spectrum disorder (ASD) (Review). Cochrane Database Syst Rev Interv 2014;3:CD008785
57. Cohen D, Raffin M, Canitano R et al. Risperidone or aripiprazole in children and adolescents with autism and/or intellectual disability: a Bayesian meta-analysis of efficacy and secondary effects. Res Autism Spectr Disord 2013;7:167-75.
58. Hirota T, Veenstra-Vanderweele J, Hollander E et al. Antiepileptic medications in autism spectrum disorder: a systematic review and meta-analysis. J Autism Dev Disord 2014;44:948-57.
59. D'Alò GL, De Crescenzo F, Amato L et al. Acceptability, equity, and feasibility of using antipsychotics in children and adolescents with autism spectrum disorder: a systematic review. BMC Psychiatry 2020;20:561.
60. Ospina MB, Seida JK, Clark B et al. Behavioural and developmental interventions for autism spectrum disorder: a clinical systematic review. PLoS One 2008;3:e3755.
61. Reichow B, Steiner AM, Volkmar F et al. Social skills groups for people aged 6 to 21 with autism spectrum disorders (ASD). Cochrane Database Syst Rev 2012;7:CD008511.
62. Tachibana Y, Miyazaki C, Ota E et al. A systematic review and meta-analysis of comprehensive interventions for pre-school children with autism spectrum disorder (ASD). PLoS One 2017;12:e0186502.
63. Nevill RE, Lecavalier L, Stratis EA. Meta-analysis of parent-mediated interventions for young children with autism spectrum disorder. Autism 2018; 22:84-98.
64. Spielmans GI, Gerwig K. The efficacy of antidepressants on overall well-being and self-reported depression symptom severity in youth: a meta-analysis. Psychother Psychosom 2014;83:158-64.
65. 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 2020;26:118-33.
66. Whittington CJ, Kendall T, Fonagy P et al. Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. Lancet 2004;363:1341-5.
67. Watanabe N, Hunot V, Omori IM et al. Psychotherapy for depression among children and adolescents: a systematic review. Acta Psychiatr Scand 2007;116:84-95.
68. Cox GR, Callahan P, Churchill R et al. Psychological therapies versus antidepressant medication, alone and in combination for depression in children and adolescents. Cochrane Database Syst Rev 2014;11:CD008324.
69. Dubicka B, Elvins R, Roberts C et al. Combined treatment with cognitive-behavioural therapy in adolescent depression: meta-analysis. Br J Psychiatry 2010;197:433-40.
70. Klein JB. Cognitive-behavioral therapy for adolescent depression: a meta-analytic investigation of changes in effect-size estimates. J Am Acad Child Adolesc Psychiatry 2007;46:1403-13.
71. Skapinakis P, Caldwell D, Hollingworth W et al. A systematic review of the clinical effectiveness and cost-effectiveness of pharmacological and psychological interventions for the management of obsessive-compulsive disorder in children/adolescents and adults. Health Technol Assess 2016;20:1-392.
72. Maneeton N, Maneeton B, Karawekpanyawong N et al. Fluoxetine in acute treatment of children and adolescents with obsessive-compulsive disorder: a systematic review and meta-analysis. Nord J Psychiatry 2020;74:461-9.
73. McGuire JF, Piacentini J, Lewin AB et al. A meta-analysis of cognitive behavior therapy and medication for child obsessive-compulsive disorder: moderators of treatment efficacy, response, and remission. Depress Anxiety 2015;32:580-93.
74. Locher C, Koechlin H, Zion SR et al. Efficacy and safety of selective serotonin reuptake inhibitors, serotonin-norepinephrine reuptake inhibitors, and placebo for common psychiatric disorders among children and adolescents: a systematic review and meta-analysis. JAMA Psychiatry 2017;74:1011-20.
75. Geller D. Which SSRI? A meta-analysis of pharmacotherapy trials in pediatric obsessive-compulsive disorder. Am J Psychiatry 2003;160:1919-28.
76. Uhre CF, Uhre VF, Lønfeldt NN et al. Systematic review and meta-analysis: cognitive-behavioral therapy for obsessive-compulsive disorder in children and adolescents. J Am Acad Child Adolesc Psychiatry 2020;59:64-77.
77. Johnco C, McGuire JF, Roper T et al. A meta-analysis of dropout rates from exposure with response prevention and pharmacological treatment for youth with obsessive compulsive disorder. Depress Anxiety 2020;37:407-17.
78. Zhou X, Zhang Y, Furukawa TA et al. Different types and acceptability of psychotherapies for acute anxiety disorders in children and adolescents: a network meta-analysis. JAMA Psychiatry 2019;76:41-50.
79. Wang Z, Whiteside SPH, Sim L et al. Comparative effectiveness and safety of cognitive behavioral therapy and pharmacotherapy for childhood anxiety disorders: a systematic review and meta-analysis. JAMA Pediatr 2017;171:1049-56.
80. Zhang H, Zhang Y, Yang L et al. Efficacy and acceptability of psychotherapy for anxious young children a meta-analysis of randomized controlled trials. J Nerv Ment Dis 2017;205:931-41.
81. Sigurvinsdóttir AL, Jensínudóttir KB, Baldvinsdóttir KD et al. Effectiveness of cognitive behavioral therapy (CBT) for child and adolescent anxiety disor-

ders across different CBT modalities and comparisons: a systematic review and meta-analysis. Nord J Psychiatry 2020;74:168-80.
82. James A, James G, Cowdrey FA et al. Cognitive behavioural therapy for anxiety disorders in children and adolescents. Cochrane Database Syst Rev 2015;11:CD013162.
83. Yang L, Zhou X, Pu J et al. Efficacy and acceptability of psychological interventions for social anxiety disorder in children and adolescents: a meta-analysis of randomized controlled trials. Eur Child Adolesc Psychiatry 2019;28:79-89.
84. Kreuze LJ, Pijnenborg GHM, de Jonge YB et al. Cognitive-behavior therapy for children and adolescents with anxiety disorders: a meta-analysis of secondary outcomes. J Anxiety Disord 2018;60:43-57.
85. Song P, Huang C, Wang Y et al. Comparison of desmopressin, alarm, desmopressin plus alarm, and desmopressin plus anticholinergic agents in the management of paediatric monosymptomatic nocturnal enuresis: a network meta-analysis. BJU Int 2019;123:388-400.
86. Caldwell PHY, Codarini M, Stewart F et al. Alarm interventions for nocturnal enuresis in children. Cochrane Database Syst Rev 2020;5:CD002911.
87. Caldwell PHY, Sureshkumar P, Wong WCF. Tricyclic and related drugs for nocturnal enuresis in children. Cochrane Database Syst Rev 2016;1:CD002117.
88. Caldwell PHY, Nankivell G, Sureshkumar P. Simple behavioural interventions for nocturnal enuresis in children. Cochrane Database Syst Rev 2013;7:CD003637.
89. Buckley BS, Sanders CD, Spineli L et al. Conservative interventions for treating functional daytime urinary incontinence in children. Cochrane Database Syst Rev 2019;9:CD012367.
90. Deshpande AV, Caldwell PH, Sureshkumar P. Drugs for nocturnal enuresis in children (other than desmopressin and tricyclics). Cochrane Database Syst Rev 2012;12:CD002238.
91. Peng CCH, Yang SSD, Austin PF et al. Systematic review and meta-analysis of alarm versus desmopressin therapy for pediatric monosymptomatic enuresis. Sci Rep 2018;8:16755.
92. Seida JC, Schouten JR, Mousavi SS et al. First- and second- generation antipsychotics for children and young adults: comparative effectiveness. Rockville: US Agency for Healthcare Research and Quality, 2012.
93. Loy J, Merry S, Hetrick S et al. Atypical antipsychotic drugs for disruptive behaviour disorders in children and youths. Cochrane Database Syst Rev 2017;8:CD008559.
94. Pringsheim T, Hirsch L, Gardner D et al. The pharmacological management of oppositional behaviour, conduct problems, and aggression in children and adolescents with attention-deficit hyperactivity disorder, oppositional defiant disorder, and conduct disorder: a systematic review and meta-analysis. Can J Psychiatry 2015;60:52-61.
95. Ipser J, Stein DJ. Systematic review of pharmacotherapy of disruptive behavior disorders in children and adolescents. Psychopharmacology 2007;191:127-40.
96. McQuire C, Hassiotis A, Harrison B et al. Pharmacological interventions for challenging behaviour in children with intellectual disabilities: a systematic review and meta-analysis. BMC Psychiatry 2015;15:303.
97. Zeeck A, Herpertz-Dahlmann B, Friederich HC et al. Psychotherapeutic treatment for anorexia nervosa: a systematic review and network meta-analysis. Front Psychiatry 2018;9:158.
98. Couturier J, Kimber M, Szatmari P. Efficacy of family-based treatment for adolescents with eating disorders: a systematic review and meta-analysis. Int J Eat Disord 2013;46:3-11.
99. Fisher CA, Skocic S, Rutherford KA et al. Family therapy approaches for anorexia nervosa. Cochrane Database Syst Rev 2019;5:CD004780.
100. van den Berg E, Houtzager L, de Vos J et al. Meta-analysis on the efficacy of psychological treatments for anorexia nervosa. Eur Eat Disord Rev 2019;27:331-51.
101. Linardon J, Wade TD, de la Piedad Garcia X et al. The efficacy of cognitive-behavioral therapy for eating disorders: a systematic review and meta-analysis. J Consult Clin Psychol 2017;85:1080-94.
102. Krause M, Zhu Y, Huhn M et al. Efficacy, acceptability, and tolerability of antipsychotics in children and adolescents with schizophrenia: a network meta-analysis. Eur Neuropsychopharmacol 2018;28:659-74.
103. Arango C, Ng-Mak D, Finn E et al. Lurasidone compared to other atypical antipsychotic monotherapies for adolescent schizophrenia: a systematic literature review and network meta-analysis. Eur Child Adolesc Psychiatry 2020;29:1195-205.
104. Sarkar S, Grover S. Antipsychotics in children and adolescents with schizophrenia: a systematic review and meta-analysis. Indian J Pharmacol 2013;45:439-46.
105. Kumar A, Datta SS, Wright SD et al. Atypical antipsychotics for psychosis in adolescents. Cochrane Database Syst Rev 2013;10:CD009582.
106. Maneeton B, Putthisri S, Maneeton N et al. Quetiapine monotherapy versus placebo in the treatment of children and adolescents with bipolar depression: a systematic review and meta-analysis. Neuropsychiatr Dis Treat 2017;13:1023.
107. Meduri M, Gregoraci G, Baglivo V et al. A meta-analysis of efficacy and safety of aripiprazole in adult and pediatric bipolar disorder in randomized controlled trials and observational studies. J Affect Disord 2016;191:187-208.
108. Liu HY, Potter MP, Woodworth KY et al. Pharmacologic treatments for pediatric bipolar disorder: a review and meta-analysis. J Am Acad Child Adolesc Psychiatry 2011;50:749-62.
109. Jochim J, Rifkin-Zybutz R, Geddes J et al. Valproate for acute mania. Cochrane Database Syst Rev 2019;10:CD004052.
110. Bloch MH, Panza KE, Landeros-Weisenberger A et al. Meta-analysis: treatment of attention-deficit/hyperactivity disorder in children with comorbid tic disorders. J Am Acad Child Adolesc Psychiatry 2009;48:884-93.
111. Yu L, Yan J, Wen F et al. Revisiting the efficacy and tolerability of topiramate for tic disorders: a meta-analysis. J Child Adolesc Psychopharmacol 2020;30:316-25.
112. Hollis C, Pennant M, Cuenca J et al. Clinical effectiveness and patient perspectives of different treatment strategies for tics in children and adolescents with Tourette syndrome: a systematic review and qualitative analysis. Health Technol Assess 2016;20:1-450
113. Zheng W, Li X, XIang Y et al. Aripiprazole for Tourette's syndrome: a systematic review and metaanalysis. Hum Psychopharmacol Clin Exp 2016;31:11-8.
114. Freeman K, Riley A, Duke D et al. Systematic review (and meta-analysis) of behavioral interventions for fecal incontinence with constipation. J Pediatr Psychol 2014;39:887-902.
115. Brazzelli M, Griffiths P, Cody JT et al. Behavioural and cognitive interventions with or without other treatments for the management of faecal incontinence in children. Cochrane Database Syst Rev 2011;12:CD0022400.
116. Miyahara M, Sl H, Pridham L et al. Task-oriented interventions for children with developmental co-ordination disorder. Cochrane Database Syst Rev 2017;7:CD010914.
117. Gillies D, Taylor F, Gray C et al. Psychological therapies for the treatment of post-traumatic stress disorder in children and adolescents (Review). Evid Based Child Health 2013;8:1004-116.
118. Scotto Rosato N, Correll CU, Pappadopulos E et al. Treatment of maladaptive aggression in youth: CERT guidelines II. Treatments and ongoing management. Pediatrics 2012;129:e1577-86.
119. Knapp P, Chait A, Pappadopulos E et al. Treatment of maladaptive aggression in youth: CERT guidelines I. Engagement, assessment, and management. Pediatrics 2012;129:e1562-76.
120. Findling RL, Robb A, McNamara NK et al. Lithium in the acute treatment of bipolar i disorder: a double-blind, placebo-controlled study. Pediatrics 2015;136:885-94.
121. Bahji A, Ermacora D, Stephenson C et al. Comparative efficacy and tolerability of pharmacological treatments for the treatment of acute bipolar depression: a systematic review and network meta-analysis. J Affect Disord 2020;269:154-84.
122. DelBello MP, Goldman R, Phillips D et al. Efficacy and safety of lurasidone in children and adolescents with bipolar I depression: a double-blind, placebo-controlled study. J Am Acad Child Adolesc Psychiatry 2017;56:1015-25.
123. Detke HC, DelBello MP, Landry J et al. Olanzapine/fluoxetine combination in children and adolescents with bipolar I depression: a randomized, double-blind, placebo-controlled trial. J Am Acad Child Adolesc Psychiatry 2015;54:217-24.
124. Skarphedinsson G, Hanssen-Bauer K, Kornør H et al. Standard individual cognitive behaviour therapy for paediatric obsessive-compulsive disorder: a systematic review of effect estimates across comparisons. Nord J Psychiatry 2015;69:81-92.
125. Andersson G, Titov N, Dear BF et al. Internet-delivered psychological treatments: from innovation to implementation. World Psychiatry 2019;18:20-8.
126. Linardon J, Cuijpers P, Carlbring P et al. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019;18:325-36.
127. Cuijpers P. Targets and outcomes of psychotherapies for mental disorders: an overview. World Psychiatry 2019;18:276-85.
128. Cuijpers P, Noma H, Karyotaki E et al. A network meta-analysis of the effects of psychotherapies, pharmacotherapies and their combination in the treatment of adult depression. World Psychiatry 2020;19:92-107.

DOI:10.1002/wps.20881

Internalizing psychopathology and all-cause mortality: a comparison of transdiagnostic vs. diagnosis-based risk prediction

Hyunsik Kim[1], Nicholas A. Turiano[2], Miriam K. Forbes[3], Roman Kotov[4], Robert F. Krueger[5], Nicholas R. Eaton[6]; HiTOP Utility Workgroup*

[1]Department of Psychology, Sogang University, Seoul, South Korea; [2]Department of Psychology, West Virginia University, Morgantown, WV, USA; [3]Department of Psychology, Macquarie University, Sydney, NSW, Australia; [4]Department of Psychiatry, Stony Brook University, Stony Brook, New York, NY, USA; [5]Department of Psychology, University of Minnesota, Minneapolis, MI, USA; [6]Department of Psychology, Stony Brook University, Stony Brook, NY, USA

*Members of the HiTOP Utility Workgroup are listed in the Appendix

Previous studies have documented the utility of a transdiagnostic internalizing factor in predicting important future outcomes (e.g., subsequent mental disorder diagnoses). To date, however, no study has investigated whether an internalizing factor predicts mortality risk. Also, while previous studies of mortality risk have emphasized its associations with particular internalizing disorders, no study has assessed how the transdiagnostic internalizing factor vs. disorder-specific variance differently predict that risk. The primary aims of this study were to explore: a) whether the internalizing factor predicts mortality risk, b) whether particular internalizing psychopathologies uniquely predict mortality risk over and beyond the transdiagnostic internalizing factor, and c) whether there is a significant interaction of internalizing with self-reported health in the prediction of mortality risk. We utilized a large national sample of American adults from the Midlife in the United States (MIDUS), a longitudinal study that examined midlife development of individuals across multiple waves between 1995 and 2015. Data were analyzed for the 6,329 participants who completed the phone interview and self-administered questionnaire in MIDUS 1 (1995-1996) and were then followed up until October 31, 2015 or until death. To investigate the association between internalizing and mortality risk, we used the semi-parametric proportional hazards Cox model, where survival time was regressed on a latent internalizing factor. Overall findings indicate that a transdiagnostic internalizing factor significantly predicts mortality risk over a 20-year period (hazard ratio, HR=1.12, 95% CI: 1.05-1.16, p<0.01) and that internalizing outperforms disorder-specific variance (e.g., depression-specific variance) in the prediction of that risk. Further, there was a significant interaction between transdiagnostic internalizing and self-reported health, whereby internalizing psychopathology had a specific association with early death for individuals with excellent self-reported health condition (HR=1.50, 95% CI: 1.17-1.84, p<0.05). This highlights the clinical utility of using the transdiagnostic internalizing factor for prediction of an important future outcome, and supports the argument that internalizing psychopathology can be a meaningful liability to explore in public health practice.

Key words: Internalizing factor, mortality, transdiagnostic prediction, diagnosis-based prediction, major depressive disorder, generalized anxiety disorder, panic disorder, neuroticism

(World Psychiatry 2021;20:276–282)

Numerous studies have reported that individuals with depressive or anxiety symptoms are at higher risk of experiencing various negative physical health conditions subsequently, compared with individuals without those symptoms. For example, depressive symptoms are associated with greater decline in physical performance in the later stages of life[1], increased risk of developing various forms of cardiovascular disease[2,3], and excessive risk of developing some forms of cancer[4]. Furthermore, childhood separation anxiety symptoms predict poor physical health in later stages of development[5]; generalized anxiety disorder (GAD) symptoms are associated with risk for coronary heart disease[6]; and GAD and post-traumatic stress disorder symptoms longitudinally predict shorter leukocyte telomere length (a biomarker for age-related diseases)[7].

Not surprisingly, a number of studies have reported the association of depression and/or anxiety with a higher risk of mortality[8-10]. For example, using survival analysis, a study investigated mortality rate in a large Danish population-based cohort (N=5,103,699), reporting that individuals with unipolar depression had a higher risk of early death[11]. Several studies also found that individuals with anxiety symptoms were exposed to a higher risk of premature death[12,13]. Additionally, some studies have indicated that individuals with a higher level of neuroticism, a personality trait with a close relation to mood and anxiety disorders[14], also have a higher mortality risk[15,16].

Although informative, a major limitation of prior research is that it mainly focused on how particular categorical diagnostic constructs were associated with mortality risk, while there has been growing evidence supporting the value of a dimensional conceptualization of psychopathology[17-21]. According to this latter approach, each mental disorder can be conceptualized as a manifestation of relatively few underlying transdiagnostic dimensions, which account for the co-occurrence among various disorders (i.e., comorbidity). For example, major depressive disorder (MDD) and GAD tend to co-occur more frequently than it is expected by chance[22]. This may indicate that they are highly correlated through the transdiagnostic internalizing factor. Indeed, numerous studies have reported that the internalizing factor accounts for the commonalities among various mood and anxiety disorders[22-25].

This framework provides the opportunity to investigate how the transdiagnostic internalizing factor, compared to particular forms of internalizing pathology (e.g., diagnostic categories), is associated with mortality risk[26]. A few prior studies have suggested a possible association of the common variance among various internalizing disorders with that risk[15,27,28]. For example, Mirza et al[27] reported that the relationship between anxiety symptoms and mortality risk was no longer significant after adjusting for comorbid depressive symptoms. This finding seems to suggest that it is the common variance that anxiety shares with depression which leads to higher mortality, and that the anxiety

disorder-specific variance may not predict mortality risk meaningfully once comorbid depression is controlled for.

There are several advantages of using the transdiagnostic internalizing factor as a predictor. Previous research has shown notable structural invariance of internalizing across different samples[29], high long-term stability of internalizing over time[30], and notable predictive validity for important future outcomes (e.g., subsequent mental disorder diagnoses)[30-32]. Given these findings, transdiagnostic internalizing could be a reliable and strong predictor of mortality risk.

It is also probable that the anticipated relationship between transdiagnostic internalizing and mortality risk is moderated by other factors. A possible moderator is one's self-reported health, given some prior studies suggesting that the association of depression and neuroticism with mortality risk varied depending on one's self-reported health condition[11,15].

Taking all the research discussed above into consideration, major limitations of the prior literature are that: a) no study has investigated whether or not a transdiagnostic internalizing dimension meaningfully predicts mortality risk, and b) previous studies have focused on the associations of individual diagnostic constructs with that risk, leaving it unclear whether these constructs have a general or a specific and unique association with early mortality[8,9,11,28,33]. This underscores the necessity to compare the prediction of mortality risk from various internalizing disorders' shared variance (i.e., transdiagnostic internalizing) versus the specific (unique) variance of each disorder, to ascertain which is a more robust predictor.

The primary aims of the current study were: a) to investigate whether the transdiagnostic internalizing factor predicts mortality risk in a longitudinal probability sample of American adults, b) to compare the utility of the transdiagnostic internalizing factor versus disorder-specific variance in the prediction of that risk, and c) to examine whether self-rated physical health moderates the association between internalizing and early mortality.

METHODS

Participants

This study utilized a large national sample of American adults from the Midlife in the United States (MIDUS)[34], which is a longitudinal study examining midlife development of individuals across multiple waves. Our study initially utilized information on the 7,108 participants who were recruited in the initial survey at MIDUS 1 (1995-1996). In order to be included in the final sample, participants needed to complete the MIDUS 1 phone interview and self-administered questionnaire, which yielded the final analytic sample of 6,329 individuals (mean age: 46.77±12.92 years; 52.64% females; 88.04% White, 4.90% African American).

These individuals were followed up until October 31, 2015 or until death. A total of 1,234 people were deceased during the study period (i.e., from 1995 to 2015). The mean survival time for all participants was 19.23±4.16 years. The mean survival time for decedents was 11.50±5.28 years.

Measures

To model a transdiagnostic internalizing factor, we included continuous symptom scores for MDD, GAD, panic disorder and neuroticism. Past 12-month MDD, GAD and panic disorder symptoms were measured using the Composite International Diagnostic Interview - Short Form (CIDI-SF) version 10, whose good diagnostic reliability and validity have been reported by numerous studies[35-37]. Neuroticism was assessed using the relevant subscale of the Midlife Development Inventory Personality Scales, whose internal consistency has been found to be good (Cronbach's alpha = .74)[38].

We chose six covariates based on the following criteria: a) whether a given covariate had been previously identified to likely influence mortality risk, and b) whether there was a large enough response rate for a given covariate (more than 6,000 responses). Based on these criteria, the six covariates chosen were: age (a standardized variable), age squared, sex (a binary variable), education level (ranged from 1 to 12, with larger numbers indicating higher educational levels), experienced severe health condition (a continuous variable ranged from 0 to 3, where higher scores indicate more severe physical health condition), and heart disease family risk (a binary variable).

Analyses

To investigate the association between the transdiagnostic internalizing factor and mortality risk, we used the semi-parametric proportional hazards Cox model. This model makes fewer assumptions about the distribution of survival time than do parametric models (e.g., Weibull, exponential models), enabling one to estimate regression coefficients and hazard ratios (HRs) even though the baseline hazard is not specified. This advantage makes it a practical and reasonable choice.

To model the transdiagnostic internalizing factor, we used confirmatory factor analysis (CFA), including four indicators (MDD, GAD, panic disorder, and neuroticism) assessed at MIDUS 1. This internalizing model (estimated from these same data and indicator variables) was previously identified as invariant across the different age cohorts and stable over time[39]. After modeling internalizing, we saved the factor scores to include them in the main Cox regression model.

The factor score approach may raise an issue of factor indeterminacy. In order to mitigate this concern, we further checked the factor determinacy index, which was calculated by the correlation between the estimated and true factor scores (ranging from 0 to 1; the higher the better representation of the true factor scores).

We performed survival analyses using maximum likelihood estimation with robust standard errors (MLR), with the latent internalizing variable standardized to have a variance of 1 and a mean of 0. All analyses were performed in Mplus version 8.0.

In order to compare the predictive validity of internalizing versus disorder-unique variance, we parameterized an explicit residual variance factor for each of the three internalizing disorders and neuroticism (i.e., the unique variance remaining in

each indicator after the common variance is accounted for by the latent internalizing variable). We then saved the factor scores from transdiagnostic internalizing and the four construct residual factors, and regressed survival time on both internalizing and the residual factor scores simultaneously.

RESULTS

The key assumption that Cox regression poses is that each predictor's multiplicative effect on the hazards function remains constant over time (proportional hazards assumption)[40,41]. We tested this assumption by assessing time-by-covariates interaction, which has been proven to be powerful for detecting non-proportionality[42]. This method involved creating the interaction term of internalizing x survival function time, including it in a Cox model with internalizing, and testing the significance of the interaction term. The result showed that the interaction term was not significant (HR=1.01, 95% CI: 0.99-1.01), indicating that the proportional hazards assumption was met.

