Global Bioethics

ISSN: 1128-7462 (Print) 1591-7398 (Online) Journal homepage: www.tandfonline.com/journals/rgbe20

Ethical considerations in the clinical application of prediction tools for severe mental disorders: perspectives from child and adolescent psychiatrists

Ivars Neiders , Jekaterina Kalēja , Ilze Mileiko , Inese Poļaka & Signe Mežinska

GLOBAL BIOETHICS
2026, VOL. 37, NO. 1, 2648374
https://doi.org/10.1080/11287462.2026.2648374

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Ethical considerations in the clinical application of prediction tools for severe mental disorders: perspectives from child and adolescent psychiatrists

Ivars Neiders, Jekaterina Kalēja, Ilze Mileiko, Inese Poļaka and Signe Mežinska

Institute of Clinical and Preventive Medicine, University of Latvia, Riga, Latvia

ABSTRACT

The ability to predict the risk of severe mental disorders holds considerable promise for individuals at risk, potentially enabling prevention and early intervention. Although the clinical application of such predictive models in child and adolescent psychiatry remains a future prospect, it is essential to consider their social and ethical implications that their use may entail. This study explores child and adolescent psychiatrists' views on these issues through a cross-sectional online survey distributed to members of the European Society for Child and Adolescent Psychiatry. Of the 81 respondents, the majority identified the most significant benefits of using prediction tools as enabling earlier intervention by healthcare professionals (81.5%), improving the quality of care (77.8%), and helping families enhance their resilience (63%). Participants also expressed concern about potential harms, particularly violations of privacy (74.1%), discrimination (92.6%), and the lack of explainability in Artificial Intelligence algorithms (74.1%). While most participants recognise potential medical benefits in the clinical use of predictive tools, numerous concerns must be addressed before such technologies can be considered viable. These include unresolved ethical challenges, such as risks related to privacy, potential of stigmatisation and discrimination, and algorithmic opacity, as well as limitations of healthcare systems.

ARTICLE HISTORY
Received 29 October 2025
Accepted 16 March 2026

KEYWORDS
Risk prediction; severe mental disorders; child and adolescent psychiatry; prediction models; Artificial Intelligence; mental health ethics

Background

Over the past decade, significant developments have occurred in precision psychiatry (Fusar-Poli et al., 2022). One of the methods envisaged as a part of precision psychiatry are personalised prediction models which "represent a novel approach, whereby statistical and machine learning models detect patterns in big data repositories of sociodemographic, clinical, behavioural, cognitive and biological (e.g. neuroimaging and genetic) information to generate probabilistic, individualised risk estimates of a particular outcome occurring." (Lane & Broome, 2022, p. 172) There are different outcomes that these models can be designed to estimate. One that has attracted a lot of attention and on which we want to focus in this study is the prediction of the risk of developing severe mental disorders such as major depressive disorder, bipolar disorder, and schizophrenia in children and adolescents. There are current attempts to develop prediction models for such purposes based on analyses of large amounts of individual and family data acquired through questionnaires, tests, interviews, and biological data, e.g. for predicting risk of severe mental illness for children whose parents (one or both) suffer from severe mental illness (van Houtum et al., 2024). The ability to predict the risk of developing mental disorders may offer significant benefits to individuals at risk, potentially enabling tailored preventive and early intervention strategies (Colizzi et al., 2020; Raballo et al., 2024). However, despite the progress that has been achieved in recent years, the prediction models are not yet ready for clinical application. Rosen et al. (2021) in their comprehensive external validation of 22 prediction models, conclude that personalised prediction of future transition in the clinical or familial high-risk state is potentially feasible, but "To reach utility within preventive psychiatry, transition models need further rounds of external validation as well as guidelines consolidating model comparability and replicability" (Rosen et al., 2021, p. 487). In addition, they point out the necessity to

CONTACT Signe Mežinska signe.mezinska@lu.lv Institute of Clinical and Preventive Medicine, University of Latvia, Riga, Latvia
Supplemental data for this article can be accessed online at https://doi.org/10.1080/11287462.2026.2648374.

develop an ethical framework for implementation. A similar conclusion is made by Hafeman et al in a study on using bipolar polygenic risk score for person level prediction (Hafeman et al., 2017).