We first examined how each of the four indicators assessed at MIDUS 1 (MDD, GAD, panic disorder, and neuroticism) was associated with mortality risk by use of hierarchical regression. A set of four two-stage hierarchical regression models were conducted where all covariates were entered at stage 1 and each of the indicators was entered one at a time at stage 2. Results showed that MDD, GAD and neuroticism significantly predicted mortality risk in this framework, while panic disorder did not (see Table 1).

We then explored whether the transdiagnostic internalizing factor predicted mortality risk. Our CFA model of internalizing showed an excellent fit to the data: root mean square error of approximation (RMSEA) = .008; comparative fit index (CFI) = .999; Tucker-Lewis index (TLI) = .998. The factor determinacy index for our CFA model was .78, which mirrored the recommended threshold of 0.80 to indicate that a model is "adequate for most scientific purposes"[43]. The Cox regression analysis showed that internalizing significantly and positively predicted mortality risk (HR=1.12, 95% CI: 1.05-1.16, p<0.01), after adjusting for age, age squared, sex, education level, experienced severe health condition, and heart disease family risk (see Table 2).

In order to compare the prediction of mortality risk from internalizing disorders' shared variance (i.e., transdiagnostic internalizing) and disorder-specific variance, we regressed mortality on internalizing and the residual variance of MDD, GAD, panic disorder and neuroticism simultaneously in the Cox regression framework (see Figure 1). The results showed that internalizing significantly predicted mortality risk across all analyses (HR ranged from 1.11 to 1.14), while MDD (HR=1.02, 95% CI: 0.98-1.06), GAD (HR=1.01, 95% CI: 0.94-1.09), panic (HR=0.94, 95% CI: 0.88-1.00), and neuroticism (HR=1.02, 95% CI: 0.92-1.12) residuals did not (see Table 3).

We then examined how the initial associations of MDD, GAD and neuroticism with mortality risk (panic disorder did not significantly predict that risk) were attenuated when adjusting for internalizing. Compared with the hierarchical regression analysis results, the comparative predictive validity analysis showed that the degrees to which MDD, GAD and neuroticism predicted mortality risk were attenuated, respectively, by 67.2%, 86.9% and 87.1%, when their shared variance captured in internalizing was accounted for.

To explore whether the association between internalizing and mortality risk was moderated by individuals' self-reported health condition, we created an interaction term and included it in our Cox regression model following the method use by Gale et al[15]. We then compared the models with and without the interaction term. Results showed that the Bayesian information criterion (BIC) was lower for the model with the interaction term included (BIC=10622.6) than for the model without the interaction term (BIC=10635.3). BIC differences of 10 between two models indicate 150:1 posterior odds in favor of the model with superior (lower) BIC.

Given the statistical significance of the interaction term, we further analyzed the association of internalizing with mortal-

Table 1 Hierarchical regression analysis of individual internalizing pathologies predicting mortality risk (regression coefficients with 95% CI)

	Stage 1	Stage 2			
		MDD	GAD	PAN	NEURO
Age	8.39 (8.19-8.59)***	8.69 (8.49-8.89)***	8.52 (8.32-8.73)***	8.47 (8.27-8.67)***	8.66 (8.45-8.86)***
Age squared	0.94 (0.66-1.21)	0.94 (0.66-1.21)	0.94 (0.66-1.21)	0.94 (0.66-1.21)	0.94 (0.66-1.20)
Sex	0.75 (0.63-0.88)***	0.74 (0.62-0.86)***	0.75 (0.63-0.87)***	0.75 (0.63-0.87)***	0.75 (0.63-0.87)***
Education level	0.82 (0.76-0.88)***	0.82 (0.76-0.88)***	0.82 (0.76-0.88)***	0.82 (0.76-0.88)***	0.82 (0.76-0.88)***
Experienced physical illness	1.52 (1.45-1.60)***	1.51 (1.43-1.58)***	1.52 (1.44-1.59)***	1.52 (1.44-1.59)***	1.50 (1.43-1.58)***
Heart disease family risk	1.23 (1.12-1.35)**	1.23 (1.12-1.35)**	1.23 (1.12-1.35)**	1.23 (1.12-1.35)**	1.23 (1.12-1.35)**
MDD		1.06 (1.02-1.09)**			
GAD			1.08 (1.02-1.15)*		
PAN				1.03 (0.96-1.09)	
NEURO					1.12 (1.02-1.21)*

MDD – major depressive disorder, GAD – generalized anxiety disorder, PAN – panic disorder, NEURO – neuroticism
*p<0.05, **p<0.01, ***p<0.001

Table 2 Results for Cox regression models of the effect of change in internalizing on mortality risk

Predictor	Hazard ratio (95% CI)
Age	5.50 (5.29-5.71)***
Age squared	0.95 (0.75-1.16)
Sex	1.28 (1.18-1.37)***
Education level	0.86 (0.81-0.91)***
Experienced physical illness	1.37 (1.31-1.44)***
Heart disease family risk	1.18 (1.09-1.28)***
Internalizing	1.12 (1.05-1.16)**
AIC	10588.07
BIC	10635.26

AIC – Akaike information criterion, BIC – Bayesian information criterion
**p<0.01, ***p<0.001

ity risk stratified by self-rated physical health level (5-point scale ranged from poor to excellent, with 5 being excellent). Results showed that internalizing significantly predicted mortality risk specifically among individuals whose self-reported physical health was excellent (HR=1.50, 95% CI: 1.17-1.84, p<0.05), but not in individuals with poorer self-rated physical health (see Table 4).

DISCUSSION

Internalizing and mortality risk

The primary aim of our study was to investigate the association between a transdiagnostic internalizing factor and mortality risk. Our findings show that higher levels of internalizing pathology are associated with a significantly increased mortality risk, even after adjusting for covariates known to affect that risk (e.g., age, sex, education level, heart disease family risk, experienced severe health condition). There was a 12.3% increase in mortality rate for every 1-standard deviation unit increment in the internalizing factor level. These findings are consistent with the previously reported close link between individual internalizing disorders and high mortality rates[8-11,28,33,44], but our study is the first to demonstrate that it is a transdiagnostic internalizing factor that predicts mortality risk, rather than the variance that is unique to MDD, GAD, panic, or neuroticism.

There are several possible explanations for why mortality rates are greater in individuals with higher levels of internalizing. One pathway is via maladaptive coping. Given that individuals with high internalizing experience frequent negative affect, they may attempt to manage their negative emotions via unhealthy coping, such as heavy drinking or drug abuse. Indeed, an internalizing pathway model has been proposed[45], in which early and persistent internalizing symptoms lead individuals to use substances as a means of coping. Issues with substance and alcohol abuse tend to emerge after trauma exposure[46], which also has additive negative effects on mental and physical health.

It is also possible that internalizing predicts mortality risk through physical inactivity. People with internalizing psychopathology tend to be physically inactive[47,48], which can lead to adverse physical health outcomes, eventually resulting in high mortality rates. Indeed, a number of prior studies indicated that physical inactivity is one of the main risk factors for cardiovascular disease[49-51], and that engaging in regular activity can meaningfully reduce the risk of premature death[52,53].

In addition, people with high levels of internalizing are more likely to experience various adverse life outcomes, such as unemployment[54], marital discord[55,56], poor social functioning[57,58], and poor quality of life[57,59], which may play a role as mediating factors in the relationship between internalizing and high mortality risk.

The superior predictive validity of the transdiagnostic internalizing factor

Contrary to the underlying assumption of traditional diagnostic systems that each mental disorder is a discrete entity, many internalizing disorders co-occur more frequently than expected

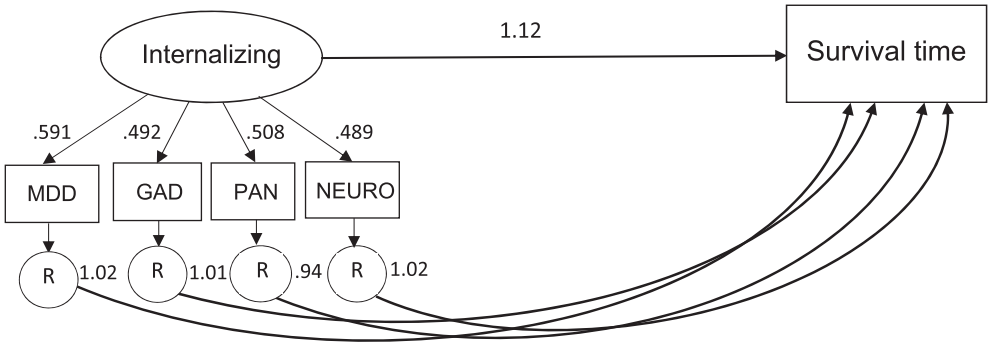

Figure 1 Comparative predictive validity analysis. Arrows flowing from the latent internalizing factor to its indicators represent factor loadings, which were all statistically significant at p<0.001. The arrow leading from internalizing to survival time represents the hazard ratio of the Cox regression model, which was significant at p<0.01. Arrows leading from each of the residual variance to survival time represent the hazard ratios of each Cox regression model, which were all non-significant. MDD – major depressive disorder, GAD – generalized anxiety disorder, PAN – panic disorder, NEURO – neuroticism, R = residual variance.

Table 3 Results of the comparative predictive validity analysis (hazard ratios with 95% CI)

	Models			
	Internalizing vs. MDD	**Internalizing vs. GAD**	**Internalizing vs. PAN**	**Internalizing vs. NEURO**
Age	5.48 (5.28-5.69)***	5.49 (5.28-5.71)***	5.47 (5.26-5.69)***	5.50 (5.29-5.72)***
Age squared	0.95 (0.74-1.17)	0.95 (0.75-1.17)	0.95 (0.74-1.17)	0.95 (0.75-1.17)
Sex	1.27 (1.18-1.37)***	1.27 (1.18-1.37)***	1.27 (1.17-1.37)***	1.28 (1.18-1.38)***
Education level	0.86 (0.81-0.92)***	0.86 (0.81-0.92)***	0.86 (0.81-0.92)***	0.86 (0.81-0.92)***
Experienced physical illness	1.37 (1.31-1.44)***	1.38 (1.31-1.44)***	1.37 (1.31-1.43)***	1.37 (1.31-1.43)***
Heart disease family risk	1.18 (1.09-1.28)***	1.18 (1.09-1.28)***	1.18 (1.09-1.28)***	1.18 (1.09-1.28)***
Internalizing	1.11 (1.03-1.18)**	1.12 (1.04-1.20)**	1.14 (1.07-1.22)***	1.12 (1.05-1.20)**
MDD residual	1.02 (0.98-1.06)			
GAD residual		1.01 (0.94-1.09)		
PAN residual			0.94 (0.88-1.00)	
NEURO residual				1.02 (0.92-1.12)

MDD – major depressive disorder, GAD – generalized anxiety disorder, PAN – panic disorder, NEURO – neuroticism
**p<0.01, ***p<0.001

by chance. A significant relationship between a particular internalizing disorder and high mortality risk can be attributed to that disorder's unique variance or to the common variance that the disorder shares with other internalizing disorders (i.e., the transdiagnostic internalizing factor). To our knowledge, our study is the first to explore which of these two sources more significantly predicts mortality risk. Our findings show that, once the commonalities among the individual diagnostic constructs are accounted for by the internalizing factor, the unique variance that is specific to each construct no longer predicts mortality risk meaningfully. The significant associations between particular internalizing pathologies and mortality risk reported in prior studies may be therefore largely attributed to an underlying internalizing factor.

Of note, internalizing accounted for 34.93% of the variance in MDD (i.e., 65.07% of the variance was MDD-specific), 24.21% in GAD (i.e., 75.79% of the variance was GAD-specific), 25.81% in panic disorder (i.e., 74.19% of the variance was panic disorder-specific), and 23.91% in neuroticism (i.e., 76.09% of the variance was neuroticism-specific). Nevertheless, none of those disorder-specific variances significantly predicted mortality risk.

It is worth further speculating about why mortality risk was mainly predicted by the commonalities among the internalizing pathologies rather than by the residual variance of each internalizing construct. Previous studies that looked at the stability of various diagnostic constructs reported a long-term instability of mood and anxiety disorders[60-62]. Given the transitory nature of these pathologies, one's particular internalizing disorder symptoms may not be a strong predictor of long-term outcomes. By contrast, a transdiagnostic internalizing dimension has high temporal stability and structural invariance over time[29,30], thus being a more reliable prospective predictor of long-term outcomes such as death.

It was especially notable to find the insignificant association between neuroticism residual variance and mortality risk, after the common variance was saturated by the internalizing fac-

Table 4 Results for Cox regression models of the effect of change in internalizing on mortality risk stratified by self-rated physical health (hazard ratios with 95% CI)

	Excellent health (N=1,217)	**Very good health (N=2,506)**	**Good health (N=2,386)**	**Fair health (N=796)**	**Poor health (N=192)**
Age	7.86 (7.21-8.51)***	2.59 (2.36-2.83)***	2.69 (2.50-2.87)***	2.28 (2.04-2.53)***	1.86 (1.57-2.15)***
Age squared	1.16 (0.54-1.78)	0.98 (0.72-1.23)	1.09 (0.87-1.32)	0.91 (0.63-1.20)	0.77 (0.43-1.12)
Sex	1.11 (0.81-1.4)	1.14 (1.04-1.24)**	1.14 (1.05-1.22)**	1.22 (1.10-1.34)**	1.07 (0.88-1.26)
Education level	0.88 (0.75-1.02)	0.92 (0.82-1.02)	0.86 (0.77-0.95)**	0.97 (0.85-1.09)	1.02 (0.84-1.19)
Experienced physical illness	1.54 (1.33-1.75)***	1.06 (1.00-1.13)*	1.19 (1.11-1.26)***	1.19 (1.07-1.31)**	1.23 (1.03-1.44)*
Heart disease family risk	1.54 (1.26-1.82)**	1.03 (0.93-1.13)	1.12 (1.03-1.20)**	1.05 (0.94-1.16)	1.01 (0.83-1.18)
Internalizing	1.50 (1.17-1.84)*	1.04 (0.93-1.14)	1.04 (0.93-1.16)	1.04 (0.89-1.19)	0.82 (0.59-1.06)

*p<0.05, **p<0.01, ***p<0.001

tor. While a number of prior studies have reported a significant connection between neuroticism and mortality risk[15,16,28], our findings indicate that such significant effects may be largely accounted for by the common variance that neuroticism shares with other internalizing pathologies, and that neuroticism itself may not be a consistent predictor of mortality risk[63].

The significant interaction between internalizing and self-rated health condition

Of note, the meaningful association between internalizing and mortality risk was moderated by one's self-reported health condition. That is, internalizing predicted mortality risk in individuals whose self-reported physical health was excellent, but not in individuals with poorer self-reported health. Individuals with excellent self-rated physical health are exposed to a 50.2% increase in the risk of premature death for every 1-standard deviation unit increment in the internalizing level, after adjusting for other covariates.

This significant interaction between internalizing and self-reported health indicates that internalizing psychopathology may not confer additional risk of early death to those with poor physical health. However, if individuals are currently physically healthy, then internalizing psychopathology is more likely to have an effect on mortality risk. This is in line with prior findings that the association between individual internalizing pathologies and mortality risk was moderated by one's self-reported health condition[11,15]. It is likely that self-reported health is a strong predictor of mortality, closely covarying with internalizing in our study.

Limitations

The current study was not without limitations. First, we included only four internalizing indicators (MDD, GAD, panic disorder, and neuroticism), since they were the only internalizing pathologies assessed in MIDUS 1. Future research needs to replicate our findings by including other internalizing disorder indicators. Second, the study was unable to test the association of a transdiagnostic externalizing factor with mortality risk, given that sufficient indicators were not available to model an externalizing factor in MIDUS. Third, although our final Cox regression model adjusted for many covariates, we were unable to control for some other covariates also known to influence mortality risk (e.g., body mass index, smoking status), due to large missing values for those variables. Fourth, given that the information regarding causes of death was not available in MIDUS, the current study was unable to further investigate the degree to which internalizing predicts mortality risk differently depending on the various causes of death. Last, our study included a binary sex variable as one of the covariates: future research could further explore how the association of internalizing with mortality risk differs in non-binary individuals.

CONCLUSIONS

Our study is the first to identify the role of the transdiagnostic internalizing factor in predicting the risk of early death. Results show that one's level of internalizing meaningfully predicts mortality risk over a 20-year period, and that internalizing outperforms disorder-specific variance in the prediction of that risk. Moreover, the significant interaction between internalizing and physical health indicates that the former dimension is more likely to have an effect on early death for currently physically healthy individuals.

These findings highlight the clinical utility of using the transdiagnostic internalizing factor for prediction of an important future outcome, and support the argument that internalizing psychopathology can be a meaningful liability to incorporate into intervention and prevention research, and to explore in public health practice.

APPENDIX

The members of the Hierarchical Taxonomy of Psychopathology (HiTOP) Utility Workgroup include: Christopher C. Conway (Fordham University), Anna R. Docherty (University of Utah), Michael Dretsch (Walter Reed Army Institute of Research), Kelsie T. Forbush (University of Kansas), Vina M. Goghari (University of Toronto), Kristian E. Markon (University of Iowa), Stephanie N. Mullins-Sweatt (Oklahoma State University), Brady Nelson (Stony Brook University), Thomas M. Olino (Temple University), and Tim Slade (University of Sydney).

REFERENCES

1. Penninx BW, Guralnik JM, Ferrucci L et al. Depressive symptoms and physical decline in community-dwelling older persons. JAMA 1998;279:1720-6.
2. Frasure-Smith N, Lespérance F, Talajic M. Depression and 18-month prognosis after myocardial infarction. Circulation 1995;91:999-1005.
3. Ford DE, Mead LA, Chang PP et al. Depression is a risk factor for coronary artery disease in men: the precursors study. Arch Intern Med 1998;158:1422-6.
4. Gross AL, Gallo JJ, Eaton WW. Depression and cancer risk: 24 years of follow-up of the Baltimore Epidemiologic Catchment Area sample. Cancer Causes Control 2010;21:191-9.
5. Battaglia M, Garon-Carrier G, Côté SM et al. Early childhood trajectories of separation anxiety: bearing on mental health, academic achievement, and physical health from mid-childhood to preadolescence. Depress Anxiety 2017;34:918-27.
6. Tully PJ, Cosh SM, Baune BT. A review of the affects of worry and generalized anxiety disorder upon cardiovascular health and coronary heart disease. Psychol Health Med 2013;18:627-44.
7. Shalev I, Moffitt TE, Braithwaite AW et al. Internalizing disorders and leukocyte telomere erosion: a prospective study of depression, generalized anxiety disorder and post-traumatic stress disorder. Mol Psychiatry 2014;19:1163-70.
8. Doyle F, Conroy R, McGee H. Differential predictive value of depressive versus anxiety symptoms in the prediction of 8-year mortality after acute coronary syndrome. Psychosom Med 2012;74:711-6.
9. Miloyan B, Bulley A, Bandeen-Roche K et al. Anxiety disorders and all-cause mortality: systematic review and meta-analysis. Soc Psychiatry Psychiatr Epidemiol 2016;51:1467-75.
10. Van Dijk MR, Utens EM, Dulfer K et al. Depression and anxiety symptoms as predictors of mortality in PCI patients at 10 years of follow-up. Eur J Prev Cardiol 2016;23:552-8.
11. 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.
12. Carriere I, Ryan J, Norton J et al. Anxiety and mortality risk in community-dwelling elderly people. Br J Psychiatry 2013;203:303-9.
13. Denollet J, Maas K, Knottnerus A et al. Anxiety predicted premature all-cause and cardiovascular death in a 10-year follow-up of middle-aged women. J Clin Epidemiol 2009;62:452-6.

14. Clark LA, Watson D, Mineka S. Temperament, personality, and the mood and anxiety disorders. J Abnorm Psychol 1994;103:103-16.
15. Gale CR, Čukić I, Batty GD et al. When is higher neuroticism protective against death? Findings from UK Biobank. Psychol Sci 2017;28:1345-57.
16. Mroczek DK, Spiro III A, Turiano NA. Do health behaviors explain the effect of neuroticism on mortality? Longitudinal findings from the VA Normative Aging Study. J Res Personal 2009;43:653-9.
17. Achenbach TM, Edelbrock CS. The classification of child psychopathology: a review and analysis of empirical efforts. Psychol Bull 1978;85:1275-301.
18. Kotov R, Krueger RF, Watson D et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): a dimensional alternative to traditional nosologies. J Abnorm Psychol 2017;126:454-77.
19. Krueger RF. The structure of common mental disorders. Arch Gen Psychiatry 1999;56:921-6.
20. Krueger RF, Caspi A, Moffitt TE et al. The structure and stability of common mental disorders (DSM-III-R): a longitudinal-epidemiological study. J Abnorm Psychol 1998;107:216-27.
21. Krueger RF, Kotov R, Watson D et al. Progress in achieving quantitative classification of psychopathology. World Psychiatry 2018;17:282-93.
22. Krueger RF, Eaton NR. Structural validity and the classification of mental disorders. In: Kendler KS, Parnas J (eds). Philosophical issues in psychiatry II: Nosology. Oxford: Oxford University Press, 2012:199-212.
23. Mineka S, Watson D, Clark LA. Comorbidity of anxiety and unipolar mood disorders. Annu Rev Psychol 1998;49:377-412.
24. Seeley JR, Kosty DB, Farmer RF et al. The modeling of internalizing disorders on the basis of patterns of lifetime comorbidity: associations with psychosocial functioning and psychiatric disorders among first-degree relatives. J Abnorm Psychol 2011;120:308-21.
25. Eaton NR, South SC, Krueger RF. The meaning of comorbidity among common mental disorders. In: Millon T, Krueger RF, Simonsen E (eds). Contemporary directions in psychopathology: scientific foundations of the DSM-V and ICD-11. New York: Guilford, 2010:223-41.
26. Conway CC, Forbes MK, Forbush KT et al. A hierarchical taxonomy of psychopathology can transform mental health research. Perspect Psychol Sci 2019;14:419-36.
27. Mirza S, Arfan Ikram M, Hofman A et al. Anxiety does not predict mortality. A population-based study. World Psychiatry 2015;14:103-4.
28. Mroczek DK, Spiro A. Personality change influences mortality in older men. Psychol Sci 2007;18:371-6.
29. Eaton NR, Keyes KM, Krueger RF et al. An invariant dimensional liability model of gender differences in mental disorder prevalence: evidence from a national sample. J Abnorm Psychol 2012;121:282-8.
30. Eaton NR, Krueger RF, Markon KE et al. The structure and predictive validity of the internalizing disorders. J Abnorm Psychol 2013;122:86-92.
31. Kim H, Eaton NR. The hierarchical structure of common mental disorders: connecting multiple levels of comorbidity, bifactor models, and predictive validity. J Abnorm Psychol 2015;124:1064-78.
32. Kessler RC, Ormel J, Petukhova M et al. Development of lifetime comorbidity in the World Health Organization World Mental Health Surveys. Arch Gen Psychiatry 2011;68:90-100.
33. Boscarino JA. Posttraumatic stress disorder and mortality among US Army veterans 30 years after military service. Ann Epidemiol 2006;16:248-56.
34. Brim OG, Ryff CD, Kessler RC. How healthy are we?: A national study of well-being at midlife. Chicago: University of Chicago Press, 2019.
35. Kessler RC, Abelson J, Demler O et al. Clinical calibration of DSM-IV diagnoses in the World Mental Health (WMH) version of the World Health Organization (WHO) Composite International Diagnostic Interview (WMH-CIDI). Int J Methods Psychiatr Res 2004;13:122-39.
36. Kessler RC, Andrews G, Mroczek D et al. The World Health Organization Composite International Diagnostic Interview Short-Form (CIDI-SF). Int J Methods Psychiatr Res 1998;7:171-85.
37. Kessler RC, Üstün 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.
38. Lachman ME, Weaver SL. The Midlife Development Inventory (MIDI) personality scales: scale construction and scoring. Waltham: Brandeis University, 1997.
39. Eaton NR, Krueger RF, Oltmanns TF. Aging and the structure and long-term stability of the internalizing spectrum of personality and psychopathology. Psychol Aging 2011;26:987-93.
40. Kleinbaum DG, Klein M. Survival analysis, 3rd ed. New York: Springer, 2010.
41. Hess KR. Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline functions. Stat Med 1994;13:1045-62.
42. Ng'andu NH. An empirical comparison of statistical tests for assessing the proportional hazards assumption of Cox's model. Stat Med 1997;16:611-26.
43. Gorsuch RL. Factor analysis: classic edition. London: Taylor & Francis, 2014.
44. Kiecolt-Glaser JK, Glaser R. Depression and immune function: central pathways to morbidity and mortality. J Psychosom Res 2002;53:873-6.
45. Hussong AM, Jones DJ, Stein GL et al. An internalizing pathway to alcohol use and disorder. Psychol Addict Behav 2011;25:390-404.
46. Forbes MK, Flanagan JC, Barrett EL et al. Smoking, posttraumatic stress disorder, and alcohol use disorders in a nationally representative sample of Australian men and women. Drug Alcohol Depend 2015;156:176-83.
47. Carek PJ, Laibstain SE, Carek SM. Exercise for the treatment of depression and anxiety. Int J Psychiatry Med 2011;41:15-28.
48. Strine TW, Mokdad AH, Dube SR et al. The association of depression and anxiety with obesity and unhealthy behaviors among community-dwelling US adults. Gen Hosp Psychiatry 2008;30:127-37.
49. Carnethon MR. Physical activity and cardiovascular disease: how much is enough? Am J Lifestyle Med 2009;3(Suppl. 1):44S-9S.
50. Kohl HW. Physical activity and cardiovascular disease: evidence for a dose response. Med Sci Sports Exerc 2001;33:S472-83.
51. Ahmed HM, Blaha MJ, Nasir K et al. Effects of physical activity on cardiovascular disease. Am J Cardiol 2012;109:288-95.
52. Woodcock J, Franco OH, Orsini N et al. Non-vigorous physical activity and all-cause mortality: systematic review and meta-analysis of cohort studies. Int J Epidemiol 2011;40:121-38.
53. Oguma Y, Shinoda-Tagawa T. Physical activity decreases cardiovascular disease risk in women: review and meta-analysis. Am J Prev Med 2004;26:407-18.
54. Borenstein M, Hedges L, Higgins J et al. Comprehensive Meta Analysis Version 2. Englewood: Biostat, 2005.
55. O'Leary KD, Christian JL, Mendell NR. A closer look at the link between marital discord and depressive symptomatology. J Soc Clin Psychol 1994;13:33-41.
56. Trudel G, Goldfarb M. Marital and sexual functioning and dysfunctioning, depression and anxiety. Sexologies 2010;19:137-42.
57. Comer JS, Blanco C, Hasin DS et al. Health-related quality of life across the anxiety disorders. J Clin Psychiatry 2011;72:43-50.
58. Hecht H, von Zerssen D, Wittchen H-U. Anxiety and depression in a community sample: the influence of comorbidity on social functioning. J Affect Disord 1990;18:137-44.
59. Sareen J, Jacobi F, Cox BJ et al. Disability and poor quality of life associated with comorbid anxiety disorders and physical conditions. Arch Intern Med 2006;166:2109-16.
60. Andrews G. Should depression be managed as a chronic disease? BMJ 2001; 322:419-21.
61. Thompson KD, Leadbeater BJ, Ames ME. Reciprocal effects of internalizing and oppositional defiance symptoms on heavy drinking and alcohol-related harms in young adulthood. Substance Abuse: Research and Treatment 2015;9(Suppl. 1):21-31.
62. Wittchen H-U. Natural course and spontaneous remissions of untreated anxiety disorders: results of the Munich Follow-up Study (MFS). In: Hand I, Wittchen HU (eds). Panic and phobias 2. Berlin: Springer, 1988:3-17.
63. Turiano NA, Graham EK, Weston S et al. Is healthy neuroticism associated with longevity? A coordinated integrative data analysis. Collabra: Psychology 2020;6:32.

DOI:10.1002/wps.20859

Psychotherapies for depression: a network meta-analysis covering efficacy, acceptability and long-term outcomes of all main treatment types

Pim Cuijpers[1], Soledad Quero[2,3], Hisashi Noma[4], Marketa Ciharova[1], Clara Miguel[1], Eirini Karyotaki[1], Andrea Cipriani[5,6], Ioana A. Cristea[7,8], Toshi A. Furukawa[9]

[1]Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; [2]Department of Basic and Clinical Psychology and Psychobiology, Universitat Jaume I, Castellón, Spain; [3]CIBER of Physiopathology of Obesity and Nutrition (CIBEROBN), Madrid, Spain; [4]Department of Data Science, Institute of Statistical Mathematics, Tokyo, Japan; [5]Department of Psychiatry, University of Oxford, Oxford, UK; [6]Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford, UK; [7]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy; [8]IRCCS Mondino Foundation, Pavia, Italy; [9]Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine, School of Public Health, Kyoto, Japan

The effects of psychotherapies for depression have been examined in several hundreds of randomized trials, but no recent network meta-analysis (NMA) has integrated the results of these studies. We conducted an NMA of trials comparing cognitive behavioural, interpersonal, psychodynamic, problem-solving, behavioural activation, life-review and "third wave" therapies and non-directive supportive counseling with each other and with care-as-usual, waiting list and pill placebo control conditions. Response (50% reduction in symptoms) was the primary outcome, but we also assessed remission, standardized mean difference, and acceptability (all-cause dropout rate). Random-effects pairwise and network meta-analyses were conducted on 331 randomized trials with 34,285 patients. All therapies were more efficacious than care-as-usual and waiting list control conditions, and all therapies – except non-directive supportive counseling and psychodynamic therapy – were more efficacious than pill placebo. Standardized mean differences compared with care-as-usual ranged from –0.81 for life-review therapy to –0.32 for non-directive supportive counseling. Individual psychotherapies did not differ significantly from each other, with the only exception of non-directive supportive counseling, which was less efficacious than all other therapies. The results were similar when only studies with low risk of bias were included. Most therapies still had significant effects at 12-month follow-up compared to care-as-usual, and problem-solving therapy was found to have a somewhat higher long-term efficacy than some other therapies. No consistent differences in acceptability were found. Our conclusion is that the most important types of psychotherapy are efficacious and acceptable in the acute treatment of adult depression, with few significant differences between them. Patient preference and availability of each treatment type may play a larger role in the choice between types of psychotherapy, although it is possible that a more detailed characterization of patients with a diagnosis of depression may lead to a more precise matching between individual patients and individual psychotherapies.

Key words: Depression, psychotherapy, network meta-analysis, cognitive behavioural therapy, behavioural activation therapy, problem-solving therapy, interpersonal psychotherapy, psychodynamic therapy, life-review therapy, "third wave" therapies

(World Psychiatry 2021;20:283–293)

Depressive disorders are common[1], costly[2,3], have a strong impact on quality of life of patients[4], and are associated with considerable morbidity and mortality[5]. Next to antidepressants, psychotherapies are first-line treatments for depression, and both treatments are effective[6,7].

Cognitive behavioural therapy (CBT) is the most examined type of psychotherapy for depression[8], but several other types of psychotherapy have also been tested in multiple trials, including interpersonal, psychodynamic, life-review, problem-solving, behavioural activation and "third wave" therapies and non-directive supportive counseling. For all these therapies, there is evidence of efficacy in comparison with care-as-usual and waiting list[9].

Head-to-head comparisons of different types of psychotherapy indicate no significant differences between them[10]. However, these findings should be considered with caution, because more than 70% of trials in this field have considerable risk of bias[9]. Furthermore, almost all comparative outcome trials are heavily underpowered[11].

Only one network meta-analysis (NMA) has examined simultaneously the effects of different psychotherapies for depression[7], confirming the comparable effects of these therapies versus control conditions. However, this previous NMA is outdated (only studies up to 2012 were included, and a considerable number of trials has been conducted since then) and did not examine acceptability of treatments. Also, the number of trials with low risk of bias was small and has substantially increased since then. Long-term outcomes of psychotherapies have also not yet been examined in an NMA. Furthermore, the methodology of NMAs has been developed considerably in the past few years, with more sophisticated techniques.

We decided, therefore, to conduct a new NMA examining the efficacy and acceptability of the main types of psychotherapy for adult depression compared to care-as-usual, waiting list and pill placebo.

METHODS

Identification and selection of studies

The protocol for the current NMA has been registered at the Open Science Foundation (https://osf.io/nxvye). We used a database of studies on psychotherapies for depression[12] which is continuously updated and covered the period from 1966 to January 1, 2020. For this database, we searched four major bibliographic sources (PubMed, PsycINFO, EMBASE and Cochrane

Library) by combining terms for depression and psychotherapies, with filters for randomized controlled trials (the full search string in PubMed is provided in the supplementary information). We also checked the references of earlier meta-analyses.

All records were screened by two independent researchers, and all papers potentially meeting inclusion criteria according to one of the researchers were retrieved as full text. The decision to include or exclude a study in the database was also done by the two independent researchers, and disagreements were solved through discussion.

We included randomized trials in which one of eight major types of psychotherapy for adult depression was compared with another major type of psychotherapy or one of three types of control conditions: waiting list, care-as-usual, and pill placebo. The definitions of the eight major types of psychotherapy were developed by experts in the field, based on the critical reading and analysis of therapies described in comparative outcomes trials of psychotherapy for depression[10].

The therapies that were examined were: CBT, behavioural activation therapy, problem-solving therapy, "third wave" therapies, interpersonal psychotherapy, psychodynamic therapy, non-directive supportive counseling, and life-review therapy. The classification of psychotherapies was made by two independent raters. Any disagreement was resolved through discussion of the two and/or in consultation with the first author. Each of these major types of psychotherapy was examined in at least ten trials comparing the therapy with a control condition.

Depression could be established by a diagnostic interview or by a score above a cutoff on a validated self-report measure. Studies of comorbid mental or physical disorders were included. Studies on inpatients were excluded[13], as were maintenance treatment studies. Psychotherapies could be delivered individually, in groups, by telephone, or as guided Internet-based treatment. Unguided interventions were excluded, because they have been found to be less effective than interventions with human contact between a patient and a therapist[14].

Quality assessment

We evaluated the included studies using four criteria of the Risk of Bias assessment tool developed by the Cochrane Collaboration[15]: adequate generation of allocation sequence; concealment of allocation to conditions; prevention of knowledge of the allocated intervention (masking of assessors); and dealing with incomplete outcome data. Assessment of risk of bias was conducted by two independent researchers, and disagreements were solved through discussion. A study was rated as low overall risk of bias when all four items were rated as low risk of bias.

Outcome measures

Treatment response, defined as a reduction of at least 50% in depressive symptomatology, was chosen as the primary outcome. When not reported, we imputed response rates using a validated method[16]. Patients randomized but not included in the analyses of responders in the original reports were assumed to be non-responders and included in the current analyses in order to abide the intention-to-treat principle.

The time point for the primary outcome was the end of the psychotherapy. When more than one depression measure was used in a study, we selected one outcome using an algorithm (see supplementary information). When a study included two or more arms of the same type of psychotherapy (e.g., individual and group CBT), the outcome data were pooled so that each study had only one outcome for one type of therapy.

We also calculated remission rates. For the selection of definitions of remission, we used the following hierarchy: a) no diagnosis of major depressive disorder; b) scoring below a specific cutoff score; c) other (e.g., significant change). In addition, we calculated the standardized mean difference (SMD) between conditions for the studies that reported means, standard deviations and number of patients at baseline and post-test, or the change score between baseline and post-test. Acceptability of the treatments was operationalized as all-cause dropout rate.

Pairwise meta-analyses

We conducted pairwise meta-analyses for all comparisons, using a random effects model. To quantify heterogeneity, we calculated the I^2-statistic with 95% confidence intervals (CIs)[17]. We tested for small study effects with Egger's test[18].

Network meta-analyses

The comparative effectiveness was evaluated using the NMA methodology via combining direct and indirect evidence for all relative treatment effects. First, we summarized the geometry of the network of evidence using network plots[19]. Second, the NMA for assessing the comparative efficacy or acceptability was conducted using contrast-based methods. Comparative odds ratios (ORs) and SMDs were reported with their 95% CIs. The ranking of treatment formats was estimated according to the "surface under the cumulative ranking" (SUCRA), based on the estimated multivariate random effects models[19].

The statistical examination of the transitivity assumption was conducted using tests of local and global inconsistency[20]. We also implemented meta-regression analyses to evaluate the influence of small study effects involving the study-specific variances as a covariate[21].

Further, we evaluated the heterogeneity in the network with tau-squared in comparison with empirically derived evidence[22,23], and conducted a multivariate meta-regression analysis to examine possible sources of heterogeneity with core characteristics of the studies.

We performed several sensitivity analyses: a) analyses with only studies with low risk of bias; b) analyses excluding life-re-

view therapy (this is only used in older adults, and may violate the transitivity assumption); and c) analyses in which studies with pill placebo were excluded (because in these studies patients could also be randomized to antidepressant medication, which may violate the transitivity assumption as well).

We assessed the certainty of evidence in network estimates of the main outcome in accordance with the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) framework.

The main analyses were conducted in Stata/SE 14.2 for Mac, except the meta-regression analyses examining small sample bias, which were conducted in OpenBUGS 3.2.3. The GRADE ratings were performed in CINeMA[24].

RESULTS

Selection and inclusion of studies

After examining 24,647 abstracts (18,217 after removal of duplicates), we retrieved 2,914 full-text papers, of which 2,583 were excluded. The PRISMA flow chart is presented in Figure 1. A total of 331 randomized controlled trials (with 34,285 patients) met inclusion criteria.

Characteristics and risk of bias of included studies

The aggregated characteristics of the 331 included studies are presented in Table 1. Most studies were aimed at adults in general (145; 43.8%). In 179 studies (54.1%), participants met criteria for a depressive disorder according to a diagnostic interview, while the other studies (152; 45.9%) included participants who scored above a cutoff on a self-rating depression scale.

CBT was examined in the majority of studies (211 trials; 63.7%), while the other therapies were examined in 13 (3.9%; life-review) to 42 (12.7%; non-directive supportive counseling) studies. Care-as-usual control condition was used in 158 studies (47.7%), waiting list in 112 studies (33.8%), and pill placebo in 10 studies (3.0%). Most interventions had an individual treatment format (145; 43.8%), 75 used a group format (22.7%), 58 used guided self-help (17.5%), and 53 used a mixed or another format (16.0%). Most studies were conducted in North America (134; 40.5%) and Europe (124; 37.5%).

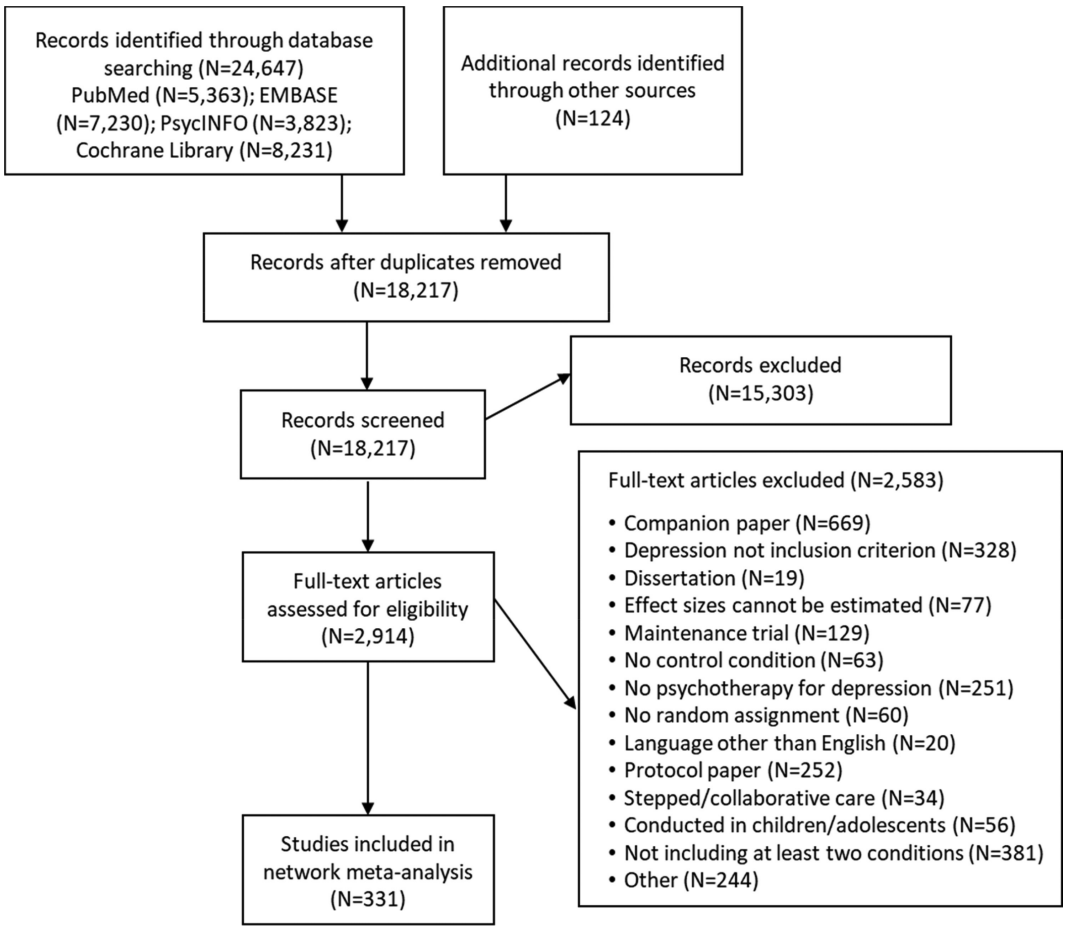

Figure 1 PRISMA flow chart for inclusion of studies

Table 1 Aggregated characteristics of the included studies (N=331)

		N	%
Recruitment	Community	148	44.7
	Clinical	86	26.0
	Other	97	29.3
Target group	Adults in general	145	43.8
	Older adults	14	4.2
	Students	32	9.7
	Perinatal depression	30	9.1
	General medical disorder	67	20.2
	Other specific group	43	13.0
Diagnosis	Depressive disorder	179	54.1
	Scoring above cutoff	152	45.9
Conditions	Cognitive behavioural therapy	211	63.7
	Behavioural activation therapy	36	10.9
	Problem-solving therapy	33	10.0
	"Third wave" therapies	29	8.8
	Interpersonal psychotherapy	35	10.6
	Psychodynamic therapy	21	6.3
	Non-directive supportive counseling	42	12.7
	Life-review therapy	13	3.9
	Care-as-usual	158	47.7
	Waiting list	112	33.8
	Pill placebo	10	3.0
Number of conditions per study	Two	296	89.4
	Three	32	9.7
	Four	3	0.9
Format	Individual	145	43.8
	Group	75	22.7
	Guided self-help	58	17.5
	Mixed/other	53	16.0
Number of sessions	<8	114	34.4
	8-12	154	46.5
	>12	63	19.0
Country	North America	134	40.5
	Europe	124	37.5
	Australia	23	6.9
	Other	50	15.1
Risk of bias	Adequate sequence generation	184	55.6
	Concealment of allocation to conditions	157	47.4
	Masking of assessors	105	31.7
	Intention-to-treat analysis	209	63.1
Risk of bias total score	Low (4)	102	30.8
	Moderate (2 or 3)	148	44.7
	High (0 or 1)	81	24.4

A total of 184 studies reported adequate sequence generation (55.6%), 157 reported allocation to conditions by an independent party (47.4%), 105 reported using blinded outcome assessors (31.7%), and 195 used only self-report outcomes (58.9%). Intent-to-treat analyses were conducted in 209 studies (63.1%). The risk of bias was low (total score: 4) in 102 studies (30.8%), moderate (total score: 2 or 3) in 148 studies (44.7%), and high (total score: 0 or 1) in 81 studies (24.4%).

Network plot

The network plot for response (Figure 2) indicated a well-connected network, with no stand-alone node. CBT was the best examined therapy and was connected to all other nodes (except life-review therapy). Non-directive supportive counseling was also connected to most other nodes. The other therapies were not connected well with each other. All therapies were connected to care-as-usual and waiting list, but not to pill placebo.

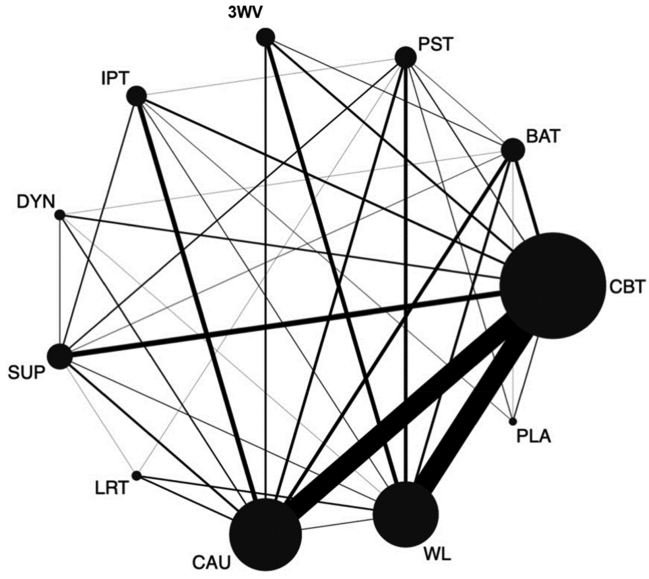

Figure 2 Network plot for response. 3WV – third wave therapies, BAT – behavioural activation therapy, CAU – care-as-usual, CBT – cognitive behavioural therapy, DYN – psychodynamic therapy, IPT – interpersonal psychotherapy, LRT – life-review therapy, PLA – pill placebo, PST – problem-solving therapy, SUP – non-directive supportive counseling, WL – waiting list

Pairwise meta-analyses

In pairwise meta-analyses (see Table 2), all therapies were more efficacious than care-as-usual (except psychodynamic therapy) and waiting list (except non-directive supportive counseling and psychodynamic therapy). There were no significant differences between therapies, except that non-directive supportive counseling was less efficacious than CBT, problem-solving therapy, and psychodynamic therapy.

Although heterogeneity was low in most comparisons, several comparisons (especially involving care-as-usual or non-directive supportive counseling) had an I^2 above 50%.

Network meta-analyses

The main results of the NMA are presented in Tables 3 to 6. The results for response indicate that all therapies are more efficacious than care-as-usual and waiting list, with few significant differences between therapies. Only non-directive supportive counseling was less efficacious than all other therapies, with ORs ranging between 0.49 to 0.65. All therapies, except non-directive supportive counseling and psychodynamic therapy, were also more efficacious than pill placebo. The results for remission and SMD are very similar to those for response. Only the results for pill placebo differ considerably, potentially related to the small number of studies.

The acceptability of all therapies (except interpersonal psychotherapy and life-review therapy) was significantly lower than waiting list, with ORs ranging between 0.49 to 0.67. Psychodynamic therapy was significantly less acceptable than care-as-usual (OR=0.64). No significant differences for acceptability were found between any of the therapies.

The global tau-squared was 0.19 for response. The design-by-treatment interaction model indicated global inconsistency in the network (p for the null hypothesis of consistency in the network <0.01). Consistency factors were examined using the loop specific approach. Considerable inconsistency was found: out of 60 loops, four showed significant inconsistency.

Because of the global inconsistency in the network, we searched for the sources of trial-level influential factors by a bootstrapping method[25]. Through the bootstrap-based evaluation, 37 trials were detected as influential outliers. After excluding these outliers, global inconsistency was no longer significant (p for the null hypothesis of consistency in the network=0.11; global tau-squared: 0.03). The results of the NMA after excluding these outliers were similar to the main analyses (see supplementary information).

Except for some comparisons mainly involving active interventions versus waiting list (CBT, behavioural activation therapy, "third wave" therapies, interpersonal psychotherapy, psychodynamic therapy, and life-review therapy vs. waiting list, and behavioural activation therapy vs. care-as-usual) which had moderate certainty, all the estimates were rated as low to very low certainty of evidence (see supplementary information).