There seems to be a consensus that although the clinical application of such models is still in the future, it is important to consider the ethical and social implications that the use of this technology might raise (Fusar-Poli et al., 2021; Fusar-Poli et al., 2022). What makes this technology novel is its reliance on Artificial Intelligence (AI) and machine learning (ML) algorithms for analysis of large amounts of data, which may raise a unique set of ethical considerations on their own, but also influence how the prediction model is used and perceived by stakeholders, including clinicians. There already exists literature that identifies and discusses some ethical considerations regarding the use of ML algorithm-based prediction models in psychiatry (Fusar-Poli et al., 2022; Koutsouleris et al., 2022; Starke et al., 2021), and although it provides many valuable insights into the issue, the vast majority of it presents theoretical reflections. Therefore, an insight into the stakeholder perspective would be timely. Furthermore, many severe mental disorders (e.g. schizophrenia) have a typical onset in adolescence or young adulthood (Walker, 2002), which means that a considerable number of predictive tests will be conducted when the individuals are minors. This raises important ethical questions, such as informed consent, assent, autonomy, the right not to know, and the psychological impact of communicating risk to children and their families. An additional group of concerns relates to the potential for stigmatisation, especially if predictive information is misunderstood or misused. Labelling a child as "at risk" may inadvertently influence how they are treated or perceived, potentially affecting their self-perception, social relationships, and access to opportunities (Fusar-Poli et al., 2022).

Given that child and adolescent psychiatrists are one of the key stakeholder groups in the possible future implementation of this technology, it is critical to understand their perspective on the clinical use of the novel prediction technology. Therefore, we aim to investigate the perspectives of child and adolescent psychiatrists on the ethical issues surrounding potential future clinical applications of novel prediction tools for assessing the risk of severe mental disorders.

Methods

To achieve the aim of the study, we conducted a cross-sectional survey. In reporting the results, we followed a Consensus-Based Checklist for Reporting of Survey Studies (Sharma et al., 2021).

Participants and data collection

The target population consisted of child and adolescent psychiatrists. The survey was administered online in English using the QuestionPro platform. It was distributed to European Society for Child and Adolescent Psychiatry (ESCAP) members through multiple channels. Participants were recruited via email invitations sent to the ESCAP member organisations, distributed through the society's newsletter, and via posts on ESCAP's official X and LinkedIn accounts. Data collection took place between 13 May and 31 May 2024.

Development of questionnaire

We developed a questionnaire (see Additional material 1) using both bottom-up and top-down approaches. First, we organised a stakeholder dialogue event in January 2024, held online, involving health care professionals and researchers from several European countries. The event focused on discussing hypothetical case scenarios involving clinical applications of prediction models for severe mental disorders. Second, we conducted a scoping review of the literature on ethical and social issues related to risk prediction in severe mental illness (Neiders et al., 2025). Based on insights from these activities, we compiled a questionnaire that consists of a list of 23 claims addressing potential concerns and benefits of using an AI-based prediction tool in clinical practice.

Participants were asked to evaluate these claims using a Likert scale ranging from "strongly agree" to "strongly disagree." The questionnaire was piloted with ten members of the FAMILY project consortium (van Houtum et al., 2024) and ESCAP members prior to distribution. The claims were preceded by a description of the prediction model. We asked the participants to imagine that:

> "[T]here is a novel prediction tool available for use in clinical practice, which would allow a (child and adolescent) psychiatrist to communicate to parents and their offspring the risk of severe mental illness later in the children's

lives. Such a tool would use artificial intelligence (AI). It would be based on analyses of large amounts of individual and family information acquired by researchers through questionnaires, tests, and interviews, and might also include biological data (e.g. genetic information, neuroimaging). It would consider both risk and resilience factors. The tool would be validated using independent samples and would be officially approved for use in clinical practice. Here, we assume that the tool could make predictions at a relatively early age, well before the manifestation of disease. The individual's overall risk, as well as their main risk and resilience factors, would be communicated by the psychiatrist during routine clinical care with the aim of providing individualised prevention and intervention programmes to the individual and their family members."