Table 2 Pairwise meta-analyses: efficacy of psychotherapies compared with each other and with control conditions

		N	OR	95% CI	I^2
CBT	BAT	12	0.97	0.74-1.26	0
	PST	4	1.00	0.61-1.61	23
	3WV	8	0.96	0.67-1.36	0
	IPT	8	0.98	0.62-1.54	57
	DYN	7	0.92	0.68-1.23	0
	SUP	20	**0.74**	0.58-0.95	15
	CAU	75	**0.47**	0.39-0.56	60
	WL	77	**0.25**	0.20-0.30	43
	PLA	4	**0.48**	0.30-0.76	20
BAT	PST	2	0.71	0.18-2.87	43
	3WV	3	0.85	0.43-1.68	0
	DYN	1	0.74	0.25-2.18	
	SUP	2	0.31	0.06-1.75	29
	CAU	13	**0.33**	0.20-0.56	46
	WL	9	**0.18**	0.11-0.32	2
	PLA	1	**0.34**	0.15-0.81	
PST	IPT	1	0.37	0.13-1.03	
	SUP	5	**0.38**	0.25-0.57	0
	LRT	1	0.51	0.18-1.50	
	CAU	10	**0.37**	0.19-0.73	77
	WL	13	**0.47**	0.29-0.76	51
	PLA	3	0.65	0.36-1.19	57
3WV	CAU	7	**0.23**	0.09-0.60	63
	WL	15	**0.30**	0.20-0.45	39
IPT	SUP	5	0.64	0.32-1.29	20
	CAU	17	**0.42**	0.26-0.68	69
	WL	3	**0.20**	0.10-0.40	0
	PLA	2	**0.46**	0.23-0.91	0
DYN	SUP	3	**0.34**	0.12-0.97	58
	CAU	5	0.77	0.52-1.12	0
	WL	1	0.16	0.01-3.85	
SUP	LRT	1	3.60	0.34-38.30	
	CAU	8	**0.56**	0.41-0.77	0
	WL	3	0.43	0.09-2.12	0
LRT	CAU	6	**0.06**	0.03-0.13	0
	WL	6	**0.35**	0.22-0.56	1
CAU	WL	3	**0.54**	1.09-2.71	46

Bold prints highlight significant differences. OR – odds ratio, CBT – cognitive behavioural therapy, BAT – behavioural activation therapy, PST – problem-solving therapy, 3WV – "third wave" therapies, IPT – interpersonal psychotherapy, DYN – psychodynamic therapy, SUP – non-directive support counseling, LRT – life-review therapy, CAU – care-as-usual, WL – waiting list, PLA – pill placebo

Table 3 Network meta-analyses: response in psychotherapies compared with each other and with control conditions

CBT										
1.20 (0.90-1.61)	BAT									
0.99 (0.75-1.31)	0.83 (0.57-1.20)	PST								
1.02 (0.76-1.38)	0.85 (0.58-1.25)	1.03 (0.70-1.51)	3WV							
1.00 (0.76-1.31)	0.83 (0.57-1.22)	1.00 (0.70-1.44)	0.98 (0.66-1.45)	IPT						
0.89 (0.62-1.29)	0.74 (0.47-1.17)	0.90 (0.58-1.40)	0.88 (0.55-1.40)	0.90 (0.58-1.39)	DYN					
0.58 (0.45-0.75)	**0.49 (0.34-0.70)**	**0.59 (0.42-0.82)**	**0.57 (0.39-0.84)**	**0.59 (0.42-0.83)**	**0.65 (0.43-0.99)**	SUP				
1.47 (0.87-2.49)	1.23 (0.68-2.20)	1.48 (0.85-2.60)	1.45 (0.81-2.60)	1.48 (0.83-2.63)	1.65 (0.88-3.10)	**2.52 (1.43-4.45)**	LRT			
0.43 (0.37-0.50)	**0.36 (0.26-0.48)**	**0.43 (0.33-0.57)**	**0.42 (0.31-0.58)**	**0.43 (0.33-0.56)**	**0.48 (0.33-0.69)**	**0.73 (0.56-0.96)**	**0.29 (0.17-0.49)**	CAU		
0.28 (0.24-0.34)	**0.24 (0.17-0.32)**	**0.29 (0.21-0.38)**	**0.28 (0.21-0.38)**	**0.28 (0.21-0.39)**	**0.32 (0.21-0.47)**	**0.48 (0.36-0.65)**	**0.19 (0.11-0.32)**	**0.66 (0.54-0.81)**	WL	
0.53 (0.34-0.83)	**0.44 (0.26-0.74)**	**0.53 (0.34-0.85)**	**0.52 (0.30-0.89)**	**0.53 (0.32-0.88)**	0.59 (0.33-1.05)	0.91 (0.55-1.50)	**0.36 (0.18-0.71)**	1.24 (0.78-1.97)	**1.87 (1.17-3.00)**	PLA

Values are odds ratios (OR) with 95% confidence intervals. OR<1 means that the row-defining intervention is less efficacious than the column-defining intervention. Bold prints highlight significant differences. CBT – cognitive behavioural therapy, BAT – behavioural activation therapy, PST – problem-solving therapy, 3WV – "third wave" therapies, IPT – interpersonal psychotherapy, DYN – psychodynamic therapy, SUP – non-directive supportive counseling, LRT – life-review therapy, CAU – care-as-usual, WL – waiting list, PLA – pill placebo

Table 4 Network meta-analyses: acceptability of psychotherapies compared with each other and with control conditions

CBT										
1.07 (0.78-1.46)	BAT									
1.05 (0.79-1.40)	0.99 (0.66-1.47)	PST								
0.99 (0.69-1.42)	0.93 (0.59-1.46)	0.94 (0.61-1.46)	3WV							
0.92 (0.68-1.24)	0.86 (0.57-1.30)	0.87 (0.59-1.28)	0.93 (0.59-1.46)	IPT						
1.38 (0.99-1.92)	1.29 (0.84-1.99)	1.31 (0.85-2.00)	1.39 (0.86-2.25)	1.50 (0.98-2.29)	DYN					
1.04 (0.76-1.42)	0.98 (0.65-1.47)	0.99 (0.68-1.44)	1.05 (0.66-1.67)	1.13 (0.77-1.68)	0.76 (0.49-1.17)	SUP				
0.82 (0.49-1.39)	0.77 (0.43-1.40)	0.78 (0.45-1.36)	0.83 (0.45-1.53)	0.90 (0.50-1.61)	0.60 (0.33-1.10)	0.79 (0.44-1.41)	LRT			
0.89 (0.77-1.03)	0.83 (0.61-1.13)	0.84 (0.63-1.13)	0.89 (0.62-1.30)	0.97 (0.73-1.28)	**0.64 (0.46-0.90)**	0.85 (0.62-1.17)	1.08 (0.64-1.83)	CAU		
0.67 (0.56-0.80)	**0.63 (0.44-0.88)**	**0.63 (0.47-0.85)**	**0.67 (0.47-0.96)**	0.73 (0.52-1.02)	**0.49 (0.33-0.70)**	**0.64 (0.45-0.90)**	0.81 (0.49-1.35)	**0.75 (0.60-0.94)**	WL	
1.38 (0.84-2.27)	1.30 (0.73-2.29)	1.31 (0.75-2.30)	1.39 (0.76-2.56)	1.50 (0.86-2.62)	1.00 (0.57-1.76)	1.33 (0.75-2.36)	1.68 (0.82-3.43)	1.56 (0.94-2.59)	**2.07 (1.23-3.49)**	PLA

Values are odds ratios (OR) with 95% confidence intervals. OR<1 means that the row-defining intervention is more acceptable than the column-defining intervention. Bold prints highlight significant differences. CBT – cognitive behavioural therapy, BAT – behavioural activation therapy, PST – problem-solving therapy, 3WV – "third wave" therapies, IPT – interpersonal psychotherapy, DYN – psychodynamic therapy, SUP – non-directive supportive counseling, LRT – life-review therapy, CAU – care-as-usual, WL – waiting list, PLA – pill placebo

Table 5 Network meta-analyses: remission in psychotherapies compared with each other and with control conditions

CBT										
1.14 (0.79-1.63)	BAT									
1.10 (0.78-1.56)	0.97 (0.61-1.54)	PST								
1.03 (0.69-1.54)	0.90 (0.55-1.49)	0.93 (0.56-1.55)	3WV							
0.88 (0.63-1.23)	0.78 (0.49-1.24)	0.80 (0.51-1.27)	0.86 (0.52-1.44)	IPT						
0.74 (0.52-1.06)	0.65 (0.41-1.05)	0.67 (0.42-1.08)	0.72 (0.43-1.22)	0.84 (0.53-1.33)	DYN					
0.59 (0.42-0.83)	**0.52 (0.33-0.82)**	**0.54 (0.35-0.83)**	**0.58 (0.35-0.96)**	0.67 (0.44-1.02)	0.80 (0.52-1.23)	SUP				
0.71 (0.33-1.52)	0.63 (0.27-1.43)	0.65 (0.29-1.42)	0.69 (0.30-1.59)	0.81 (0.36-1.82)	0.96 (0.42-2.19)	1.20 (0.53-2.73)	LRT			
0.35 (0.29-0.43)	**0.31 (0.21-0.45)**	**0.32 (0.22-0.46)**	**0.34 (0.22-0.53)**	**0.40 (0.29-0.55)**	**0.47 (0.33-0.68)**	**0.60 (0.42-0.82)**	0.49 (0.23-1.07)	CAU		
0.25 (0.20-0.32)	**0.22 (0.15-0.33)**	**0.23 (0.16-0.33)**	**0.25 (0.16-0.37)**	**0.29 (0.15-0.42)**	**0.34 (0.22-0.52)**	**0.43 (0.29-0.63)**	**0.36 (0.17-0.74)**	**0.72 (0.54-0.96)**	WL	
0.58 (0.33-1.52)	**0.51 (0.27-0.99)**	**0.53 (0.30-0.93)**	0.57 (0.28-1.13)	0.66 (0.35-1.24)	0.78 (0.42-1.48)	0.98 (0.52-1.86)	0.82 (0.32-2.07)	1.65 (0.92-2.96)	**2.30 (1.26-4.19)**	PLA

Values are odds ratios (OR) with 95% confidence intervals. OR<1 means that the row-defining intervention is less efficacious than the column-defining intervention. Bold prints highlight significant differences. CBT – cognitive behavioural therapy, BAT – behavioural activation therapy, PST – problem-solving therapy, 3WV – "third wave" therapies, IPT – interpersonal psychotherapy, DYN – psychodynamic therapy, SUP – non-directive supportive counseling, LRT – life-review therapy, CAU – care-as-usual, WL – waiting list, PLA – pill placebo

The results of the SUCRA are shown in Table 7, separately for response, remission, SMD and acceptability. Life-review and behavioural activation therapy ranked highest for response and SMD; behavioural activation and problem-solving therapy ranked highest for remission; while non-directive supportive counseling and psychodynamic therapy ranked lowest for response, remission and SMD. Psychodynamic therapy ranked lowest for acceptability, while life-review and interpersonal psychotherapy ranked highest.

Sensitivity and meta-regression analyses

In the sensitivity analyses in which we only included studies with low risk of bias, we found outcomes comparable to the main analyses. Only the differences between non-directive supportive counseling and most other therapies were no longer significant, and non-directive supportive counseling was no longer significantly better than care-as-usual and waiting list. The other sensitivity analyses resulted in no materially different outcomes from the main analyses.

In meta-regression analyses, only five predictors were found to be statistically significant (diagnosed depressive disorder for CBT vs. interpersonal psychotherapy, and CBT vs. waiting list; number of sessions for CBT vs. behavioural activation therapy; Western vs. non-Western countries for CBT vs. care-as-usual; and risk of bias for CBT vs. behavioural activation therapy) (see supplementary information). Because of their correlational nature and the large number of analyses conducted, these findings should be interpreted with caution.

In the meta-regression analysis to assess the influences of small study effects, the overall results were comparable with the main analysis.

Long-term effects

We conducted an NMA with the 90 studies that reported outcomes for response at 12 (±6) months after randomization (see Table 7). The results indicated that CBT, behavioural activation therapy, problem-solving therapy, interpersonal psychotherapy, and psychodynamic therapy had significant effects compared with care-as-usual at follow-up. The same therapies, except behavioural activation therapy, had also significant effects compared to waiting list. Problem-solving therapy was significantly more effective than CBT, "third wave" therapies and non-directive supportive counseling at follow-up. Interpersonal psychotherapy was also significantly more effective than non-directive supportive counseling at follow-up.

Only nine studies reported outcomes at more than 18 months after randomization. Because of the small number of studies and different periods, we did not conduct any analyses with these studies.

Table 6 Network meta-analyses: standardized mean difference for psychotherapies compared with each other and with control conditions

CBT										
0.06 (–0.15 to 0.28)	BAT									
–0.03 (–0.27 to 0.20)	–0.09 (–0.39 to 0.20)	PST								
0.02 (–0.21 to 0.25)	–0.04 (–0.33 to 0.25)	0.05 (–0.26 to 0.37)	3WV							
–0.13 (–0.36 to 0.10)	–0.19 (–0.49 to 0.10)	–0.10 (–0.40 to 0.21)	–0.15 (–0.46 to 0.16)	IPT						
–0.17 (–0.47 to 0.13)	–0.23 (–0.69 to 0.13)	–0.13 (–0.50 to 0.24)	–0.19 (–0.56 to 0.18)	–0.04 (–0.40 to 0.33)	DYN					
–0.35 (–0.55 to –0.15)	**–0.41 (–0.69 to –0.13)**	**–0.32 (–0.60 to –0.03)**	**–0.37 (–0.67 to –0.07)**	–0.22 (–0.49 to 0.06)	–0.18 (–0.52 to 0.15)	SUP				
0.14 (–0.21 to 0.48)	0.08 (–0.32 to 0.47)	0.17 (–0.23 to 0.57)	0.12 (–0.28 to 0.52)	0.27 (–0.13 to 0.67)	0.30 (–0.14 to 0.75)	**0.49 (0.10 to 0.87)**	LRT			
–0.67 (–0.79 to –0.56)	**–0.73 (–0.95 to –0.52)**	**–0.64 (–0.88 to –0.40)**	**–0.69 (–0.93 to –0.45)**	**–0.54 (–0.76 to –0.32)**	**–0.50 (–0.81 to –0.20)**	**–0.32 (–0.53 to –0.11)**	**–0.81 (–1.15 to –0.46)**	CAU		
–0.97 (–1.09 to –0.84)	**–1.03 (–1.26 to –0.80)**	**–0.93 (–1.17 to –0.69)**	**–0.99 (–1.22 to –0.76)**	**–0.84 (–1.08 to –0.59)**	**–0.80 (–1.12 to –0.48)**	**–0.62 (–0.84 to –0.39)**	**–1.10 (–1.45 to –0.76)**	**0.29 (0.14 to 0.45)**	WL	
–0.64 (–1.16 to –0.11)	**–0.70 (–1.25 to –0.15)**	**–0.61 (–1.16 to –0.05)**	**–0.66 (–1.23 to –0.09)**	–0.51 (–1.06 to 0.05)	–0.47 (–1.07 to 0.13)	–0.29 (–0.85 to 0.27)	–0.78 (–1.40 to –0.15)	–0.03 (–0.56 to 0.50)	0.33 (–0.21 to 0.86)	PLA

Values are standardized mean differences (SMDs) with 95% confidence intervals. Negative values indicate that the row-defining intervention is less efficacious than the column-defining intervention. Bold prints highlight significant differences. CBT – cognitive behaviour therapy, BAT – behavioural activation therapy, PST – problem-solving therapy, 3WV – "third wave" therapies, IPT – interpersonal therapy, DYN – psychodynamic therapy, SUP – non-directive supportive counseling, LRT – life-review therapy, CAU – care-as-usual, WL – waiting list, PLA – pill placebo

Table 7 Ranking of psychotherapies and control conditions according to the "surface under the cumulative ranking" (SUCRA) for response, standardized mean difference (SMD), remission and acceptability

	Response	SMD	Remission	Acceptability
Cognitive behavioural therapy	64.0	72.8	75.1	48.4
Behavioural activation therapy	85.2	82.1	86.3	39.1
Problem-solving therapy	62.9	67.2	83.5	40.8
"Third wave" therapies	66.5	75.7	76.3	51.1
Interpersonal psychotherapy	64.6	52.0	62.3	62.1
Psychodynamic therapy	52.8	49.2	46.3	10.0
Non-directive supportive counseling	26.6	30.8	30.5	42.3
Life-review therapy	93.1	87.1	46.5	72.5
Care-as-usual	12.0	14.5	10.7	71.8
Waiting list	0.0	1.10	0.2	97.2
Pill placebo	22.3	17.4	32.2	14.6

DISCUSSION

In this NMA, we compared the effects of the eight most common types of psychotherapy for depression with each other and with major control conditions in 331 controlled trials. We found that all therapies had significant effects compared to care-as-usual and waiting list control condition. The effects of the therapies did not differ significantly from each other, except for non-directive supportive counseling, that was less effective than all the other types of therapy. These results were broadly confirmed in a series of sensitivity analyses.

These findings are in line with previous meta-analytic research on psychotherapies for depression[7,10]. However, in contrast to previous meta-analyses, we could include a considerable number of studies with low risk of bias, which broadly confirmed the main results of this NMA.

Non-directive supportive counseling was less effective than the other therapies, but these findings were no longer significant when we only included studies with low risk of bias. This is in line with previous meta-analytic work[26]. However, these findings may be related to the fact that, in many studies, counseling was used as a control condition, and therapists may not have delivered optimal treatments.

Life-review therapy was not included in previous meta-analyses, because the number of studies was too small. This psychotherapy is mostly used in older adults, but it has also been used successfully in cancer patients[27,28], and it could very well be used in other populations without general medical disorders. Because of the small number of studies and the low quality of most of them, more research is clearly needed. However, life-review therapy can be considered a promising intervention that is probably efficacious in depression.

Overall, the findings of this NMA suggest that all psychotherapies that were examined, except non-directive supportive coun-

Table 8 Long-term response to psychotherapies compared with each other and control conditions

CBT									
0.97 (0.62-1.52)	BAT								
1.69 (1.08-2.66)	1.75 (0.97-3.14)	PST							
0.77 (0.46-1.30)	0.80 (0.43-1.49)	**0.46 (0.23-0.90)**	3WV						
1.35 (0.92-1.99)	1.40 (0.78-2.49)	0.80 (0.45-1.41)	1.75 (0.93-3.31)	IPT					
1.02 (0.63-1.66)	1.05 (0.55-2.02)	0.60 (0.32-1.14)	1.32 (0.65-2.67)	0.75 (0.41-1.38)	DYN				
0.78 (0.56-1.09)	0.81 (0.46-1.40)	**0.46 (0.27-0.79)**	1.01 (0.55-1.86)	**0.58 (0.36-0.94)**	0.76 (0.44-1.33)	SUP			
0.90 (0.19-4.33)	0.93 (0.18-4.75)	0.53 (0.10-2.71)	1.16 (0.22-6.05)	0.67 (0.13-3.33)	0.88 (0.17-4.52)	1.15 (0.24-5.48)	LRT		
0.59 (0.50-0.70)	**0.61 (0.39-0.96)**	**0.35 (0.23-0.53)**	0.76 (0.45-1.29)	**0.43 (0.30-0.63)**	**0.58 (0.36-0.93)**	0.75 (0.54-1.06)	0.65 (0.14-3.15)	CAU	
0.49 (0.29-0.83)	0.51 (0.26-1.01)	**0.29 (0.15-0.58)**	0.63 (0.31-1.29)	**0.36 (0.19-0.69)**	**0.48 (0.24-0.98)**	0.63 (0.34-1.16)	0.55 (0.11-2.69)	0.84 (0.48-1.44)	WL

Values are odds ratios (OR) with 95% confidence intervals. OR<1 means that the row-defining intervention is less efficacious than the column-defining intervention. Bold prints highlight significant differences. CBT – cognitive behavioural therapy, BAT – behavioural activation therapy, PST – problem-solving therapy, 3WV – "third wave" therapies, IPT – interpersonal therapy, DYN – psychodynamic therapy, SUP – non-directive supportive counseling, LRT – life-review therapy, CAU – care-as-usual, WL – waiting list, PLA – pill placebo

seling, are efficacious and can be used in routine care. The fact that all psychotherapies can be efficacious means that, when choosing a therapy, patient's preferences can have a prominent role. Mental health professionals need to facilitate access to evidence-based updated information about the effects of treatment interventions and to involve patients more in their day-to-day care, with a focus on carefully acknowledging the risk and outlining potential effects while managing expectations[29]. It is possible that a more detailed characterization of each patient with a diagnosis of depression may lead to a more precise matching between individual patients and individual psychotherapies[30].

One important finding of this study is that several psychotherapies still have significant effects at one-year follow-up, including CBT, behavioural activation therapy, problem-solving therapy, interpersonal psychotherapy, and psychodynamic therapy. We also found that problem-solving therapy may be somewhat more efficacious than some other therapies at follow-up, although this should be considered with caution, because of the relatively small number of studies and the considerable risk of bias in most studies. It is important for clinicians and patients that therapies work considerably longer than the therapy lasts.

In a recent NMA published in this journal[31], combined psychotherapy and pharmacotherapy was more effective than either of them alone in achieving response, also in chronic and treatment-resistant depression. Combined treatment and psychotherapy alone were also more acceptable than pharmacotherapy. Combined treatments seem therefore to be the best choice for patients with moderate to severe depression.

This study has several important strengths, but also some limitations. One strength is the large number of trials (N=331) that could be included. This is the largest NMA ever conducted in psychotherapies for depression. Although most studies were focused on CBT, care-as-usual and waiting list, we have sufficient studies comparing most other therapies and control conditions with each other. One important limitation is that the proportion of studies with low risk of bias was still relatively small (30.8%), although this was enough to conduct sensitivity analyses. Another important limitation is that we found some discrepancies between direct and indirect evidence, and only after excluding outliers the direct and indirect evidence pointed in the same direction. A final limitation is that only a relatively small number of trials reported longer-term outcomes, which makes these effects uncertain.

Despite these limitations, we can conclude that the most important types of psychotherapy, including CBT, behavioural activation therapy, problem-solving therapy, "third wave" therapies, interpersonal psychotherapy, psychodynamic therapy and life-review therapy, can be effective and acceptable in the treatment of adult depression, with no significant differences between them.

ACKNOWLEDGEMENTS

Supplementary information on the study is available at https://osf.io/7rmgj/.

REFERENCES

1. Steel Z, Marnane C, Iranpour C et al. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. Int J Epidemiol 2014;43:476-93.
2. Bloom DE, Cafiero E, Jané-Llopis E et al. The global economic burden of noncommunicable diseases. Geneva: World Economic Forum, 2011.
3. Hu TW. Perspectives: an international review of the national cost estimates of mental illness, 1990-2003. J Ment Health Policy Econ 2006;9:3-13.
4. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017;390:1211-59.
5. 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.
6. 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.
7. Barth J, Munder T, Gerger H et al. Comparative efficacy of seven psychotherapeutic interventions for depressed patients: a network meta-analysis. PLoS Med 2013;10:e1001454.
8. Cuijpers P, Berking M, Andersson G et al. A meta-analysis of cognitive behavior therapy for adult depression, alone and in comparison to other treatments. Can J Psychiatry 2013;58:376-85.
9. Cuijpers P, Karyotaki E, de Wit L et al. The effects of fifteen evidence-supported therapies for adult depression: a meta-analytic review. Psychother Res 2020;30:279-93.
10. Cuijpers P, van Straten A, Andersson G et al. Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. J Consult Clin Psychol 2008;76:909-22.
11. Cuijpers P. Are all psychotherapies equally effective in the treatment of adult depression? The lack of statistical power of comparative outcome studies. Evid Based Ment Health 2016;19:39-42.
12. Cuijpers P, Karyotaki E, Ciharova M. A meta-analytic database of randomised trials on psychotherapies for depression. www.osf.io/825c6.
13. Cuijpers P, Clignet F, van Meijel B et al. Psychological treatment of depression in inpatients: a systematic review and meta-analysis. Clin Psychol Rev 2011;31:353-60.
14. Cuijpers P, Noma H, Karyotaki E et al. Individual, group, telephone, self-help and internet-based cognitive behavior therapy for adult depression; a network meta-analysis of delivery methods. JAMA Psychiatry 2019;76:700-7.
15. Higgins JPT, Altman DG, Gøtzsche PC et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.
16. Furukawa TA, Cipriani A, Barbui C et al. Imputing response rates from means and standard deviations in meta-analyses. Psychopharmacology 2005;20:49-52.
17. Orsini N, Bottai M, Higgins J et al. Heterogi: Stata module to quantify heterogeneity in a meta-analysis. Boston: Statistical Software Components, 2006.
18. Egger M, Smith GD, Schneider M et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629.
19. Hutton B, Wolfe D, Moher D et al. Reporting guidance considerations from a statistical perspective: overview of tools to enhance the rigour of reporting of randomised trials and systematic reviews. Evid Based Ment Health 2017;20:46-52.
20. Chaimani A, Higgins JP, Mavridis D et al. Graphical tools for network meta-analysis in STATA. PLoS One 2013;8:e76654.
21. Chaimani A, Salanti G, Leucht S et al. Common pitfalls and mistakes in the set-up, analysis and interpretation of results in network meta-analysis: what clinicians should look for in a published article. Evid Based Ment Health 2017;20:88-94.
22. Turner RM, Davey J, Clarke MJ et al. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. Int J Epidemiol 2012;41:818-27.
23. Rhodes KM, Turner RM, Higgins JP. Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. J Clin Epidemiol 2015;68:52-60.
24. Nikolakopoulou A, Higgins JPT, Papakonstantinou T et al. CINeMA: an approach for assessing confidence in the results of a network meta-analysis. PLoS Med 2020;17:1-19.

25. Noma H, Gosho M, Ishii R et al. Outlier detection and influence diagnostics in network meta-analysis. Res Synth Meth 2020;11:891-902.
26. Cuijpers P, Driessen E, Hollon SD et al. The efficacy of non-directive supportive therapy for adult depression: a meta-analysis. Clin Psychol Rev 2012;32:280-91.
27. Kleijn G, Lissenberg-Witte BI, Bohlmeijer ET et al. The efficacy of Life Review Therapy combined with Memory Specificity Training (LRT-MST) targeting cancer patients in palliative care: a randomized controlled trial. PLoS One 2018;13:e0197277.
28. Zhang X, Xiao H, Chen Y. Effects of life review on mental health and well-being among cancer patients: a systematic review. Int J Nurs Stud 2017;74:138-48.
29. Tomlinson A, Boaden K, Cipriani A. Withdrawal, dependence and adverse events of antidepressants: lessons from patients and data. Evid Based Ment Health 2019;22:137-8.
30. Maj M, Stein DJ, Parker G et al. The clinical characterization of the adult patients with depression aimed at personalization of management. World Psychiatry 2020;19:269-93.
31. Cuijpers P, Noma H, Karyotaki E et al. A network meta-analysis of the effects of psychotherapies, pharmacotherapies and their combination in the treatment of adult depression. World Psychiatry 2020;19:92-107.

DOI:10.1002/wps.20860

Explaining the missing heritability of psychiatric disorders

Evidence from family, twin and adoption studies indicates that psychiatric disorders are substantially heritable. Heritability is usually expressed as the proportion of trait variance attributable to additive genetic factors (narrow sense heritability: h^2). The h^2 estimates for schizophrenia, attention-deficit/hyperactivity disorder, autism spectrum disorder and bipolar disorder are all >0.66, and are substantial for a range of other psychiatric conditions[1].