In the second part of the questionnaire, participants were presented with the following hypothetical case:

"Johann was diagnosed with schizophrenia at twenty years old, and some of his close relatives have also experienced schizophrenia. Johann is now 55, has been married for some time now, with three children. During discussions with his psychiatrist, he opened up about the challenges his family faces. Johann's eldest son, David, has become a successful businessman, with two children. Johann's daughter Emma excels at her studies and currently shows no signs of mental illness. Although she aspires to have children in the future, there is lingering concern that they might inherit Johann's condition. Johann's youngest child, Dorothea, who is thirteen, is navigating the turbulent waters of adolescence. As a rebellious teenager, she occasionally exhibits strong emotional behaviour, struggles with sleep issues, and from time to time experiments with alcohol. Johann expresses deep concern about Dorothea's behaviour and mental well-being, highlighting the need for support and guidance especially in the context of the complexities of his family's mental health dynamics."

Based on this case they were asked whether they would recommend using the prediction tool for the family and to specify for whom they would suggest its use. Participants could also provide short justifications for their responses.

Subsequent questions asked participants to rank the importance of different outcomes of using the prediction tool and to indicate whether they would recommend its use for all or only some of the family's children. Finally, participants were asked about their clinical experience and years of professional practice.

Data analysis

Descriptive statistics were used to summarise the results, using JASP software (JASP - A Fresh Way to Do Statistics, n.d.). Data visualisation was performed using the RStudio packages sjPlot (Lüdecke, 2018) and ggplot2 (Wickham, 2016). Statistical significance of differences in proportions was assessed using binomial tests. Differences between mean values of normally distributed variables (e.g. age) in two groups were evaluated using t-tests. The result was considered statistically significant if $p < 0.05$. Free-text responses were analysed using thematic content analysis. An inductive approach was applied, allowing themes to emerge directly from the data rather than relying on predefined categories. This process enabled the identification of key themes that reflect participants' perspectives (Mayring, 2014).

Ethical approval

The study protocol was reviewed and approved by the Research Ethics Committee for Social Sciences and Humanities of the University of Latvia.

Results

In total, 81 participants ($N = 81$, $\text{Mean}_{\text{age}} = 50.13$, $\text{SD}_{\text{age}} = 12.60$) from 33 countries completed the survey, and 78 (96.3%) of them had clinical experience in psychiatry. Given the distribution method of the survey, calculating an exact response rate is not possible; only approximate estimates can be provided. The ESCAP newsletter is sent to approximately 5000 members. Of these, 142 (2.8%) initiated the survey, and 61 dropped out, resulting in a completion rate of 57%. Consequently, the estimated overall response rate is approximately 1.6%. Anonymous data file and materials are available via the OSF link: https://doi.org/10.17605/OSF.IO/PWRNE.

The first seven claims (see Figure 1) in the survey addressed the potential benefits that the use of the prediction tool may bring.

The majority of participants (81.5%) agreed that prediction would allow early interventions to postpone or prevent the onset of a disease for those who are at risk (Q7). The claim that the use of such a prediction tool would allow healthcare professionals to improve the quality of care (Q4) had a slightly lower level of support – 77.8% of respondents expressed their agreement with the statement. On the other hand, only 25.9% agreed that clinical use of the tool would decrease stigmatisation of people who live with such conditions (Q5). 44.4% of respondents agreed that clinical use of such tools would provide relief to people who are at risk and their families (Q6).

The next eight claims (Q8–Q15) addressed potential harms the use of the prediction tool may cause (see Figure 2).