This evidence has motivated the application of increasingly sophisticated genomic approaches, including genome-wide association studies (GWAS) and next generation sequencing, that have identified a large number of genetic risk factors across a range of psychiatric conditions[2]. These studies revealed that psychiatric disorders are highly polygenic, with the major component of the heritability captured so far coming from common alleles (population frequency >0.01) detected in GWAS.

While this is extremely encouraging, and has set up an empirical platform upon which future progress towards precision psychiatry can be built[2], estimates of h^2 accounted for by the genetic variants identified in GWAS have always been substantially lower than the estimates of h^2 from family, twin and adoption studies. This shortfall is not a peculiarity of psychiatric disorders; it is also seen in many polygenic diseases and traits, and has been termed the "missing heritability".

Three main explanations for this missing heritability have been proposed[3,4]. First, it is possible that the estimates of h^2 from family, twin and adoption studies were inflated due to confounding factors such as shared environment. Second, estimates of h^2 from genomic studies may be deflated as they do not account for non-additive genetic effects such as dominance and gene-gene interactions. Finally, it may be the case that many risk alleles have simply not been identified by GWAS, either because their effects are too small or because they are too uncommon.

While all of these hypotheses remain plausible, the last one has received support from recent studies of polygenic traits and diseases, suggesting that many causal variants remain unidentified. In order to understand this, a brief explanation of GWAS is required. These studies involve genotyping single nucleotide polymorphisms (SNPs) that are common in the population (typically 500,000 - 1 million SNPs with a population frequency >5%). Because common SNPs tend to be correlated with their neighbours – a phenomenon known as linkage disequilibrium (LD) – the genotypes of additional SNPs can be inferred through a statistical process known as "imputation". This greatly increases the number of SNPs available to GWAS (typically >10 million SNPs with a population frequency >1%). When researchers seek associations in GWAS, they need to correct for the large number of statistical tests by taking a stringent threshold for statistical significance (known as genome-wide significance). This greatly reduces the occurrence of false positives, but at the expense of causing many real associations to be missed.

Early studies that revealed the missing heritability focused only on SNPs that met genome-wide significance. Subsequent studies have shown that more accurate and larger estimates of h^2 can be obtained by considering all available SNPs together, including imputed as well as directly genotyped SNPs, and by using data from reference samples that have undergone whole-genome sequencing (WGS) to allow better imputation of rare variants.

When these approaches are implemented, the proportion of h^2 that is captured increases to around one- to two-thirds of that expected in polygenic traits and diseases[4], with h^2 estimates for schizophrenia, bipolar disorder and autism being 0.23, 0.25 and 0.17, respectively[5]. This indicates that a proportion of the missing heritability was carried by SNPs that currently lie below the genome-wide significance threshold and also those that were insufficiently correlated with common SNPs to allow accurate imputation. It is, therefore, anticipated that the increased power of GWAS obtained from a substantial increase in both the number of common SNPs and the sample size will result in many more risk variants of small effect meeting genome-wide significance, as well as improving estimates of heritability[4].

However, the ability of common SNPs used in GWAS to capture the effects of variants with which they are in low LD is limited. The application of exome sequencing and WGS to complex disease cohorts has confirmed the presence in the human genome of a large number of rare genetic variants (defined as having a population frequency <1%). Importantly, these are not well correlated through LD with common SNPs and are therefore not accurately imputed in GWAS.

Recent work applying WGS to a large population cohort[6] has shown that estimates of heritability made using rare as well as common variants are much closer to those predicted from family studies for both height and body mass index, with much of the increase coming from SNPs that could not be accurately imputed from GWAS.

It is well recognized that, when compared to height and body mass index, many psychiatric disorders are under greater negative selection, and this is expected to result in a greater contribution from rare risk alleles. It is, therefore, plausible that rare genetic variants could be particularly relevant to psychiatric disorders, meaning that future WGS studies in large samples could prove to be particularly fruitful.

The prospect of large scale WGS studies in psychiatry is certainly exciting and will likely reveal much about genetic architecture and biology, as well as delivering better predictive tools. Short-read sequencing (SRS), based on compiling reads from <150bp segments, is currently the most widely used approach to WGS, because of its low cost and high throughput. It is particularly powerful in identifying rare single nucleotide variants and small insertion/deletions[7]. Robust approaches have been recently introduced to detect structural variants such as duplications, deletions, inversions, and other changes involving larger DNA segments (generally greater than 50-100 bases long) that are likely to be relevant to psychiatric disorders[8].

While SRS will undoubtedly be increasingly and fruitfully ap-

plied in psychiatric genomics in the coming years, it has limitations imposed by the fact that it works by stitching together short reads *in silico*. This means that there are regions of the genome which are difficult or impossible to read, such as those containing large structural variants, repetitive sequences, extreme guanine-cytosine content, or sequences with multiple homologous elements within the genome. This is sometimes known as the "dark genome".

There are now a number of long-read sequencing (LRS) platforms that allow the analysis of segments of the human genome up to 200kb, and these are capable of shining a light into the dark genome. Emerging studies using LRS are identifying larger, more harmful structural variants and long repetitive elements[7,9], both of which are candidates for involvement in psychiatric disorders.

Psychiatric genomics is a work in progress. GWAS have been hugely successful in identifying the role of multiple common variants, but recent work on missing heritability suggests a need to focus now on rare variants, and in the next few years we can expect studies based upon both SRS and LRS technologies to do this.

Fully characterizing the genetic architecture of psychiatric disorders is likely to improve polygenic risk prediction for both screening and stratification, allow a better understanding of the underlying biological mechanisms of disease, and broaden the landscape of pharmaceutical targets[2].

Michael J. Owen, Nigel M. Williams
MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK

1. Sullivan PF, Daly MJ, O'Donovan M. Nat Rev Genet 2012;13:537-51.
2. Rees E, Owen MJ. Genome Med 2020;12:43.
3. Manolio TA, Collins FS, Cox NJ et al. Nature 2009;461:747-53.
4. Visscher PM, Wray NR, Zhang Q et al. Am J Hum Genet 2017;101:5-22.
5. Cross-Disorder Group of the Psychiatric Genomics Consortium. Nat Genet 2013;45:984-94.
6. Wainschtein P, Jain D, Yengo L et al. bioRxiv 2019;588020.
7. Chaisson MJP, Sanders AD, Zhao X et al. Nat Commun 2019;10:1784.
8. Collins R, Brand H, Karczewski K et al. bioRxiv 2019;578674.
9. Beyter D, Ingimundardottir H, Eggertsson H et al. Nature 2020;581:444-51.

DOI:10.1002/wps.20870

Toward a systems-based approach to understanding the role of the sympathetic nervous system in depression

The sympathetic nervous system (SNS) has an essential role in the prototypical stress response. Stress, stressors, and stress responses are central themes in most prominent theories of depression etiology and maintenance. Yet, the SNS is not a commonly targeted mechanism in depression research. Here we propose a dynamic, systems-level approach that contextualizes SNS-mediated stress responsivity within a regulatory framework. We believe that this conceptualization hews closer to the role of the SNS as a time-varying, context-driven regulatory system, and provides clinicians and researchers with a model for understanding its relevance to depression.

Interest in the SNS in depression is not new. A host of methods and markers have been used to try to delineate the role of the SNS in depression, including cardiac measures such as heart rate and pre-ejection period, skin conductance, salivary alpha-amylase, and urinary and serum measures of catecholamines. However, evidence for tonically-elevated SNS arousal in depression has been inconsistent and equivocal[1].

We propose three reasons for this equivocality. First, because the SNS is embedded in a larger set of regulatory systems, analysis of absolute levels should be augmented with – if not eschewed altogether for – a systems perspective that incorporates dynamic interrelations between system components. Second, the temporal dynamics of the stress response have been well documented[2], with SNS effects occurring relatively rapidly and ephemerally (compared to those of glucocorticoids), and attempts should be made to capture these time-dependent fluctuations. Third, there are likely to be individual differences in the dynamics and calibration of cognitive, affective and physiological regulatory systems. Thus, attempts should be made to identify subgroups.

Cognitive theories of depression have long posited the importance of depressogenic schemas – internal working models of the self, others, and the world – that magnify and distort the perception of ambiguous stimuli[3]. The presence of these schemas can increase the likelihood of threat appraisals (e.g., perceptions of external stressors) and the elicitation of negative emotional responses. The aversive arousal from negative emotions has been proposed to amplify memory for negative events[2] and provide experiential feedback that supports and reinforces the initial threat appraisal[3]. Thus, individuals with depression may be more likely to perceive environmental stressors, which elicit negative emotional reactions that reinforce the threatening nature of the stimulus and enhance memory encoding of the experience.

Inherent to this positive feedback loop between perceptions, appraisals and arousal is the physiological stress response to perceived stressors. This response serves an adaptive function to mobilize energy, stimulate immune activation, and increase cardiovascular tone through vasoconstriction and increases in heart rate and contractility. The stress response is composed of coordinated actions of the hypothalamic-pituitary-adrenal (HPA) axis, the SNS and the parasympathetic nervous system (PNS).

Compared to the SNS, there has been an abundant amount of research on the HPA axis and the PNS in depression, and studies have found evidence for HPA dysfunction[4] and reduced heart rate variability[1,5] in depressed patients. However, contradictory and null findings have also been common. We raise the possibility that inconsistent findings may stem from the isolation of system components in lieu of the whole. For instance, the HPA axis

can play permissive or suppressive roles in determining the magnitude of SNS stress reactivity[2]. Because HPA axis functionality can precede and inform the nature of the sympathetic stress response, incorporating preceding levels of available cortisol *in situ* during moments of emotional strain may help to better calibrate measurements of SNS reactivity in dysphoric individuals.

Moreover, it has been demonstrated that the doctrine of reciprocal antagonism between the PNS and SNS – the notion that more of one inherently means less of the other – does not universally hold[6]. This conclusion dictates that the two systems can exhibit concurrent and interactive behavior and should be measured and modeled as separate and distinct dimensions. A systems approach to stress responsiveness may help to better define and measure the individual components.

However, methodological challenges remain regarding the measurement and timing of different system components. The cascade of HPA axis hormonal actions has been well documented, with peak effects following roughly 20 min after stress exposure[2]. Meanwhile, PNS effects such as vagal withdrawal can operate on a scale of milliseconds to seconds, and SNS effects typically take place on a scale of seconds to minutes. We propose that research into autonomic functioning in human subjects should be pursued via time series analysis of electrophysiological measurements.

Common inputs such as respiratory sinus arrhythmia, pre-ejection period, and heart rate can be binned in epochs as small as 30 sec. Thus, data collection periods as short as an hour can produce time series of 120 observations. Time series analyses such as vector-autoregression and network analysis can model the relationships between system components. Moreover, ambulatory technologies exist allowing researchers to capture autonomic functioning in emotionally salient scenarios during an individual's day-to-day life.

Measures such as pre-ejection period require academic-research-grade equipment and expertly placed electrodes. However, innovations in mobile assaying of salivary measures could yield ambulatory measurements of SNS markers such as salivary alpha-amylase. Additionally, one could foresee the development of a fingerstick system for capillary blood measurement of catecholamines akin to blood glucose monitoring systems. It has been shown that reliable measurements of catecholamines can be derived from as little as 100 μL of capillary blood[7].

Understanding the role of the SNS in the treatment of depression may be important for patients' psychological and physical health. As noted above, depressed individuals have been shown to have significantly decreased parasympathetic cardiac regulation, and depression has long been associated with an increased incidence of coronary heart disease[1,5]. Although the evidence for sympathetic predominance in depressed patients has been equivocal, there is some evidence that antidepressant medications may affect this predominance[8]. Of note, one study has found that cognitive behavioral treatment may increase heart rate variability[9]. Clearly, more work is needed to understand the effects of psychotherapy and antidepressant medications on the SNS and sympathetic cardiac control.

Finally, we noted at the outset the likely existence of heterogeneous subpopulations of depressed individuals, some of whom may experience elevations in sympathetic arousal and some may not. It follows that individuals could exhibit more complex individual differences in the calibration of stress responsivity among cognitive, affective and physiological system components.

From this perspective, the SNS could play a primary role in driving phenomenological and physiological consequences for some individuals. For instance, during the transduction of cognitive-emotional stimuli into physiological responses, an adrenergic gain factor might serve to amplify moderate signals into more robust responses. In a time series context, competing directional models of SNS arousal, subjective affect, and cognitive appraisal could be tested. Moreover, such evaluations could be carried out on a person-by-person basis.

The SNS plays a calibrating role, exerting effects in response to shifting external demands and emotional conditions. It may be more fruitful to examine these dynamic, time-varying relationships with other stress-response systems, rather than mean group differences.

Aaron J. Fisher, Jiyoung Song, Peter D. Soyster
University of California, Berkeley, CA, USA

1. Carney RM, Freedland KE, Veith RC. Psychosom Med 2005;67:S29-33.
2. Sapolsky RM, Romero LM, Munck AU. Endocrine Rev 2000;21:55-89.
3. Disner SG, Beevers CG, Haigh EA et al. Nat Rev Neurosci 2011;12:467-77.
4. Keller J, Gomez R, Williams G et al. Mol Psychiatry 2017;22:527-36.
5. Kemp AH, Quintana DS, Gray MA et al. Biol Psychiatry 2010;67:1067-74.
6. Berntson GG, Cacioppo JT, Quigley KS. Psychol Rev 1991;98:459.
7. Baumgartner H, Ritsch R, Luz O et al. Pediatr Res 1992;31:579-82.
8. Koschke M, Boettger MK, Schulz S et al. Psychosom Med 2009;71:852-60.
9. Carney RM, Freedland KE, Stein PK et al. Psychosom Med 2000;62:639-47.

DOI:10.1002/wps.20872

Cardiac vagal tone: a neurophysiological mechanism that evolved in mammals to dampen threat reactions and promote sociality

The evolutionary journey from asocial reptiles to social mammals is highlighted by a reorganized autonomic nervous system with unique structural and functional changes in the vagus. These changes enable mammals to suppress defensive strategies in order to support and express sociality. The product of this transition is an autonomic nervous system with capacities to self-calm, to socially engage others, and to mitigate threat reactions in ourselves and others through social cues.

For mammals, whose survival is dependent on their ability to cooperate, to connect, and to co-regulate, the ancient defense programs dependent on sympathetic activation supporting fight/flight behaviors, and vagal activation supporting death feigning, had to be harnessed and repurposed. This process resulted in a reorganized brainstem area, the ventral vagal complex, from which a unique branch of the vagus nerve enabled the expression of several uniquely mammalian features, including the ability to calm and to signal safety. Thus, sociality became embedded within specific neurobiological processes that had capabilities to mitigate threat and support mental and physical health. When this "calming" system is disrupted, prominent markers of chronic stress and core features shared by several psychiatric conditions are expressed (e.g., flat facial affect, poor vocal prosody, hypervigilance, hyperreactivity, auditory and visual hypersensitivities).

Anatomically, this vagal pathway is myelinated and originates in the brainstem structure called nucleus ambiguus. It provides the primary vagal regulation of organs above the diaphragm. This is distinct from the vagal pathways originating in the dorsal vagus nucleus, which are unmyelinated and provide the primary vagal regulation of organs below the diaphragm. The ventral vagal complex also regulates the striated muscles of the face and head and is greatly influenced by afferent pathways traveling through the vagus, trigeminal and facial nerves. Thus, in mammals, the brainstem areas regulating the heart and bronchi are interconnected with the areas regulating ingestion, facial expression, listening, breathing and vocalizations, to form an integrated social engagement system. In fact, intonations of vocalizations are mediated by the vagus, enabling prosodic features of voice to convey a relatively accurate index of vagal regulation of the heart[1].

Following the work of Jackson[2], the polyvagal theory[3] assumes a phylogenetic hierarchy in which the newer circuits inhibit the older. Thus, when the ventral vagus and the social engagement system are dampened or go offline, which frequently is observed during chronic stress and in response to threat, the autonomic nervous system moves into a sympathetic state that supports mobilization (e.g., fight/flight). If this functional shift in state does not lead to a positive outcome, then the autonomic nervous system may abruptly shut down via the dorsal vagal circuit (e.g., syncope, death feigning).

Jackson described this process of sequentially disinhibiting older structures as *dissolution* or evolution in reverse. He used dissolution to explain the consequence of brain damage and disease, while polyvagal theory applies the principle to adaptive autonomic reactions to cues of threat, which may be reversible by cues of safety. In the realm of mental health, loss of access to the ventral vagus may be a product of chronic threat or a measurable core feature of several psychiatric disorders (e.g., post-traumatic stress disorder, PTSD), developmental disabilities (e.g., autism, Prader Willi syndrome), and disabling chronic pain.

To survive, mammalian offspring must initially nurse as the primary mode of ingesting food. To nurse the infant must suck, a process dependent on a brainstem circuit involving the ventral vagal complex. Survival is dependent on the infant's nervous system efficiently and effectively coordinating suck-swallow-breathe-vocalize behaviors with vagal regulation of the heart through the ventral vagal pathways originating in the nucleus ambiguus. Through maturation and socialization, this "ingestive" circuit provides the structural neural platform (i.e., social engagement system) for sociality and co-regulation as major mediators to optimize homeostatic function leading to health, growth and restoration.

In mammals, there is a dependency between reactions to contextual cues and the function of this circuit. Cues of threat may disrupt, while cues of safety may enhance function. The sensory branches of the facial and trigeminal nerves provide major input into the ventral vagal complex. Functionally, changes in the state of this circuit, through the process of dissolution, will either "disinhibit" phylogenetically older autonomic circuits to support defense (e.g., predator, disease, physical injury) or inform all aspects of the autonomic nervous system, including the enteric system[4], to optimize homeostatic function.

Polyvagal theory introduces "neuroception", a neural process that evaluates risk and safety and reflexively triggers shifts in autonomic state without requiring conscious awareness. This reflexive process, distinct from perception, detects environmental and visceral features that are safe, dangerous or life-threatening[5]. Although many vertebrates have a capacity to detect pain and threat, mammals repurposed the neuroception capacity of their reptilian ancestors to not only react instantaneously to threat, but also to calm instantaneously to cues of safety.

It is this latter feature that enables mammals to downregulate defensive strategies to promote sociality by enabling psychological and physical proximity without the consequences of injury. It is this calming mechanism that adaptively adjusts the central regulation of autonomic function to dampen sympathetic activation and to protect the oxygen-dependent central nervous system, especially the cortex, from the metabolically conservative defensive reactions of the dorsal vagal complex (e.g., syncope, diarrhea).

This potential to calm autonomic state via the social engagement system is compromised in many psychiatric conditions, and leads to a variety of autonomic dependent comorbidities, including irritable bowel syndrome, migraine and fibromyalgia. However, being a common feature of several disorders limits the potential utility of measures of ventral vagal function in differential diagnoses, although it would highlight the potential of recruiting the ventral vagal pathway as a portal for treatment via technologies (e.g., vagal nerve stimulation).

Our research documents that the quantification of the respiratory-related component of heart rate variability, known as respiratory sinus arrhythmia, provides a sensitive metric of the ventral vagus function (i.e., cardiac vagal tone)[6]. Applications of our method confirmed that respiratory sinus arrhythmia was even more sensitive than the assumed "gold standard" of cardiac vagal tone (i.e., changes in heart rate in response to vagal blockade).

Respiratory sinus arrhythmia is a physiological phenomenon with an identifiable underlying neural mechanism reflecting ventral vagal control of the heart. With an accurate measure of ventral vagal function, there is the possibility to monitor autonomic adjustments to threat and safety. From a clinical perspective, the ability to monitor dampened vagal regulation would provide insight

into understanding the mechanisms underlying clinical features. For example, chronic stress, clinical depression, or a life-threatening traumatic experience that may lead to PTSD could profoundly dampen ventral vagal regulation of the heart and the structures regulated by ventral vagal complex constituting the social engagement system[7,8].

Disrupting the brainstem locus of social engagement system would functionally impair social communication and co-regulation by reducing vocal prosody and facial affect, and, through the loss of neural tone to the middle ear muscles, influence auditory processing by inducing hypersensitivity to low frequency background sounds and hyposensitivity to voice. In concert with these changes, brainstem communication with higher brain structures would impair cognitive function and affect regulation, while supporting the defense strategies of fight or flight or shutdown (e.g., syncope, dissociation).

Monitoring ventral vagal function may provide an objective neurophysiological marker of clinical improvement[9].

Stephen W. Porges
Traumatic Stress Research Consortium at the Kinsey Institute, Indiana University, Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

1. Porges SW, Lewis GF. Handb Behav Neurosci 2010;19:255-64.
2. Jackson JH. BMJ 1884;1:703-7.
3. Porges SW. Psychophysiology 1995;32:301-18.
4. Kolacz J, Porges SW. Front Med 2018;5:145.
5. Porges SW. Ann NY Acad Sci 2003;1008:31-47.
6. Lewis GF, Furman SA, McCool MF et al. Biol Psychol 2012;89:349-64.
7. Hage B, Britton B, Daniels D et al. World J Biol Psychiatry 2019;20:359-67.
8. Fanning J, Silfer JL, Liu H et al. J Behav Med 2020;43:308-17.
9. Porges SW, Macellaio M, Stanfill SD et al. Int J Psychophysiol 2013;88:261-70.

DOI:10.1002/wps.20871

Psychiatric comorbidity in immune-mediated inflammatory diseases

Chronic immune-mediated inflammatory diseases (IMIDs) are a group of conditions characterized by immune dysregulation and aberrant organ system inflammation. Common examples of these conditions include rheumatoid arthritis, inflammatory bowel disease (including Crohn's disease and ulcerative colitis) and multiple sclerosis. Although these conditions affect different organ systems, they are all characterized by recurrent relapses and potentially debilitating disease progression.

Collectively, IMIDs affect more than 1 in 20 people worldwide, and substantially burden affected persons, their families and societies. The adverse impacts of IMIDs include symptoms such as pain and fatigue, impairments in relationships and social participation, loss of employment, increased health care utilization, and reduced life expectancy. Comorbid conditions are common in people with IMIDs and also contribute substantially to their burden.

Comorbid psychiatric disorders, including depression, anxiety disorders and bipolar disorder, are of particular interest. A growing body of evidence indicates that the incidence and prevalence of psychiatric disorders are elevated in persons with IMIDs as compared to the general population. For example, a population-based cohort of persons with rheumatoid arthritis, inflammatory bowel disease or multiple sclerosis had an elevated incidence of depression (incidence rate ratio, IRR=1.71; 95% CI: 1.64-1.79), anxiety (IRR=1.34; 95% CI: 1.29-1.40), bipolar disorder (IRR=1.68; 95% CI: 1.52-1.85) and schizophrenia (IRR=1.32; 95% CI: 1.03-1.69) compared to age-, sex- and geographically-matched controls[1].

The association between IMIDs and psychiatric disorders appears to be bidirectional, and the increased incidence of psychiatric disorders is not simply due to the challenges of living with a chronic disease. In a population-based study from Denmark involving 1,016,519 individuals, those with depression had a significantly higher risk of developing any IMID in the subsequent 11 years than individuals without depression[2]. In a population-based study from Canada, individuals with rheumatoid arthritis, inflammatory bowel disease or multiple sclerosis had an increased incidence of any psychiatric disorder, including depression, anxiety, bipolar disorder and schizophrenia, for 8-10 years prior to the diagnosis of their IMID, even after accounting for sociodemographic factors and number of physician visits[3].

Broadly, two health conditions may be comorbid (co-occur) for several reasons. Chance alone may account for comorbidity. Surveillance bias may also occur, wherein a person affected by one chronic health condition uses more health care services and consequently is more likely to get diagnosed with a second condition. Furthermore, conditions may co-occur due to "true etiologic mechanisms". These mechanisms may include common genetic or environmental factors, or direct causation of the second condition by the first one. Finally, both conditions could be caused by an unrecognized third condition.

Epidemiological and biological evidence suggests that IMIDs and psychiatric disorders are comorbid due to "true etiologic mechanisms". In a cohort of 5,727,655 individuals, incident depression was associated with an increased risk of incident Crohn's disease (hazard ratio, HR=2.11; 95% CI: 1.65-2.70) and ulcerative colitis (HR=2.23; 95% CI: 1.92-2.60) after adjusting for age, sex, socioeconomic status, comorbid conditions, smoking status and use of antidepressants[4]. Notably, treatment with antidepressants was protective of developing Crohn's disease or ulcerative colitis among individuals with depression.

The role of inflammation and immune dysregulation in IMIDs is well-recognized. Emerging evidence is highlighting the importance of immune dysfunction in psychiatric disorders as well, including depression, bipolar disorder, schizophrenia and anxiety disorders[5]. These latter disorders are associated with dysregulation of T cell function and pro-inflammatory cytokines, including interleukin-6 (IL-6), IL-2 receptor, IL-1β, IL-17A, and

C-reactive protein; altered microglial activation; and disruption of the blood-brain barrier[5,6]. Pharmacological and non-pharmacological therapies for depression are associated with reductions in peripheral inflammatory markers. Relatedly, the role of immunomodulatory therapies in the treatment of psychiatric disorders is also being explored. A randomized placebo-controlled trial in persons with major depression suggested that infliximab, a tumor necrosis factor antagonist, might improve depressive symptoms in persons who had elevated levels of C-reactive protein at enrollment[7].