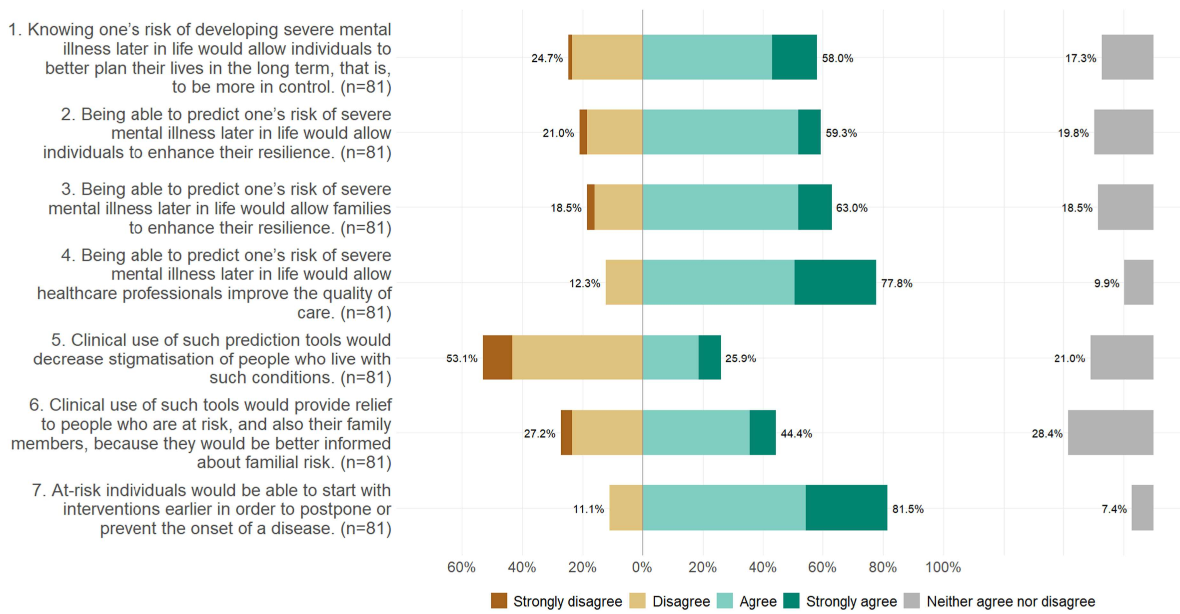

Figure 1. Potential benefits of using prediction tool.

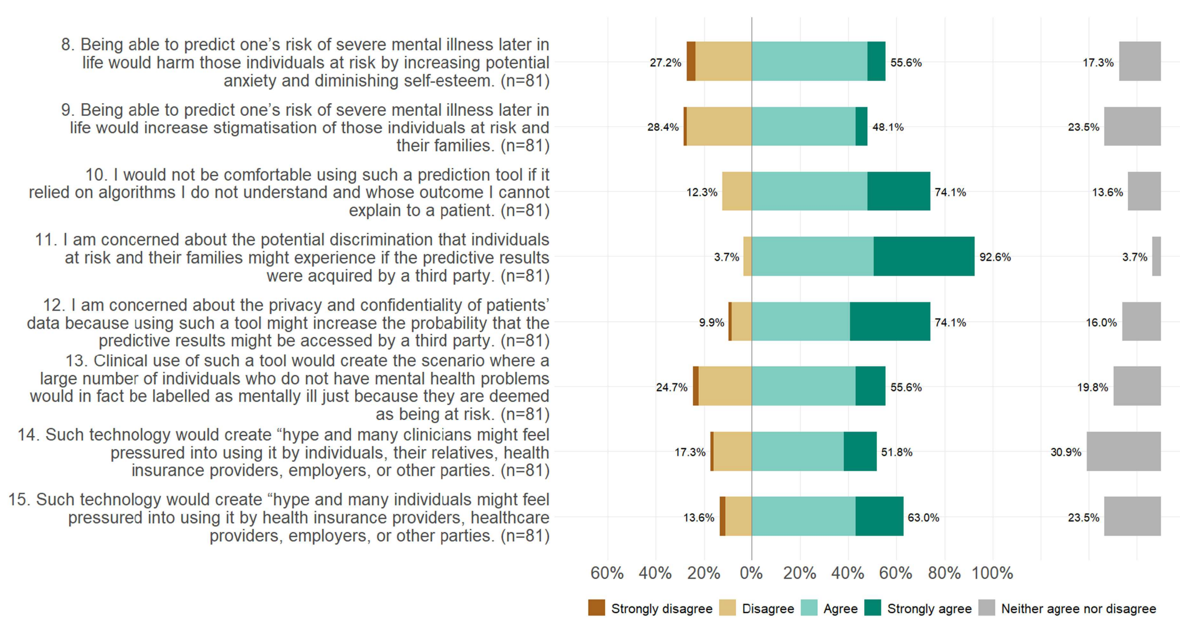

Figure 2. Potential harms of using prediction tool.

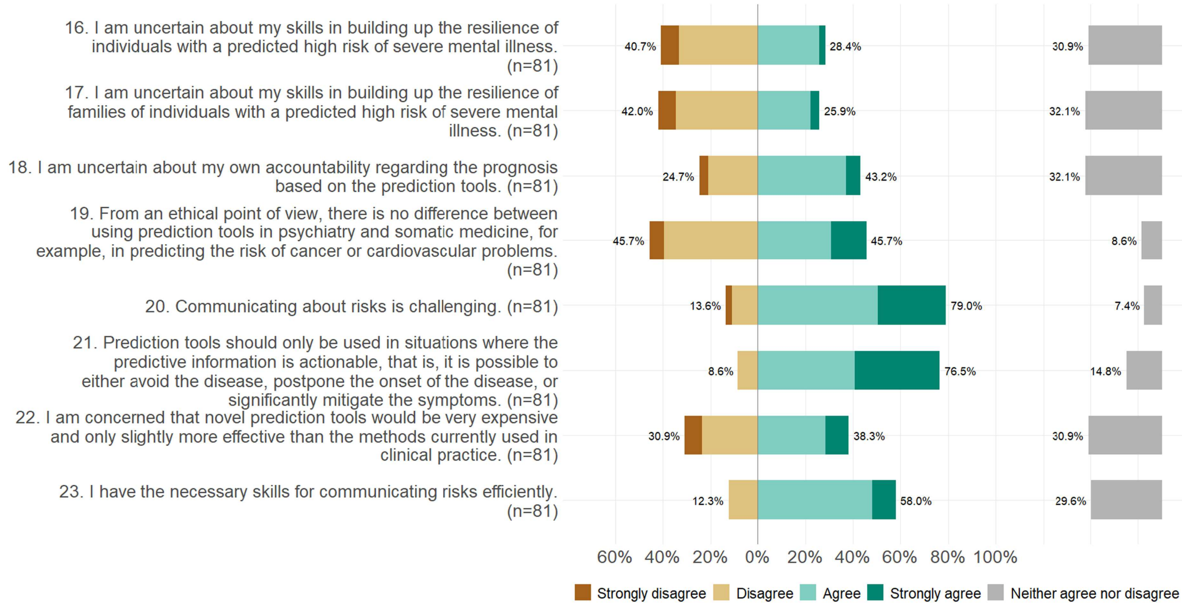

Figure 3. Concerns regarding use of prediction tool in clinical practice.

The biggest concern according to participants was the risk of discrimination –92.6% of the respondents agreed that clinical use of the tool might facilitate discrimination of the individuals at risk and their families (Q11). A slightly lesser proportion of the participants (74.1%) were concerned about the explainability of algorithms used by the tool (Q10) and the risks of violation of privacy and confidentiality (Q12). A total of 55.6% of the participants agreed that prediction of risk might harm individuals who are at risk by increasing anxiety and diminishing their self-esteem (Q8). The same proportion of the respondents expressed their agreement with the statement that using a prediction tool could lead to a significant number of persons who will be labelled as mentally ill just because they are at risk (Q13). Forty-eight percent of the participants agreed that early prediction of severe mental illness would increase stigmatisation of individuals at risk and their families (Q9). It is also worth noting that a considerable proportion of the respondents (23.5%) refused to either agree or disagree with this statement (Q9). A similar trend can be observed in the case of the two claims about the influence of hype created by this technology that might affect the decisions of clinicians (Q14) and individuals (Q15). Specifically, 30.9% and 23.5% chose "neither agree nor disagree" as their answers to those statements, respectively.

The next eight claims (Q16–Q23) addressed various concerns regarding the use of the technology in clinical practice (see Figure 3).

Two claims in this section received approximately equal support –79% of respondents agreed that communication about risks is challenging (Q20), and a slightly lower proportion (76.5%) indicated their agreement with the claim that prediction tools should be used only in cases where the predictive information is actionable (Q21). A total of 58% of the respondents in our survey agreed that they have the necessary skills for effective communication of risks (Q23), although 29.6% opted to "neither agree nor disagree". A total of 45.9% of the participants agreed with the view that ethically there is no difference between using risk prediction in psychiatry and somatic medicine (Q19). A total of 38.3% of the participants agreed that they were concerned that novel prediction tools might be very expensive, with only a slight improvement in efficiency (Q22). Here again, there was a considerably high proportion (30.9%) of those who chose to "neither agree nor disagree" with the statement.

To provide participants with the opportunity to express any additional concerns they might have, we included a free-text response option in the questionnaire. The answers revealed several additional concerns, such as the risk of perceiving the prediction as unavoidable, scepticism about the utility of the prediction tool in professional practice, limitations imposed by an underfunded healthcare system, uncertainty about whether the tool should be used for healthy individuals, commercialisation, data quality, and eugenics (see Box 1 for categories and illustrative examples).

Box 1: **Categories and illustrative quotes from analysis of free-text responses on concerns regarding use of the prediction tool**.