With respect to common etiologic factors, several pleiotropic genetic loci are jointly associated with the risk of psychiatric disorders and IMIDs, as shown by an analysis of genome-wide association studies of five psychiatric disorders (major depressive disorder, bipolar disorder, schizophrenia, autism spectrum disorder and attention-deficit/hyperactivity disorder) and seven immune-mediated disorders (Crohn's disease, rheumatoid arthritis, inflammatory bowel disease, multiple sclerosis, type 1 diabetes, systemic lupus erythematosus and psoriasis)[8]. Notably, shared genetic loci related to immune function were prominent.

In addition to genetic factors, psychosocial factors also influence immune system function and inflammation. Acute stress leads to activation of the autonomic nervous system and hypothalamic-pituitary-adrenal axis, and upregulates inflammation. Chronic stress, such as childhood maltreatment, increases inflammation and suppresses cellular and humoral immunity. In turn, these changes increase the risk of chronic diseases such as IMIDs and psychiatric disorders. Social support can mitigate the adverse effects of psychosocial stressors on the risk of chronic diseases. Concordant with these observations, psychosocial interventions, in particular cognitive behavioral therapy and multimodality therapies, are associated with sustained improvements in immune system function as measured by pro-inflammatory cytokines and immune cell counts[9].

Thus, epidemiological data support bidirectional relationships between psychiatric disorders and IMIDs, and inflammation and immune dysregulation are common to these conditions. Increasingly, treatment approaches applied to IMIDs are being tested for psychiatric disorders. However, much remains to be understood about the interface between psychiatric disorders and IMIDs.

Ruth Ann Marrie[1,2], Charles N. Bernstein[1]
[1]Department of Internal Medicine, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, Canada; [2]Department of Community Health Sciences, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, Canada

1. Marrie RA, Walld R, Bolton JM et al. J Psychosom Res 2017;101:17-23.
2. Andersson NW, Gustafsson LN, Okkels N et al. Psychol Med 2015;45:3559-69.
3. Marrie RA, Walld R, Bolton JM et al. Epidemiol Psychiatr Sci 2019;28:333-42.
4. Frolkis AD, Vallerand IA, Shaheen A-A et al. Gut 2019;68:1606-12.
5. Pape K, Tamouza R, Leboyer M et al. Nat Rev Neurol 2019;15:317-28.
6. Vieira MMM, Ferreira TB, Pacheco PAF et al. J Neuroimmunol 2010;229:212-8.
7. Raison CL, Rutherford RE, Woolwine BJ et al. JAMA Psychiatry 2013;70:31-41.
8. Wang Q, Yang C, Gelernter J et al. Hum Genet 2015;134:1195-209.
9. Shields GS, Spahr CM, Slavich GM. JAMA Psychiatry 2020;77:1031-43.

DOI:10.1002/wps.20873

Patients with schizophrenia are under-vaccinated for COVID-19: a report from Israel

After the first report published in this journal[1], several other studies conducted in the US, France, Korea and Israel have confirmed that individuals with severe mental illness (SMI), especially those diagnosed with schizophrenia, are at increased risk for COVID-19-related severe morbidity and mortality[e.g., 2,3]. These reports have led to the call to prioritize these patients for early COVID-19 vaccination[4].

While prioritization is especially pivotal during periods of vaccine deficiency, there are several reasons to suspect that, when vaccinations become widely available, they will not be fully utilized in individuals with SMI. Studies indicate that these patients are less likely to receive available standard levels of care for most of their medical diseases[5], and overall receive less treatment for diseases they are more susceptible to suffer from[6]. Furthermore, rates of vaccination for diseases such as influenza, which is mostly available to the public, have been reported to be low among individuals with SMI[7].

Israel has been highly proactive in engaging citizens to follow its mass COVID-19 vaccination plan[8]. Vaccinations became available to all citizens above the age of 16 by the end of January 2021. In a recent study from this country, we found that individuals with schizophrenia were more likely to suffer from COVID-19 morbidity and mortality compared to age and gender matched controls[3]. To explore whether patients with this diagnosis are being vaccinated to the same extent as their matched controls, we revisited the cohort of patients and updated their medical registry with information regarding vaccination rates.

The original cohort included 25,539 patients with schizophrenia and their matched controls (overall N=51,078). Deceased cases were omitted from the analysis, thus leading to a total of 50,240 cases (25,120 cases of schizophrenia and their age and gender matched controls). The study utilized the databases of Clalit Health Services (CHS), the largest operating health care organization in Israel[9]. These databases are regularly updated with real-time information derived from patients' medical registries, and undergo routine validation procedures for medical and psychiatric diagnoses.

The diagnosis of schizophrenia in this study was made by a senior psychiatrist in the patient's medical registry or was listed on a psychiatric hospital's discharge letter. Matched control participants comprised individuals with no diagnosis of schizophrenia randomly sampled at a 1:1 ratio. The study was approved by the CHS institutional review board, where informed consent was waived due to the anonymous nature of data extraction.

For the purposes of the current analysis, vaccination was considered as implemented if the patient received at least one dose. Univariate logistic regressions were employed to assess the odds of being vaccinated, and odds ratios (ORs) and 95% confidence intervals (CIs) were reported. The dataset was stratified for age and gender groups. All statistical analyses were performed using SPSS software, version 25.

The odds of receiving COVID-19 vaccination were significantly lower in the schizophrenia group compared to the control group (OR=0.80, 95% CI: 0.77-0.83, p<0.0001). No significant differences were observed in the 16-21 age subsample. Differences between the two groups were more profound as age increased: OR=0.90, 95% CI: 0.83-0.97, p<0.0001 in the 21-40 age subsample; OR=0.83, 95% CI: 0.79-0.88, p<0.0001 in the 40-60 age subsample; and OR=0.61, 95% CI: 0.57-0.64, p<0.0001 in the subsample at age 60 and above. The odds of being vaccinated were lower in the schizophrenia group for both male and female participants, with males showing slightly greater gaps in vaccination rates (OR=0.79, 95% CI: 0.75-0.82, p<0.0001) than females (OR=0.82, 95% CI: 0.77-0.87, p<0.0001).

These results indicate that individuals with schizophrenia, although well known by the scientific community for their medical and social vulnerabilities, are being under-vaccinated for COVID-19 in Israel compared to the rest of the population. This inequality is especially pronounced in people aged 60 and above, where the convergence of risk factors may create an additional accumulating mortality risk.

The lack of significant differences in the 16-21 age subsample may be related to the overall low rates of vaccination in young people. On the other hand, the increasing gap between the schizophrenia and control groups as age increases indicates that, when vaccination is more available (as older age groups could be vaccinated immediately upon the launch of the national plan), schizophrenia patients are more profoundly disadvantaged.

A variety of factors previously described as barriers to immunization in SMI people, such as lack of awareness and knowledge, fear, and lack of active recommendation from primary caregivers[7], may also serve as barriers to COVID-19 vaccination. Proactive efforts should be made to provide SMI people with easier access to vaccination as part of routine medical care policy. Such access can be obtained by, for example, providing *ad-hoc* vaccination to patients presenting for psychiatric examinations or follow-ups, who are interested in being vaccinated. Patients should also be actively monitored for completing the vaccination plan so as to make sure that they follow through on the recommendations made by the vaccine producers.

The results of this study are based on analyses of associations; therefore, no causality can be inferred from the study design. Future studies should explore whether accessibility to vaccination is associated with specific chronic diseases, as well as with other sociodemographic factors. They should also assess the mediating factors associating schizophrenia with under-vaccination for COVID-19.

The lower rates of vaccination among patients with schizophrenia reported in this study should alert public health policy entities to provide better care in the form of easier access to COVID-19 mitigation/prevention efforts for individuals with schizophrenia.

Dana Tzur Bitan
Department of Behavioral Sciences, Ariel University, Ariel, Israel; Shalvata Mental Health Center, Hod Hasharon, Israel

The author is grateful to D. Comaneshter, I. Gabay, A. Nachman, Y. Schonmann, O. Weinstein, A. Cohen, K. Kridin, I. Krieger and E. Horowitz Leibowitz for their ongoing collaboration and assistance.

1. Wang Q, Xu R, Volkow ND. World Psychiatry 2021;20:124-30.
2. Li L, Li F, Fortunati F et al. JAMA Netw Open 2020;3:e2023282.
3. Tzur Bitan D, Krieger I, Kridin K et al. Schizophr Bull (in press).
4. De Hert M, Mazereel V, Detraux J et al. World Psychiatry 2021;20:54.
5. De Hert M, Correll CU, Bobes J et al. World Psychiatry 2011;10:52-77.
6. Tzur Bitan D, Krieger I, Berkovitch A et al. Gen Hosp Psychiatry 2019;58:1-6.
7. Miles LW, Williams N, Luthy KE et al. J Am Psychiatr Nurses Assoc 2020; 26:172-80.
8. McKee M, Rajan S. Isr J Health Policy Res 2021;10:1-4.
9. Israeli Ministry of Health. Annual report of healthcare-providing companies for 2018. https://www.health.gov.il.

DOI:10.1002/wps.20874

Some good news for psychiatry: resource allocation preferences of the public during the COVID-19 pandemic

The COVID-19 pandemic has put tremendous strain on health care systems all over the world and has particularly challenged mental health services. During the first wave of the pandemic, for reasons of both infection control and resource allocation, many mental health services have been downsized or even closed worldwide. A rapid assessment of 130 World Health Organization member states revealed that more than 60% of countries fully or partially closed community-based mental health services, and more than 40% fully or partially closed inpatient services for substance use disorders[1].

At the same time, it has been widely recognized that the pandemic increases the burden on people with mental illness and puts many healthy people at risk of developing mental health problems[2]. Maintaining adequate mental health services and adapting the way mental health care is delivered during the pandemic is thus of tremendous importance[3,4].

Previous population studies have shown that mental disorders enjoy low standing in the public opinion when it comes to allocation of financial resources[5,6], so there is reason to suspect that the current shortage of health care resources puts people with mental disorders at risk of structural discrimination. In this study, we examined how public priorities on health care spending have evolved from 2001 through 2011 to 2020.

From July to September 2020, a representative face-to-face survey was carried out among the adult population in Germany (N=1,200, response rate: 57%). The survey was a methodologically identical replication of surveys in 2001 (N=5,025, response rate: 65%) and 2011 (N=1,232, response rate: 64%)[7]. In 2020, respondents were asked: "In order to have sufficient resources for the care of patients with the coronavirus disease, it may become necessary to cut budgets for the care of people with other diseases. Please choose from the following list those three conditions where, in your opinion, it would by no means be acceptable to reduce funding for patient care" . They were then presented with a list of nine diseases, including physical conditions such as diabetes, rheumatism, cancer, AIDS and cardiovascular diseases, as well as mental disorders such as Alzheimer's disease, alcoholism, depression and schizophrenia. In 2001 and 2011, the question had been posed similarly, only with the first sentence being framed in more general terms: "There is an increasing shortage of financial resources within the health care system. Please choose from the following list…"

In 2020, depression ranked fourth – after cancer (84%), cardiovascular diseases (60%) and diabetes (41%) – among conditions for which funding should by no means be reduced, with 25% of the respondents selecting it to be spared from budget cuts. Its rise from the 8th position in 2001 and 6th position in 2011 mostly reflected two developments: a growing share of respondents indicating a funding preference for depression (up from 6% in 2001 and 21% in 2011), and a declining share of people giving priority to the funding of AIDS care, which started at 47% in 2001 and went down to 35% in 2011 and 20% in 2020.

Schizophrenia, although remaining on the 8th position in the list, was nevertheless chosen by 17% to be spared from financial cuts in 2020, about doubling its share from 9% in 2001 and 8% in 2011. Alcoholism, in contrast, remained firmly at the bottom of the list, chosen by 5% in 2001, 8% in 2011, and 6% in 2020.

Our results show that, under the unprecedented pressure of the coronavirus pandemic on our health care systems, resources for the treatment of people with mental disorders have solid support among the general public, at least in Germany. Probably, this reflects the extensive coverage of the mental health consequences of the pandemic both in the public media and medical journals[8], and possibly also the personal experience of psychological vulnerability during the crisis.

Comparing our recent survey with those from 2001 and 2011, there is evidence for a trend of growing support for mental health care funding, especially for the treatment of depression. It is striking, however, that alcoholism remains firmly excluded from this supportive public sentiment, despite evidence for an increased burden due to substance use during the pandemic[9].

Our findings are thus reassuring with respect to funding priorities for depression and schizophrenia, with little indication of public support for structural discrimination of people with these disorders. They are worrying, however, with regard to alcohol use disorders. Despite their high prevalence, considerable burden, and available treatment options, people with these latter conditions remain at particular danger to be neglected when competing for

treatment resources, even more so during the current pandemic.

Georg Schomerus[1], Eva Baumann[2], Christian Sander[1], Sven Speerforck[1], Matthias C. Angermeyer[3]
[1]Department of Psychiatry and Psychotherapy, University of Leipzig Medical Center, Leipzig, Germany; [2]Department of Journalism and Communication Research, Hannover University of Music, Drama, and Media, Hannover, Germany; [3]Center for Public Mental Health, Gösing am Wagram, Austria

This study was funded by Fritz Thyssen Stiftung (Az. 10.18.2.009SO).

1. World Health Organization. The impact of COVID-19 on mental, neurological and substance use services: results of a rapid assessment. Geneva: World Health Organization, 2020.
2. Pfefferbaum B, North CS. N Engl J Med 2020;383:510-2.
3. Adhanom Ghebreyesus T. World Psychiatry 2020;19:129-30.
4. Moreno C, Wykes T, Galderisi S et al. Lancet Psychiatry 2020;7:813-24.
5. Matschinger H, Angermeyer MC. Eur Psychiatry 2004;19:478-82.
6. Schomerus G, Matschinger H, Angermeyer MC. Soc Psychiatry Psychiatr Epidemiol 2006;41:369-77.
7. Angermeyer MC, Matschinger H, Link BG et al. Soc Sci Med 2014;103:60-6.
8. Venkatesh A, Edirappuli S. BMJ 2020;369:m1379.
9. Czeisler MÉ, Lane RI, Petrosky E et al. Morb Mortal Wkly Rep 2020;69:1049-57.

DOI:10.1002/wps.20875

A prevalence assessment of prolonged grief disorder in Syrian refugees

Although many studies indicate the elevated rates of mental disorders in refugees[1], relatively little attention has been given to prolonged grief. This is surprising, considering that refugees commonly experience bereavement arising from war, torture, detention, or in the process of fleeing persecution[2].

There has been an increasing focus on problematic grief reactions in recent years, culminating in the new diagnosis of prolonged grief disorder (PGD) being introduced into the ICD-11[3]. PGD is defined as persistent yearning for the deceased, and associated emotional pain, difficulty in accepting the death, a sense of meaninglessness, bitterness about the death, and difficulty in engaging in new activities, persisting beyond six months after the death. The disorder is estimated to occur in 7% of bereaved people[4], but one population-based study of re-settled refugees reported an incidence of 15.8%[5].

There are no large-scale representative sampling studies of PGD in refugees directly affected by war. This is a serious omission in the evidence, because there are millions of refugees directly affected by war, conflict and persecution, and understanding the rates of PGD in this group would help shape better mental health policies to assist those experiencing bereavement. To fill this gap in current knowledge, this study aimed to determine the rate of PGD in a representative sample of adult Syrian refugee parents residing in a camp in Jordan.

Participants were recruited in the process of screening for eligibility for a trial testing the effectiveness of a psychological intervention (ACTRN12619001386123). The screening assessments included adult (>18 years) Syrian refugees who had at least one child residing in two villages in the Azraq Refugee Camp, which hosts approximately 35,000 Syrian refugees. Arabic-speaking interviewers approached each consecutive caravan in two villages in the camp between January 2019 and February 2020, and interviewed one randomly selected adult in the household. The interview included experience and timing of bereavement, cause of death, and relation to the deceased.

PGD was assessed using a 5-item interview that has been used in a previous survey following a major disaster[6], and is consistent with the ICD-11 definition. PGD was operationalized as satisfying the following criteria over the past month for a bereavement that had occurred at least six months earlier: a) yearning for the deceased at least "once a day"; b) at least two of the following three symptoms occurring "quite a lot": bitterness about the loss, difficulty accepting the loss, or feeling that life is meaningless; and c) endorsing functional impairment as a result of the grief. Additionally, psychological distress was assessed with the Kessler Psychological Distress Scale (K10), on which severe mental disorder was categorized as scores ≥30. Disability was evaluated using the WHO Disability Assessment Schedule 2.0 (WHODAS 2.0), on which disability was defined as scores ≥17. The study was approved by the Institutional Review Board of the King Hussein Cancer Centre, Jordan.

We assessed 955 participants (67.3% females). Of these, 564 (59.1%) reported bereavement, with cause of death including war (26.5%), accident (8.3%), natural causes (64.6%) or other causes (2.6%). The deceased was a parent (32.2%), spouse (4.2%), child (5.7%) or relative/friend (57.9%).

Among those experiencing bereavement, 85 (15.1%) met criteria for PGD, which comprised 8.9% of the entire sample. In terms of the specific symptoms of PGD, 478 (84.8% of the bereaved sample) reported persistent yearning, 447 (79.3%) bitterness about the loss, 251 (44.5%) feeling meaningless, 167 (29.6%) difficulty accepting the death, and 149 (26.4%) impairment resulting from their grief.

There was no relationship between marital status, educational level, cause of death or relationship to the deceased and the likelihood of developing PGD. In terms of the association between PGD and ongoing problems, after controlling for cause of death, refugees with PGD were more likely to have a serious mental disorder (68.2% vs. 56.3%: OR=1.6, 95% CI: 1.0-2.6). PGD was not associated with greater disability.

These findings are significant as this is the first study reporting on the prevalence of PGD, by representative sampling, in a population of refugees directly affected by war. The percentage of 15.1% of bereaved refugees experiencing PGD is markedly higher than

those reported in other population-based studies, although it is commensurate with one prior population-based study of refugees[5].

The finding that 8.9% of the screened refugees had PGD highlights the importance of bereavement-related mental health issues affecting refugees exposed to war. This is arguably because a large proportion of Syrian refugees are exposed to events that can cause traumatic or unexpected death, and the nature of the deaths increases the risk for PGD.

It was somewhat surprising that bereavement caused by deaths incurred by war or accident was not more predictive of PGD, as violent and accidental deaths have previously been shown to be associated with a higher risk for PGD[7]. This lack of association may be explained by extreme conditions due to war and forced displacement, potentially causing many natural deaths. These deaths, which are nonetheless unexpected, potentially predispose refugees to PGD. The observation that refugees with PGD are more likely to have a serious mental disorder underscores the psychological costs associated with this condition.

The rate of PGD in war-affected refugees points to the need for greater attention to managing the disorder in this population. Considering that there are millions of refugees similarly affected worldwide, it is important to develop and evaluate programs that can address PGD. Although there are proven treatments for PGD based on cognitive behavioral strategies[8], these are lengthy, rely on mental health specialists, and are costly for health systems, which limits the capacity for poorly resourced countries to implement them on a large scale.

Recent initiatives that rely on task-shifting, in which non-specialists are trained to deliver mental health programs[9], can be implemented to reduce common mental disorders. However, these have yet to address the clinical needs of people with PGD. The increasing evidence of the widespread psychological problems caused by bereavement in refugees highlights the need to develop scalable interventions that can address PGD.

Richard A. Bryant[1], Ahmad Bawaneh[2], Luana Giardinelli[2], Manar Awwad[2], Hadeel Al-Hayek[2], Aemal Akhtar[1,3]
[1]School of Psychology, University of New South Wales, Sydney, NSW, Australia; [2]Jordan Country Office, International Medical Corps, Amman, Jordan; [3]Clinical, Neuro and Developmental Psychology, Vrije Universiteit, Amsterdam, The Netherlands

This project was funded by a National Health and Medical Research Council (NHMRC) - European Union Collaborative Grant (APP1442605).

1. Charlson F, van Ommeren M, Flaxman A et al. Lancet 2019;394:240-8.
2. Tay AK, Rees S, Chen J et al. Soc Psychiatry Psychiatr Epidemiol 2016;51:395-406.
3. Reed GM, First MD, Kogan CS et al. World Psychiatry 2019;18:3-19.
4. Lundorff M, Holmgren H, Zachariae R et al. J Affect Disord 2017;212:138-49.
5. Bryant RA, Edwards B, Creamer M et al. Epidemiol Psychiatr Sci 2019:e44.
6. Shear MK, McLaughlin KA, Ghesquiere A et al. Depress Anxiety 2011;28:648-57.
7. van Denderen M, de Keijser J, Kleen M et al. Trauma Violence Abuse 2015;16:70-80.
8. Shear MK, Reynolds CF 3rd, Simon NM et al. JAMA Psychiatry 2016;73:685-94.
9. Bryant RA. World Psychiatry 2019;18:259-69.

DOI:10.1002/wps.20876

Clinical implications of co-occurring prolonged grief disorder in patients with treatment-resistant major depressive disorder

Prolonged grief disorder (PGD) is now an official diagnosis in both the ICD-11 and the DSM-5-TR[1]. It is a distressing and disruptive condition that often occurs concomitantly with depressive and other psychiatric disorders, yet the consequences of this comorbidity are not fully understood[2]. We examined the importance of co-occurring PGD in a large sample of patients with treatment-resistant major depressive disorder (TRD).

The VA Augmentation and Switching Treatments for Depression (VAST-D) Cooperative Study included 1,522 psychiatric outpatients with TRD (85.2% males; 69.2% White, 25.6% African American; mean age 54.4±12.2 years) who were randomized to switch the index antidepressant to bupropion SR, combine the index antidepressant with bupropion SR, or augment with aripiprazole[3]. Of these patients, 1,416 (93.0%) had experienced the death of a loved one in their lifetime. The mean time since the death was 11.9 years, yet 600 (42.4% of the bereaved) felt that their grief still interfered, at least somewhat, with their lives.

The study's primary outcome, remission, was operationalized as a 16-item Quick Inventory of Depressive Symptomatology-Clinician Rated (QIDS-C_{16})[4] score of ≤5 at two consecutive visits during the 12-week acute treatment phase. Suicide risk was assessed with the Columbia-Suicide Severity Rating Scale (C-SSRS)[5] both at baseline and at the end of the acute treatment phase. For this report, we utilized the number and percentage of participants who endorsed recent (within the past 3 months) passive suicidal ideation (e.g., wish to be dead) and active suicidal ideation (actual intent and/or plan).

At the baseline visit, bereaved VAST-D participants completed the 5-item self-rated Brief Grief Questionnaire[6] (difficulty accepting the death, grief interfering with life, intrusive thoughts of the person or the death, avoiding reminders of the loss, and feeling cut off or distant from others) and two additional items for yearning and grief intensity. Each of the seven items was rated on a 3-point scale: 0 = not at all, 1 = somewhat, 2 = a lot.

According to the ICD-11[7], a PGD diagnosis was assigned when a participant indicated all of the following: a) time since death of the loved one ≥6 months; b) endorsed "a lot" on at least one of the following: grief intensity or yearning for the deceased; c) endorsed "a lot" on at least one of the following: trouble accepting the death, troublesome images of the lost person or his/her death, avoiding reminders, feeling cut off or distant from others;

d) endorsed "somewhat" or "a lot" on "grief continues to interfere with life". Participants fulfilling these requirements were 276 (19.5% of the bereaved). Patients bereaved >6 months without PGD were categorized as "ordinary long-term grief" (N=1,041, 73.5%). Those bereaved less than 6 months were categorized as "acute grief" (N=99, 7.0%).

PGD was significantly more common among African American bereaved participants (96/343, 28.0%) than White ones (145/948, 15.3%) (X^2=37.26, p<0.0001). There were no differences between bereaved participants with and without PGD on age, gender, employment, age of first treatment for depression, or duration of current episode. Time since the death of the loved one was not significantly different between those with PGD and those with ordinary grief.

Among participants bereaved >6 months, those with PGD were more likely to feel that their depression was related to their grief. They had at baseline significantly more severe depression on the QIDS-C_{16} (p<0.0001), more anxiety on the Beck Anxiety Scale (p<0.0001), worse quality of life on the Quality of Life Enjoyment and Satisfaction Questionnaire (p<0.0001), more passive and active suicidal ideation on the C-SSRS (p<0.05), and more post-traumatic stress disorder on the Mini International Neuropsychiatric Interview (p<0.0001). In addition, they had experienced more early childhood life adversity as assessed by Adverse Childhood Experiences Survey (ACES) (p<0.0001).

Remission of depression at the end of the treatment period was significantly less likely in patients with PGD (N=50, 18.1%) than in those with ordinary long-term grief (N=296, 28.4%) and those with acute grief (N=20, 20.2%) (X^2=13.9, p<0.001). Also, with or without depression remission, active suicidal plans and/or intent at the end of the treatment period were more frequently reported in patients with PGD (N=13, 4.7%) than in those with ordinary long-term grief (N=15, 1.4%) and those with acute grief (N=3, 3.0%) (X^2=11.2, p<0.01).

To our awareness, this is the first study to systematically assess the effects of co-occurring grief in a large sample of patients with treatment-resistant depression. We found that patients with co-occurring PGD were less likely to remit from their depressive episode than patients with ordinary long-term grief. Those with acute grief were somewhere in between the two other groups. The same pattern was true for active suicidal ideation.

These findings underscore the importance of an accurate diagnosis of PGD in patients with treatment-resistant depression, so that targeted clinical attention can ensue. The only study assessing the effects of antidepressants in depressed patients with or without PGD found that medications relieved depressive symptoms in those with PGD, but only if they were also receiving a grief-targeted psychotherapy[8]. This suggests that patients with treatment-resistant depression who also have PGD would greatly benefit from a grief-focused intervention in addition to the depression-focused treatment.

Sidney Zisook[1,2], Somaia Mohamad[3], Gary Johnson[4], Ilanit Tal[2], Gerardo Villarreal[5,6], James Allen Wilcox[7], Katherine M. Shear[8,9]
[1]Department of Psychiatry, University of California San Diego, La Jolla, CA, USA; [2]VA San Diego Healthcare System, La Jolla, CA, USA; [3]Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA; [4]Cooperative Studies Program Coordinating Center, VA Connecticut Healthcare System, West Haven, CT, USA; [5]VA New Mexico Healthcare System, Albuquerque, NM, USA; [6]Department of Psychiatry, University of New Mexico, Albuquerque, NM, USA; [7]Southern Arizona VA Healthcare System, Tucson, AZ, USA; [8]Center for Complicated Grief, New York, NY, USA; [9]Columbia University School of Social Work, College of Physicians and Surgeons, New York, NY, USA

This study was supported by the Cooperative Studies Program (CSP), Department of Veterans Affairs, VA Office of Research and Development. The contents of this letter do not necessarily represent the views of the US Department of Veterans Affairs or the US Government.

1. Prigerson HG, Boelin PA, Xu J et al. World Psychiatry 2021;20:96-106.
2. Zisook S, Shear K. World Psychiatry 2009;8:67-74.
3. Mohamed S, Johnson GR, Chen P et al. JAMA 2017;318:132-45.
4. Rush AJ, Trivedi MH, Ibrahim HM et al. Biol Psychiatry 2003;54:573-83.
5. Posner K, Brent D, Lucas C et al. Columbia-Suicide Severity Rating Scale (C-SSRS). New York: Columbia University Medical Center, 2008.
6. Shear KM, Jackson CT, Essock SM et al. Psychiatr Serv 2006;57:1291-7.
7. Reed GM, First MB, Kogan CS. World Psychiatry 2019;18:3-19.
8. Shear KM, Reynolds CF, Simon NM et al. JAMA Psychiatry 2016;73:685-94.

DOI:10.1002/wps.20877

Self-interpretation and meaning-making processes: re-humanizing research on early psychosis

Dimensional, dynamic and transdiagnostic approaches to explaining, classifying and treating psychopathological symptoms are currently leading the way into a new era of early intervention for psychosis research and practice – beyond the "at risk mental state" concept[1].

While research into the clinical and biological predictors of psychosis onset is slowly advancing, we argue that the role of self-interpretation and meaning-making processes remains relatively neglected. If humans are self-interpreting agents, constantly interacting with others through culture, then meanings are likely to be a significant driving factor in the transition from vulnerable selves to persons affected by psychosis.

Notably, robust evidence emerging from cognitive-oriented studies points to a central role of appraisal (i.e., the meaning and interpretation of anomalous experiences) in re-directing the phenotypic trajectory of illness towards a "need-for-care" status, as compared to a more benign non-clinical course[2]. This is consistent with person-centred dialectical models of schizophrenia[3], emphasizing the active role of the person in shaping psychopathological syndromes. We argue that prediction research could benefit from exploring the effect of personal meanings, family narratives and cultural signifiers on the clinical outcome of pre-psychotic experiences.

As a response to growing consensus in the field that current

syndromic classifications do not fit the ebb and flow of emerging microphenotypes, new dynamic models of the onset of mental disorders have been suggested[4]. These models propose the application of mathematical principles governing complex systems to the study of emerging psychopathology in psychosis. By capturing the fast-moving and unstable nature of psychopathological states, the hope is that such models will aid prediction research and allow for the optimization of early treatments.

While some of these models have been used successfully to predict the behaviour of complex systems (e.g., ecosystems and financial markets), they do not account for the fact that human beings will actively interpret and respond to changes in the coordinates of the system, thereby contributing to increasing complexity. These interpretations will partly depend upon the experiential and immediately felt quality of self/world anomalous experiences (SWAE), but also upon the collectivity of meanings, attitudes, beliefs and values as a function of the person's unique biographical, sociocultural and historical context.

SWAE are increasingly recognized as core psychopathological features of psychosis, particularly within the schizophrenia spectrum. According to the ipseity disturbance model of schizophrenia[5], the longitudinal progression of symptoms is driven by a gradual exacerbation of diminished self-affection and disturbed grip or hold on the world. These latter features are thought to be fairly stable and enduring vulnerabilities affecting the conventional or common-sense way of inhabiting the world. However, very little is known – and even less has been investigated empirically – about how individuals understand, make sense of, and appropriate emerging SWAE in the early stages of psychosis. How do people take a position (i.e., cope with or respond to SWAE)? What meaning resources do they turn to? What role do meaning-making processes play in the shaping of psychopathological phenomena and clinical outcomes?

The idea that human beings are self-interpreting agents is not new in the phenomenological literature. On the foundations laid by K. Jaspers' account of the "patient's attitude to his illness", also called "position-taking" (*Stellungnahme*), other classical authors have developed dialectical models of the onset of schizophrenia (Pinel, Bleuler, Wyrsch, de Clérambault, Blankenburg, Mayer-Gross, J.S. Strauss, and others). Their models emphasize the active and dynamic interplay between the person and his/her basic experiential anomalies, suggesting that such interaction might influence the development of different schizophrenic phenotypes from the common root of self-disorders. Stemming from this classical literature, a person-centred dialectical approach to schizophrenia has been proposed[6], which is concerned not only with the phenomenal level of experience (i.e., changes in the sense of self and lived world), but also with the different ways in which patients "take a position" in the face of emerging psychotic-like phenomena.

This dynamic sense-making activity in the face of anomalous experiences is clearly reflected in the first-person narratives of individuals in the early stages of psychosis[7]. From this growing body of qualitative literature, two inextricable dimensions seem to emerge. On one level, an all-enveloping sense of unreality, threat, disorientation and uncertainty often accompanies disordered experiences of self, time, space and body. A number of phenomenologists and psychopathologists have referred to "background feelings"[8] as a way to capture the subtle but all-encompassing, pervasive, atmospheric qualities of subjective life that drive our interpretation of the world. On another level, an ongoing reflective activity in the form of "searching for meaning", "making meaning" and "finding meaning" appears overwhelmingly central to the process of articulating the subjective experience of illness onset. According to the most influential phenomenological models of the onset of schizophrenia, these self-interpreting efforts are triggered by a state of "perplexity", whereby reality loses its taken-for-granted, common-sense meanings and endangers the core sense of self.

Understanding how people make sense of these puzzling experiences and how these meanings impact on their behaviour is crucial in order to develop adequate dynamic models of emerging psychosis. In addition to a nuanced and in-depth psychopathological examination and clinical phenotyping, these models should also reflect the agentic role of the person in shaping the course of illness within a specific meaning-giving context.

But how to account for these meanings in a way that can be effectively operationalized to inform predictive models while avoiding the de-humanization of psychiatry? Even if research shows that psychosis evolves in a staging fashion akin to the biological process of growing and spreading of cancer, we still need to deal with a distinct set of features that places the mental realm apart from the physical realm. In this sense, meanings are not only something mechanically generated by the mind in the same way as the immune system defensively produces antibodies. Meanings also respond to a fundamental human quest for coherence and purpose, prospectively shaping our way of being in the world within a value-laden framework of significance.

One way in which the cognitive aspects of this human response have been operationalized is through the conception of appraisal (i.e., the meaning attributed to pre-psychotic/psychotic-like experiences). As an alternative phenomenologically-informed approach, we suggest the integration of appraisal into the construct of position-taking, which also takes into account the "background states of the person that exist before the experience and contribute to determining the experience itself"[9]. The principal dimensions of position-taking are: a) the person's emotional tone, i.e. the background feelings associated with the SWAE, and b) the person's working-through, i.e. the cognitive and narrative elaboration of his/her SWAE.

We believe that this second level of assessment might also help shed light on the transdiagnostic occurrence of SWAE across schizophrenia, depersonalization/derealization disorder and panic disorder, but also non-clinical states such as intense introspection. Given that some of these experiential anomalies might be indistinguishable at a phenomenal level, the subsequent clinical-diagnostic trajectory might be (at least partly) driven by the kind of position-taking adopted, further influencing subjective distress and need-for-care.

Improved understanding of the relevance of meaning-making

in psychosis research can ultimately enhance our ability to predict the dynamic evolution of clinical high-risk states without losing focus on the person as a self-interpreting being interacting with his/her sociocultural world.

Rosa Ritunnano[1], Giovanni Stanghellini[2,3], Matthew R. Broome[1]
[1]Institute for Mental Health, University of Birmingham, Birmingham, UK; [2]"G. d'Annunzio" University, Chieti, Italy; [3]"D. Portales" University, Santiago, Chile

1. Shah JL, Scott J, McGorry PD et al. World Psychiatry 2020;19:233-42.
2. Peters E, Ward T, Jackson M et al. Lancet Psychiatry 2017;4:927-36.
3. Stanghellini G, Bolton D, Fulford WKM. Schizophr Bull 2013;39:287-94.
4. Nelson B, McGorry PD, Wichers M et al. JAMA Psychiatry 2017;74:528.
5. Sass L, Borda JP, Madeira L et al. Schizophr Bull 2018;44:720-7.
6. Stanghellini G, Rosfort R. Curr Opin Psychiatry 2015;28:256-63.
7. McCarthy-Jones S, Marriott M, Knowles R et al. Psychosis 2013;5:1-16.
8. Stanghellini G, Raballo A. J Affect Disord 2015;171:171-8.
9. Woods A, Wilkinson S. Lancet Psychiatry 2017;4:891-2.

DOI:10.1002/wps.20878

Clinical relevance of general and specific dimensions in bifactor models of psychotic disorders

Increasing empirical evidence indicates that psychotic disorders are more accurately understood as continuously distributed dimensions of psychopathology, which can be arranged in a hierarchical structure framework[1].

Bifactor models are a subset of hierarchical models that can inform several important questions about the structure and external validity of psychopathology, including the nature of the general factor given the pattern of symptom loadings, the amount of variance in the specific factors due to the general factor, and the unique external correlates of the general and specific factors[2,3].

While previous studies of bifactor models in psychotic disorders outlined the clinical and research relevance of separating general and specific psychopathology[4,5], most of them have been limited by one or more problems regarding sample size, sample representativeness, symptomatic coverage and systematic external validation. Of particular concern are the conflicting results in the literature regarding the nature and validity indicators of the general psychosis factor.

Using a broad sample of subjects with the full spectrum of psychotic disorders, we tested the construct validity of alternative bifactor models of symptoms, and the criterion validity of the factors in the best-fitting model against 21 external validators, including antecedents, illness-related features, treatment-related variables and psychosocial impairment.

The study population consisted of 2,240 subjects with a diagnosis of a DSM-IV functional psychotic disorder from a defined catchment area in Navarra (Spain)[6], and included three subsamples: first-episode admissions (N=486), multi-episode admissions (N=660) and outpatients (N=1,094).

In the whole study sample, 57% of subjects were male, the mean age was 38.7±15.2 years, and the mean Global Assessment of Functioning scale score was 65.4±20.8. DSM-IV diagnoses included schizophrenia (N=908), schizophreniform disorder (N=180), delusional disorder (N=120), brief psychotic disorder (N=179), schizoaffective disorder (N=124), bipolar disorder (N=345), major depressive disorder (N=245) and psychotic disorders not otherwise specified (N=139). Subjects underwent an extensive clinical examination by the Comprehensive Assessment of Symptoms and History (CASH).

Bifactor models were estimated in each subsample and in the total sample to establish the best-fitting model using the omegaSem routine in the R package. Seven alternative models were estimated using multiple information criteria, including X^2, log-likelihood, Akaike information criterion (AIC), Bayesian information criterion (BIC), sample-size adjusted Bayesian information criterion (SABIC), confirmatory fit index, Tucker-Lewis index and root mean square error of approximation. Following identification of the best-fitting model, we examined its validity against the external indicators using linear regression. As an omnibus test of the associations between each factor and the validators, we used the general linear model, with partial η^2 indicating the percentage of the unique variance in each of the factors that is explained by the validators. To avoid false positive findings, p was set at <0.01.

For all fit measures compared, the best-fitting model in each subsample and in the whole sample was a model comprising a general factor and six specific factors of diminished expressivity, avolition-anhedonia, reality distortion, disorganization, mania and depression. The difference in AIC, BIC and SABIC parameters between this model and that with the second better fit was >211 in each subsample and in the whole sample, indicating very strong support for the superiority of this model over the others.

Factor loadings on the general factor were relatively large for disorganization symptoms (mean λ=.58), moderate for negative symptoms (mean λ=.44), low for reality distortion symptoms (mean λ=.23) and very low for mood symptoms (mean λ=.16). The amount of variance in the symptom scores that was accounted for by the general and specific factors was 61% and 31%, respectively.

The general factor exhibited more significant associations with the validators than any other factor, since it predicted most of the validators and explained 54% of their variance. The general factor score was significantly related to familial risk of schizophrenia (β=.056, 95% CI: .020-.125), poor educational performance (β=.165, 95% CI: .125-.205), childhood adversity (β=.065, 95% CI: .024-.105), poor premorbid adjustment (β=.155, 95% CI: .118-.191), duration of untreated psychosis (β=.110, 95% CI: .072-.174), early age at illness onset (β=.194, 95% CI: .153-.234), chronic illness onset (β=.140, 95% CI: .102-.179), number of hospitalizations

(β=.107, 95% CI: .067-.148), cognitive impairment (β=.286, 95% CI: .249-.322), chronic illness course (β=.129, 95% CI: .092-.176), dose-years of antipsychotics (β=.060, 95% CI: .019-.100), current dose of antipsychotics (β=.596, 95% CI: .562-.622), poor treatment response (β=.176, 95% CI: .138-.213), and four domains of psychosocial functioning impairment (β=.223, 95% CI: .138-.290).

In decreasing order of importance, the contribution of the individual specific factors in explaining the validity indicators was as follows: avolition-anhedonia 36%, diminished expressivity 26%, depression 23%, reality distortion 18%, mania 12%, and disorganization 8%. The strongest association of both diminished expressivity and avolition-anhedonia factors was with social impairment (β=.212, 95% CI: .168-.252, and β=.436, 95% CI: .402-.488, respectively). The strongest association of reality distortion, disorganization, mania and depression factors was with current dose of antipsychotics (β=.364, 95% CI: .336-.394), poor response to treatment (β=.136, 95% CI: .099-.173), familial risk of bipolar disorder (β=.227, 95% CI: .182-.251) and polypharmacy (β=.438, 95% CI: .397-.478), respectively.

In summary, the construct validity of the bifactor model with six specific factors was verified by statistical fit indices. The psychosis general factor accounted for a majority of the variance among symptoms and was the factor most often related to external indicators, explaining 54% of their variation, all of which posits the general factor as a core construct of psychotic pathology. The observed association pattern of the general factor with external validators suggests that it may be indexing both a neurodevelopmental signature of psychotic disorders and a marker of overall illness severity. Furthermore, these findings are broadly consistent with the hypothesis that the general psychosis factor might be at the extreme end of severity of the general psychopathology factor[8].

Despite not being an explicit component of psychiatric nosology, the general psychosis factor provides an organizing model for understanding psychotic disorders that might have important implications for predicting relevant risk factors and clinical features[9]. Although a direct causal interpretation is not possible, the external correlates of the general psychosis factor might be meaningful for clinical practice.

Assessing for the general factor could allow the identification of relevant background variables and clinical features in subjects with psychotic disorders, who may benefit from careful clinical monitoring and possibly more tailored interventions. For instance, a high score on the general factor, irrespective of diagnostic categories, may alert the clinician about the risk of a broad range of poor clinical outcomes.

Victor Peralta[1,2], Gustavo J. Gil-Berrozpe[2,3], Ana Sánchez-Torres[2,3], Manuel J. Cuesta[2,3]
[1]Mental Health Department, Servicio Navarro de Salud-Osasunbidea, Navarra, Spain; [2]Navarrabiomed and Instituto de Investigación Sanitaria de Navarra, Navarra, Spain; [3]Psychiatry Service, Complejo Hospitalario de Navarra, Navarra, Spain

The study was funded by the Spanish Ministry of Economy, Industry and Competitiveness (grant PI16/02148) and the Regional Government of Navarra (grant 31/17).

1. Kotov R, Jonas KG, Carpenter WT et al. World Psychiatry 2020;19:151-72.
2. Bornovalova MA, Choate AM, Fatimah H et al. Biol Psychiatry 2020;88:18-27.
3. Moore TM, Kaczkurkin AN, Durham EL et al. J Abnorm Psychol 2020;129:677-88.
4. Reininghaus U, Priebe S, Bentall RP. Schizophr Bull 2013;39:884-95.
5. Reininghaus U, Bohnke JR, Chavez-Baldini U et al. World Psychiatry 2019;18:67-76.
6. Peralta V, Gil-Berrozpe GJ, Librero J et al. Schizophr Bull Open 2020;1:sgaa008.
7. Andreasen NC, Flaum M, Arndt S. Arch Gen Psychiatry 1992;49:615-23.
8. Caspi A, Moffitt TE. Am J Psychiatry 2018;175:831-44.
9. Lahey BB, Krueger RF, Rathouz PJ et al. World Psychiatry 2017;16:142-4.

DOI:10.1002/wps.20879

The WPA Code of Ethics for Psychiatry

In October 2020, the WPA General Assembly adopted the Association's first Code of Ethics for Psychiatry[1]. Developed by the Standing Committee on Ethics and Review, the Code was drafted and revised, with input from psychiatric societies worldwide, over almost a decade prior to being finalized and adopted. Its four sections cover ethical issues in the major areas of psychiatric endeavor: the clinical practice of psychiatry, psychiatric education, psychiatric research, and psychiatric participation in the promotion of public health, including public mental health. The Code now stands as WPA's official statement on the ethics of psychiatry, having superseded the Declaration of Madrid.

The formulation of this Code of Ethics was stimulated by the recognition that WPA's previous ethics documents, beginning with the Declaration of Hawaii in 1977 and culminating in the Declaration of Madrid in 1996, were incomplete in their coverage of the principal areas of psychiatric ethics. Thus, they lacked one of the primary attributes of an ethics code, i.e., a systematic approach to defining the parameters of professional ethical behavior[2]. To ensure coverage of the major areas of psychiatric ethics, the Standing Committee on Ethics and Review began by assembling all existing WPA documents related to ethics, as well as available ethics codes from Member Societies around the world.

Key issues, especially if they were reflected in multiple sources – suggesting international relevance – were compiled and organized into relevant sections, and approaches reflecting the generally agreed-upon ethical principles were defined. Within each of the four areas of psychiatric activity noted above, the Code's provisions were organized around five principles of medical ethics: beneficence, respect for patients (autonomy), non-maleficence, and the imperatives to improve standards of psychiatric practice and to apply psychiatric expertise to the service of society (including seeking equity in prevention, treatment and rehabilitation of psychiatric disorders). The resulting draft of the Code was then reviewed by the WPA Executive Committee and revised in response to its comments, and subsequently circulated to all Member Societies for their input. Initial discussion at the WPA General Assembly in Berlin in 2017 was followed by clarification of the goals of the document, which was then approved by the Executive Committee and the General Assembly in 2020.

As with other codes of professional ethics, the new WPA Code serves several functions[2]. Individual psychiatrists, especially in countries where national psychiatric societies have not yet formulated ethics codes, can draw guidance from the Code when faced with ethical challenges in their professional activities. The public and members of other medical specialties and other health professions can look to the Code to shape their expectations of their interactions with psychiatrists. Member Societies of the WPA can compare their existing codes with the new WPA Code to identify gaps that may need to be addressed, and those Societies without codes can use the Code as the foundation of their efforts to develop their own. Finally, the Code will alert governments to the ethical boundaries of psychiatric practice and provide support for psychiatrists who may be pressured to act unethically to support political ends.

The Code is not meant to supplant national codes of ethics, which can better address the particular circumstances of each country and incorporate societal values. However, Member Societies have been asked to endorse the principles embodied in the WPA Code and to confirm that their codes are not in conflict with them. Individual psychiatrists will continue to be subject to the provisions of their national societies' codes.

We encourage psychiatrists to review the Code, which is easily accessible online[1]. Its provisions are framed as affirmative statements of psychiatric behavior. Here, we provide some illustrative examples from each of its sections:

- Ethics in the clinical practice of psychiatry
  - Psychiatrists recognize that their primary obligation in the clinical setting is to pursue the wellbeing of their patients, in light of the best available evidence and clinical experience.
  - Psychiatrists are sensitive to the needs of patients' families, carers, and others who are affected by patients' disorders. They provide education and support to these groups, empowering them to assist patients in coping with their disorders and achieving their personal goals. Psychiatrists recognize that optimal clinical care is rendered through collaboration among patients, carers and clinicians, along with other team members, and they work to resolve differences and encourage cooperation among them.
- Ethics in psychiatric education
  - Psychiatrists recognize an obligation to share their knowledge of biological, psychological and social determinants of mental health; of psychiatric diagnosis, treatment and prevention; and of systems of mental health care with trainees and practitioners in psychiatry, other medical specialties, other mental health professions, and the general public. They fulfill this responsibility in a professional manner that reflects up-to-date, evidence-based knowledge of the field.
  - Acknowledging the vulnerable position of students and trainees and the trust that they place in their teachers, psychiatrists avoid exploitation in their educational roles, e.g., they do not take credit for work done by students and trainees, appropriately balance education and requirements for service, and do not abuse their relationship with their students and trainees in any way.
- Ethics in psychiatric research and publication
  - Psychiatrists recognize that research and publication are vital in improving care for current and future patients and improving the health of the population as a whole. Hence, they acknowledge their responsibility to help advance knowledge about the nature of psychiatric disorders,

including risk and protective factors, and their treatment. Not all psychiatrists will be interested in or carry out research, but everyone should be able to understand, interpret and apply research findings, when appropriate, in a manner consistent with psychiatric ethics.
    - Psychiatrists present the results of their research fairly, calling attention to both positive and negative results, and focusing both on the potential value of their findings and the limitations of the conclusions that can be drawn from their data.
- Ethics in psychiatric public health
    - Psychiatrists take every opportunity to combat the stigma of psychiatric disorders in the practice of their profession and participate in public health activities that target the stigma of psychiatric disorders to the extent of their abilities to do so.
    - In their commitment to advancing mental health, psychiatrists promote distributive justice, including equitable allocation of resources for the prevention, treatment and rehabilitation of psychiatric disorders. Psychiatrists advocate in particular for support for mental health programs, especially in but not limited to developing countries and in areas where care for persons with psychiatric disorders is non-existent or rudimentary.

The WPA explicitly recognizes that an ethics code needs to be a living document, responsive to changes in knowledge of psychiatric disorders and approaches to psychiatric treatment. Moreover, professional ethics codes have always been responsive to societal changes as well, which often lead to reconceptualization of professional obligations or rebalancing among conflicting principles[3]. Hence, it is anticipated that changes to the Code will be proposed by Member Societies and/or the Standing Committee on Ethics and Review and reflected in subsequent versions of the document.

Paul S. Appelbaum[1], Samuel Tyano[2]
[1]Department of Psychiatry, Columbia University, New York State Psychiatric Institute, New York, NY, USA; [2]Sackler School of Medicine, Tel Aviv University Medical School, Tel Aviv, Israel

Both authors are past chairs of the WPA Standing Committee on Ethics and Review. They acknowledge the other members of the Committee who contributed to the original drafting of the Code (J. Arboleda-Flórez, A. Javed and C. Soldatos), as well as those who contributed to the subsequent revisions (S. Galderisi, M.J. Wise, F. Lieh-Mak, D. Moussaoui, N. Sartorius and D. Stewart).

1. World Psychiatric Association. Code of Ethics for Psychiatry. https://www.wpanet.org/policies.
2. Beyerstein D. In: Winkler E, Coombs JR (eds). Applied ethics: a reader. Hoboken: Wiley, 1993.
3. Backof JF, Martin CL. J Bus Ethics 1991;10:99-110.

DOI:10.1002/wps.20861

Future WPA President's public mental health agenda

The devastating global burden of mental disorders continues unabated and, in fact, has been significantly exacerbated by the COVID-19 pandemic. More than one billion people suffered from mental and addictive disorders prior to the pandemic[1]. The true burden lies not only among identified patients, but also within the general population, where stress, burnout, depression, anxiety and sleep disorders are often unrecognized, untreated, and seldomly prevented[2]. Now, overwhelmingly, these conditions are compounded by the impact of the pandemic, including death of loved ones and associated grief[3-5]. Most preventable ill-health conditions have major consequences not only for the individual well-being but also for every nation's economic prosperity.

Several well-known factors amenable to intervention contribute to the global burden of mental disorders. First and foremost are stigma and discrimination[6,7]. The impact of both environmental and societal factors on the public's mental health are generally not sufficiently recognized[8,9], nor adequately addressed. Scepticism remains commonplace about the role of healthy lifestyles in promoting and preserving good mental health. In 2015, the 2030 United Nations (UN) Agenda for Sustainable Development, with seventeen sustainable development goals (SDGs), was agreed upon by Member States, requiring every country to act in a global effort towards a better future and health for all[10]. While that agenda has limitations, and some would suggest insurmountable barriers, it provides a useful global framework for action.

The WPA has an obligation to play an important role in contributing to the achievement of the SDGs. Most notably, the third SDG, "Health and Wellbeing", includes decreasing suicide rates as an indicator of progress. Such a reduction should be a high priority for the WPA. Mental health stigma and discrimination due to age, race, ethnicity, nationality, religion, gender, sexual orientation and other factors remain widespread. The tenth SDG, "Reducing Inequality" is another important goal for the WPA, to ensure that no one is left behind. The seventeenth SDG, "Partnership between Governments, Associations, the Private Sector and Civil Society", deserves the WPA's attention to secure better public mental health. Surely the COVID-19 pandemic has demonstrated the centrality of this SDG.

As WPA President-Elect, my vision includes increasing collaboration with UN agencies, to increase awareness about public mental health, and to facilitate WPA Member Societies' contributions to the achievement of the SDGs. Collectively, we must influence not only UN bodies, but also national and local politicians. This can be achieved by learning from each other through a shared focus on collaborative educational and research activities, devoted to improving public mental health and carried out in parallel with improved recognition and treatment of psychiatric disorders.

I was humbled by the major obstacles cited when listening to the needs of national psychiatric associations during my WPA presidential campaign. It became painfully apparent that lack of fluency in English for a substantial portion of the membership hindered communication

and utilization of scientific and health promoting materials, often published in English only. Therefore, one of my priorities is to develop educational hubs, based on the six World Health Organization (WHO) official languages (Arabic, Chinese, English, French, Russian and Spanish), focused on high priority topics such as schizophrenia, substance use, depression and suicide prevention. These and many other topics will be addressed in response to the expressed needs of Member Societies.

Increasing understanding of what constitutes public mental health among psychiatrists and the general public, including collaboration with patient and family organizations, is a related goal, as it requires providing materials in the appropriate language to Member Societies, so they can assist local communities in their prevention and intervention efforts. The resulting community collaboration, I believe, should be an integral part of the everyday activities of psychiatrists. Treatment activities and public mental health promotion initiatives should go hand in hand, reinforcing each other to achieve optimal outcomes.

Utilizing existing materials created by WPA Member Societies will be advantageous, and new materials will be created as needed. All materials disseminated within the six educational language hubs will also be culturally adapted. Ideally, these efforts will come from students, residents and WPA members, including those who have retired, thus allowing participants to also serve as ambassadors and mentors to local and regional public mental health staff and programmes.

My experience in building international collaborations as Head of the Department of Public Health Sciences at Karolinska Institute and Director of the WHO Suicide Prevention Centre will aid in this process. Presently, I lead and participate in clinical and community projects throughout the world, particularly on developing mental health services and suicide prevention during the COVID-19 pandemlc[11]. It is my hope that this work can be expanded.

To the extent possible, I would like to mentor WPA members, so they too can assume international leadership positions. I know that shared knowledge about public mental health and a constant dialogue among colleagues from different cultural contexts plays a vital role in the quality of our work. Moreover, it has the potential to expand the perspective of our members, which contributes to increased leadership when communicating with politicians and decision-making bodies. It also assists in identifying allies and securing funding. Overall, my goal is the further development and enhancement of a global network of psychiatrists in the WPA, who can assume leadership positions, locally and internationally, as we jointly seek better outcomes for all.

Finally, I want to acknowledge that the action plans and mosaic of current and previous WPA activities constitute the critical platform necessary for actualizing my own vision. All the ongoing high-quality activities at the WPA are worthy of our full support. As happened with my predecessors, each of us brings to the WPA his/her own unique gifts and possibilities. Personally, I have had the privilege of pursuing my interests in suicide prevention as the core of my professional life. I look forward to enlisting the collaboration of WPA members in actualizing the goals for the WPA articulated here, as I simultaneously pursue my lifelong dedication to the prevention of suicide, hopefully with the help of many new colleagues[12].

Danuta Wasserman
WPA President Elect

1. Global Burden of Disease Collaborative Network. GBD results tool. http://ghdx.healthdata.org/gbd-results-tool.
2. Rehm J, Shield KD. Curr Psychiatry Rep 2019; 21:10.
3. Adhanom Ghebreyesus T. World Psychiatry 2020;19:129-30.
4. Stewart DE, Appelbaum PS. World Psychiatry 2020;19:406-7.
5. Wasserman D, Iosue M, Wuestefeld A et al. World Psychiatry 2020;19:294-306.
6. Vigo D, Thornicroft G, Atun R. Lancet 2016; 3:171-8.
7. Stangl AL, Earnshaw VA, Logie CH et al. BMC Med 2019;17:31.
8. World Health Organization. Multisectoral action for mental health. Geneva: World Health Organization, 2019.
9. World Health Organization. Preventing suicide: a global imperative. Geneva: World Health Organization, 2014.
10. United Nations. The 17 goals. New York: United Nations, 2015.
11. National Centre of Suicide Research and Prevention. COVID-19 BIC study to prevent suicide. https://ki.se/en/nasp/covid-19-bic-study-to-prevent-suicide.
12. Wasserman D. Oxford textbook of suicidology and suicide prevention, 2nd ed. Oxford: Oxford University Press, 2021.

DOI:10.1002/wps.20862

WPA Secretariat's work during the COVID-19 pandemic

The COVID-19 pandemic has inexorably entered the life of the largest psychiatric organization in the world, the WPA, which unites under its wing more than 250,000 specialists from 121 countries and represents 145 national societies of psychiatrists[1].

The task that the WPA President and Executive Committee set for me – as the new WPA Secretary General – is to reorganize the work of the Secretariat, based on today's reality, to optimize it, to revise job responsibilities, and to highlight priorities. This includes restructuring the entire communication process; defining development prospects, strategy and coordination of our media resources; and introducing and using new digital technologies, including the operational management of our website (www.wpanet.org)[2,3].

The next priority task is to improve communication with the 18 Zonal Representatives. I already knew many of them from our previous joint work in the WPA, or from other joint projects. My six years of experience in the same position will also help me in this respect. Furthermore, an important and time-consuming mission of the Secretariat is working with Member Societies[4]. Formally, there are 145 of them, but not all are active enough, not all of them pay dues on a regular basis, and with some we have just lost contact. That is why, together with Zonal Representatives, we are going to implement an "inventory" of the information base of the Member Societies. Close communica-

tion with them is one of the foundations of a well-coordinated organization.

The next important task of the WPA Secretariat is the logistic support to the effective work of the Executive Committee[4]. This includes the preparation of materials for meetings; verification and approval of reports; and many other daily and routine activities. As the WPA Secretary General, I am also expected to ensure the effective work of the WPA Council, maintaining constant communication with its chairperson. The responsibilities of the Secretary General also include the leadership of the Accreditation Committee, which starts its active work during the pre-election period.

An important component of my upcoming work will also be the assistance in the work of the WPA Standing Committees[5] on Planning, Publications, Education, Sections, Meetings, Ethics and Review, Finances and Fundraising. I hope that my previous three years of experience in the Committee on Planning will help me in this respect.

At present, all our efforts are focused on the implementation of the WPA Action Plan for 2020-2023[6]. The success of this implementation will largely depend on the well-coordinated interaction of all components of the WPA structure, and its Secretariat is called to help in this.

The key features of the Action Plan are: to promote psychiatry as a medical specialty in clinical, academic and research areas, and to promote public mental health as a guiding principle; to highlight the specific role of psychiatrists in working with other professionals in health, public health, legal and social aspects of care; to ensure WPA's positive engagement with Member Societies and WPA components, mental health professionals and general health care workers.

The six areas of the WPA Action Plan 2020-2023 include: public mental health; child, adolescent and youth mental health; addressing comorbidity in mental health; developing partnerships for collaborative work and strengthening partnerships with mental health and other organizations; capacity building and training in global mental health and public mental health; continuation and completion of previous WPA Action Plans[7,8].

I plan to assist the President in searching for educational grants and funds, and in preparing and monitoring specific proposals for sponsors. I will also assist in finding sites for Regional and Thematic Congresses. I also plan to prepare a report on the importance of a constant collaboration with the mass media, with the aim of exploring the possibility to create a working group, or establish a Section, to develop a charter for relations with the press. The ultimate goal would be to sign such a charter with the International Press Association.

Promoting mental health and increasing the popularity of psychiatry as a medical specialty are among the objectives of the WPA Action Plan. I intend to submit working materials in this area to the Executive Committee. I am also planning regular appearances in the mass media with the aim to promote psychiatry as a medical specialty in clinical, academic and research areas and to promote public mental health as a guiding principle, as well as to disseminate information on WPA activities.

Concerning partnership with other professional mental health organizations, I am going to participate as an observer in the work of the World Health Organization (WHO) General Assembly and Executive Board, to cooperate with the World Medical Association, and – as a member of the Council of the European Psychiatric Association and one of the founders of the European College of Neuropsychopharmacology (ECNP) – to cultivate relationships with these latter associations.

Using my experience in the WHO and the ECNP, I am also planning to contribute to the Action Plan goal of "Promoting evidence-based psychopharmacotherapy". I also expect to continue working in the WPA Expert Group on COVID-19 and the care for people with mental illness.

I am confident that the WPA will successfully overcome the difficulties associated with this critical period and continue its effective work. The WPA Secretariat will make every effort to contribute to this, relying on the full support from all components of the Association.

Petr V. Morozov
WPA Secretary General

1. Herrman H. World Psychiatry 2020;19:404-6.
2. Kallivayalil RA. World Psychiatry 2020;19:407-8.
3. Kallivayalil RA. World Psychiatry 2020;19:124.
4. Kallivayalil RA. World Psychiatry 2019;18:239.
5. Herrman H. World Psychiatry 2019;18:368-9.
6. Javed A. World Psychiatry 2020;19:411-2.
7. Herrman H. World Psychiatry 2019;18:113-4.
8. Bhugra D. World Psychiatry 2014;13:328.

DOI:10.1002/wps.20863

Plan of the WPA Secretary for Scientific Meetings for the triennium 2020-2023

It has been a great honor to be elected as the WPA Secretary for Scientific Meetings and to join the WPA Executive Committee. I am poised to do my utmost to promote the mission of the WPA and look forward to contributing to the achievements and success of the Association, supporting and advocating for the members and further advancing the best clinical care, education and research in psychiatry globally[1,2].

The WPA Secretary for Scientific Meetings is expected to: a) work with the Executive Committee and the Secretariat to oversee and co-ordinate all official scientific meetings of the WPA and manage applications for WPA co-sponsorship of scientific meetings; b) be responsible for the development of proposals to host the World Congresses of Psychiatry and other WPA scientific meetings in accordance with the Association's policies; c) assist in all aspects of the organization of World Congresses and other WPA scientific meetings[3].

The WPA goals for scientific meetings, as delineated in the policy approved by the Executive Committee in May 2019, are

to: a) increase the exchange of information between psychiatrists from different parts of the world, including networking, training and mentoring of early career psychiatrists[4]; b) contribute to the education of different categories of mental health workers by providing up-to-date scientific information[5]; c) increase exchange and collaboration between psychiatrists and their community, professional, government and development partners in all parts of the world; d) boost collaborative research by bringing together psychiatrists and others interested in research from various parts of the world; e) strengthen links among WPA Member Societies and between WPA and international and regional organizations in psychiatry; f) increase the visibility of psychiatry nationally and internationally; g) contribute to WPA funds.

The WPA Standing Committee for Scientific Meetings will continue to implement and improve the tasks and functions of the WPA related to congresses and co-sponsored meetings by: a) improving the scientific quality of these meetings with state-of-the-art presentations; b) working in close collaboration with the WPA Secretary for Education and Secretary General to provide continuing medical education (CME) credits for WPA meetings; c) working in close collaboration with the WPA Secretary for Finances to improve the financial income and stability of the WPA through sponsored events; d) increasing the number of WPA co-sponsored meetings to involve all the four Regions and eighteen Zones of the WPA, reaching high-, middle- and low-income countries; e) disseminating WPA information, knowledge, educational programs and expertise to all the WPA Regions, in coordination with the Association's Scientific Sections[6,7]; f) focusing on evidence-based knowledge by research oriented and educationally oriented presentations; g) addressing the mental health issues during and after the COVID-19 era[8,9].

From the very beginning of the current triennium 2020-2023, the COVID-19 pandemic has disrupted planned medical conferences across the entire world. While facing these uncertain, unpredictable and unprecedented times, the WPA remains sensitive and respectful not only to those affected by COVID-19 but also to our own safety and well-being[10]. Based on close monitoring of the global risk assessment by the World Health Organization regarding the COVID-19 pandemic, a number of our congresses and co-sponsored meetings have since been cancelled, postponed or transformed into innovative online formats.

At present, it is difficult to predict when the pandemic will be mitigated and we will be able to travel safely and resume in-person meetings without jeopardizing our and others' personal health. We will closely monitor the future development of the COVID-19 pandemic and diligently make appropriate and sagacious adjustments in planning for future congresses. The WPA will not succumb to the "pandemic fatigue" and detour its path, but will move forward.

The Association is confident that, by working closely together, we will be able to face and overcome the challenges. While trusting the pandemic shall pass, we now set our sights on 2021 and beyond, begin to review and reflect on the relevant issues of the current year, especially mental health during the COVID-19 pandemic, then strategize how to plan what the future will hold for us and embrace whatever the new norm for the post-COVID-19 era would be.

The very first WPA scientific event of the new triennium was the Thematic Congress on Intersectional Collaboration entitled "Psychological Trauma: Global Burden on Mental and Physical Health", held virtually from 11 to 13 December 2020. The following WPA Congresses that have been held or are confirmed or proposed include: the World Congress of Psychiatry "Psychiatry in a Troubled World", held virtually from 10 to 13 March 2021; the Regional Congress "Interdisciplinary Understanding of Comorbidity in Psychiatry: from Science to Integrated Care", held virtually from 16 to 18 May 2021; the Regional Congress "Psychopathology in Periods of Transition", scheduled in Kyiv, Ukraine, from 7 to 9 July 2021; the World Congress of Psychiatry "New World, New Challenges for Psychiatry and Mental Health", to take place in Cartagena, Colombia, from 18 to 21 October 2021; the World Congress of Psychiatry to be held in Bangkok, Thailand, from 3 to 6 August 2022; and a Thematic Congress scheduled in Moscow, Russia, in October 2022.

The future congresses and meetings may be transformed into virtual or "hybrid" formats until the world recovers from the pandemic, and it becomes safe to travel and to convene face-to face meetings. The WPA has called for each zone to think of a thematic meeting and/or a regional meeting in collaboration with other zones. The latter provides a viable opportunity to get closer to the other zones in the region. There are more WPA congresses and co-sponsored meetings in the pipeline and we will regularly provide updates.

Edmond H. Pi
WPA Secretary for Scientific Meetings

1. Javed A. World Psychiatry 2020;19:411-2.
2. Hermann H. World Psychiatry 2020;19:404-6.
3. Okasha T. World Psychiatry 2013;12:182-3.
4. Pinto da Costa M. World Psychiatry 2020;19:127-8.
5. Ng RMK. World Psychiatry 2020;19:257-8.
6. Schulze TG. World Psychiatry 2020;19:123-4.
7. Bertelli MO, Salvador-Carulla L, Munir KM et al. World Psychiatry 2020;19:260.
8. Adhanom Ghebreyesus T. World Psychiatry 2020;19:129-30.
9. Unützer J, Kimmel RJ, Snowden M. World Psychiatry 2020;19:130-1.
10. Stewart DE, Appelbaum PS. World Psychiatry 2020;19:406-7.

DOI:10.1002/wps.20864

Update from the WPA Secretary for Education

The year 2020 has been most challenging. The COVID-19 pandemic has adversely affected the physical and mental health of virtually everyone in the world. There is an anticipated increased demand for mental health care. However, given the limited mental health capacities around the world, especially in low- and middle-income countries, this will imply that more

people with mental health problems will be deprived of optimal care[1-5].

In the current triennium, one of the items of the WPA Action Plan[6] is to enhance the mental health capacities around the world. Education and training in evidence-based clinical knowledge and skills are essential to achieve this goal. With Internet and smartphone technologies gaining popularity worldwide, the WPA Secretary for Education's work plans have also focused more on Internet-based education.

The WPA has been recently able to purchase a licence for a learning management platform (known as the "LMS"), reorganize and update some of its many existing educational materials, recruit a part-time educational coordinator, and develop several new online educational modules for the LMS with the kind support from experts in various Association's components.

The availability of these modules in the WPA educational portal – accessible on the Association's website (www.wpanet.org) – provides mental health professionals around the globe (especially those in low- and middle-income countries) with new opportunities to learn and to update their psychiatric knowledge and skills. In particular, we have been able to develop modules on several vital topics – some examples being telepsychiatry, intimate partner violence, and basic psychotherapy. Such enhancement of knowledge and skills will hopefully translate into enhanced mental health capacity and improved access to care by patients, especially in underserved populations.

Within days of launching the system, more than 1,000 psychiatrists registered to access our learning materials (more than double our initial technical subscription). More than 3,000 website visitors read WPA news stories related to the educational portal and close to 4,000 viewed the educational portal webpage. The profile of registered users is vast, with close to 50 countries represented from right across the globe.

It has become quickly clear that the WPA educational portal will be vital to our future work and to that of our 145 Member Societies and our mental health partners. We are eager to maintain the momentum created during this launch phase and to build on it to ensure a valuable and sustainable resource well into the future. In the coming year, the WPA would also like to further develop and expand the number of educational materials available, commission new materials on evolving areas and issues, expand the search functionality and language capability of the system, refine the registration process, and diversify the information gathered.

Apart from didactic learning through the modules available in the WPA educational portal, experiential learning is obviously important to enhance the translation of clinical knowledge into actual skills in patient care. In this respect, an important ongoing development is the further rollout of the WPA volunteering programme[7]. Due to travel restrictions related to COVID-19, online volunteering will be the main focus in the year 2021.

Through face-to-face interaction with expert volunteers online, participants from host organizations will benefit from real-time lectures, Q & A sessions, role play of clinical skills, and supervisory feedback on their clinical performance. Given the potential benefits to the volunteers and their hosts, this educational initiative has been upgraded to become the target of one of the WPA Action Plan's work groups.

Comprised of experts on volunteering from different countries, this work group has just prepared guidelines and recommendations on the attributes of expert volunteers, the roles and responsibilities of the volunteering and host organizations, and the advisory and supportive roles played by the WPA. The work group is now planning to conduct a pilot volunteering project with an aim of field-testing and fine-tuning the practical logistics, and evaluating the effectiveness and satisfaction with the project. These exciting developments will soon be available on the website of the WPA.

Roger M.K. Ng
WPA Secretary for Education

1. Adhanom Ghebreyesus T. World Psychiatry 2020; 19:129-30.
2. Stewart DE, Appelbaum PS. World Psychiatry 2020;19:406-7.
3. Wasserman D, Iosue M, Wuestefeld A et al. World Psychiatry 2020;19:294-306.
4. Wang Q, Xu R, Volkow ND. World Psychiatry 2021; 20:124-30.
5. De Hert M, Mazereel V, Detraux J et al. World Psychiatry 2021;20:54-5.
6. Javed A. World Psychiatry 2020;19:411-2.
7. Ng RMK. World Psychiatry 2020;19:257-8.

DOI:10.1002/wps.20865

The WHO's Global Clinical Practice Network and the ICD-11 implementation

The WPA is one of the professional organizations collaborating with the World Health Organization (WHO)'s Global Clinical Practice Network, whose ultimate mission is improving mental health care and services in the various regions of the world (https://gcp.network).

This Network now includes more than 16,000 clinicians from 159 countries (51% psychiatrists, 30% psychologists; 40% from Europe, 25% from Western Pacific, 24% from the Americas, 5% from Southeast Asia, 3% from Eastern Mediterranean, and 3% from Africa; 63% from high-income countries, 37% from middle- and low-income countries).

The Network has been collaborating with the WHO Department of Mental Health and Substance Use in studies informing the development of the chapter on mental and behavioural disorders of the ICD-11. In particular, the Network has been involved in the Internet field trials, which have compared the diagnostic agreement and the clinical utility for several groups of disorders in the ICD-11 vs. the ICD-10.

A Network-based study focusing on the ICD-11 diagnostic guidelines for disorders specifically associated with stress[1] found that the addition of complex post-traumatic stress disorder (PTSD) and prolonged grief disorder to the diagnostic system represented significant developments with

respect to the ICD-10. Clinicians were able to distinguish these disorders from similar conditions and from normality. Their ability to differentiate between PTSD and adjustment disorder also improved with respect to the ICD-10. However, participants had some difficulties in making the distinction between symptoms of re-experiencing the trauma in the present and memories in PTSD. This finding brought to a revision in the diagnostic guidelines. Indeed, a very important feature of this and other ICD-11 field trials is that they were conducted before the finalization of the text of the guidelines, so that they really informed the final step of the development of the text.

Another Network-based study, focusing on the ICD-11 diagnostic guidelines for feeding and eating disorders[2], found that the addition of the new categories of binge eating disorder and avoidant-restrictive food intake disorder (ARFID) significantly improved diagnostic consistency with respect to the ICD-10. Furthermore, for all diagnostic categories, clinicians rated the clinical utility of ICD-11 guidelines (including ease of use, goodness of fit, diagnostic confidence, and clarity) more favourably than the ICD-10. However, the results of the study highlighted the need for, and led to, some revisions in the final version of the diagnostic guidelines, such as adding additional qualifiers related to underweight status in the definition of recovery in anorexia nervosa, and adding a clearer specification that the subjective experience of loss of control over eating and related distress is a diagnostic feature of binge eating, even when the person does not consume an objectively large amount of food.

The Network is now serving as a catalyst for other research collaborations, in addition to contributing to the implementation of the ICD-11.

The chapter on mental and behavioural disorders of the ICD-11 has been adopted unanimously by the 72nd World Health Assembly in Geneva on May 25, 2019, although reporting of health statistics based on the new classification will begin in Member States only on January 1, 2022.

Innovations and changes in this chapter have been presented in detail in a paper published in this journal[3]; the involvement of the WPA in the development of the chapter has been also described in the journal[4-6], and several contentious issues (such as the role of a dimensional approach) debated in the process leading to the finalization of the chapter have been addressed in the journal as well[7-12].

The translation of the ICD-11 chapter in several languages and the training of professionals in its use are now ongoing. Educational courses have been conducted at the 18th and 19th World Congresses of Psychiatry (Mexico City, Mexico, September 27-30, 2018; and Lisbon, Portugal, August 21-24, 2019), and a more comprehensive online 20-hr training course was conducted in relation to the 29th Congress of the European Psychiatric Association in April 2021, with the participation of 150 psychiatrists (selected from almost 500 applicants) representing 78 different countries. This training course was organized by the Naples WHO Collaborating Centre.

At the 19th World Congress of Psychiatry, a plenary session dealt with the implementation of the new classification system, that will involve the interaction of the system with each country's laws, policies, health care organization and information infrastructure. K.M. Pike, from the Columbia University, New York, illustrated the multiple modalities developed for training a vast array of international health professionals. M. Maj, who chaired the session with G. Reed, the coordinator of the process of development of the new system, summarized some lessons that should be learnt from the implementation of previously developed classification systems.

The session emphasized the strong collaboration between the WHO and the WPA in all the steps of the development and testing of the ICD-11 chapter on mental and behavioural disorders, and the long-term partnership that will now be established between the two organizations for the implementation of the diagnostic system.

Giuseppe Piegari
WHO Collaborating Centre for Research and Training in Mental Health, University of Campania L. Vanvitelli, Naples, Italy

1. Keeley JW, Reed GM, Roberts MC et al. Int J Clin Health Psychol 2016;16:109-27.
2. Claudino AM, Pike KM, Hay P et al. BMC Med 2019;17:93.
3. Reed GM, First MB, Kogan CS et al. World Psychiatry 2019;18:3-19.
4. Giallonardo V. World Psychiatry 2019;18:115-6.
5. Pocai B. World Psychiatry 2019;18:371-2.
6. Perris F. World Psychiatry 2020;19:263.
7. van Os J, Guloksuz S, Vijn TW et al. World Psychiatry 2019;18:88-96.
8. Fusar-Poli P, Solmi M, Brondino N et al. World Psychiatry 2019;18:192-207.
9. Forbes MK, Wright AGC, Markon KE et al. World Psychiatry 2019;18:272-3.
10. Kotov R, Jonas KG, Carpenter WT et al. World Psychiatry 2020;19:151-72.
11. Maj M, Stein DJ, Parker G et al. World Psychiatry 2020;19:269-93.
12. Sanislow CA. World Psychiatry 2020;19:311-2.

DOI:10.1002/wps.20866

Acknowledgement
This publication has been partially supported by
an unrestricted educational grant from Lundbeck Italia SpA,
which is hereby gratefully acknowledged.