Impact on self-perception:
"Young people and their families can see a predicted risk as inevitable and therefore reduce or fail to take actions that would be protective or mitigate impact in the long term."
Utility of the prediction tool:
"I do not think it is helpful to use a predictive tool. Most of my work is with adolescents who are already aware of some degree of risk for themselves, but this awareness does not increase their capacity to protect."
Limitations of health care systems:
"I think this questionnaire lacks the impact of current health care where even if we would decide that it is a good idea to predict severe mental health problems currently there is to my knowledge not one health care system that can guarantee a good follow up with efficient and sufficient use of clinicians and interventions to offer good care for patients involved."
"A great part of a prevention effort lies outside the health care system, and it is very difficult to get access to relevant interventions from schools, education institutions, public community administration etc."
Scope of use:
"The question is whether you should use these tools in adolescents without any mental health problems."
Quality of data, commercialisation, eugenics:
"The quality of the data used. Commercial interests (selling this tool and make a profit out of it). Political interests (eugenics)."

When answering the questions regarding the case description 45.7% of respondents chose the option to suggest to the family to use the prediction tool, and 54.3% said that they would not. A binomial test showed that there is no statistically significant difference between the groups ($p = 0.505$). Therefore, we cannot say that one answer is preferred over the other.

The participants were asked to justify their answers (see Box 2 for categories and some illustrative examples). Those who said they would recommend using the prediction tool for the family explained their decision by highlighting the benefits of knowledge, prevention and intervention, and the strengthening of resilience.

Box 2: **Illustrative quotes of justifications of the positive and negative answers about the vignette**.

Yes, I would suggest to this family to employ the prediction tool.
Value of knowing:
"It is better to know than not to know."
Benefits of prevention and early intervention:
"Having more in-depth and individualised information would help inform the older son of the risks of his dependency and, in the case of the younger daughter who presents risk behaviours and, perhaps, high-risk mental state, could help to design an intervention plan and give greater weight to the recommendations, in case high risk is confirmed."
"I think it would enhance the factor of prevention of the development of a possible mental illness."
Resilience:
"It would especially help if the resilient factors are also put in view; so you can help the parents to support this resilience and teach them when to call for extra help."
No, I would not suggest to this family to employ the prediction tool.
Lack of justification to use the tool:
"I do not see enough risk factors to legitimise this."
"The discomfort of the youngest daughter can be addressed without predictions, no doubt that resilience, positive parenting and health psychoeducation must be accompanied and worked on. No need for AI for this."
"I do not believe it would significantly change intervention with the family, in terms of psychoeducation and mental health evaluation."
Bad use of the resources:
"The question is for whom you would use this tool. For Emma I think it would be considered as it might lead to an actionable decision: wanting a biological child yes/no. For Dorothea the plan of action in my opinion would not change so I would not recommend it. I think it would be far wiser to invest in good quality care for children and the parent here.
Impact on self-perception:
"The risk for Dorothea to assume a pathological identity ('I'm schizophrenic') at such an important age deeply concerns me."
Medicalisation:
"Risk of turning into disorders the turmoil of adolescence."

While those who said they would not recommend the tool to the family justified their stance by pointing to a lack of compelling reasons, concerns about inefficient use of resources, the possibility that the prediction could negatively affect self-perception, and the risk of medicalisation.

Figure 4 represents word rain visualisation (Skeppstedt et al., 2024) comparing the language used in free-text responses by participants who would not suggest the use of the tool to the family (No) versus those who would (Yes). In both cases, participants were equally talking about risks. However, those who said "no" more frequently mentioned family-related terms, but those who said "yes" emphasised help.

Among those who said they would suggest that the family use the prediction tool, 43.2% said they think the tool should be used for Dorothea only, 35.1% would use it for all the family's children, and 21.6% specified that they would use it only for Dorothea and Emma, which is statistically significantly less than the first two choices ($p < 0.001$).

Those who said that they would suggest that the family use the prediction tool attached the following ranks to three different outcomes (see Figure 5). The results suggest that participants did not express a preference for any particular outcome over the others.

Binomial regression analysis was carried out to determine which questions are the most significant to predict the answer to the question "Would you suggest to this family to employ the prediction tool described above?" The analysis showed that there are five such claims (Q1, Q10, Q2, Q23 and Q18). From the answers to these questions, we were able to predict the answer to "Would you suggest to this family to

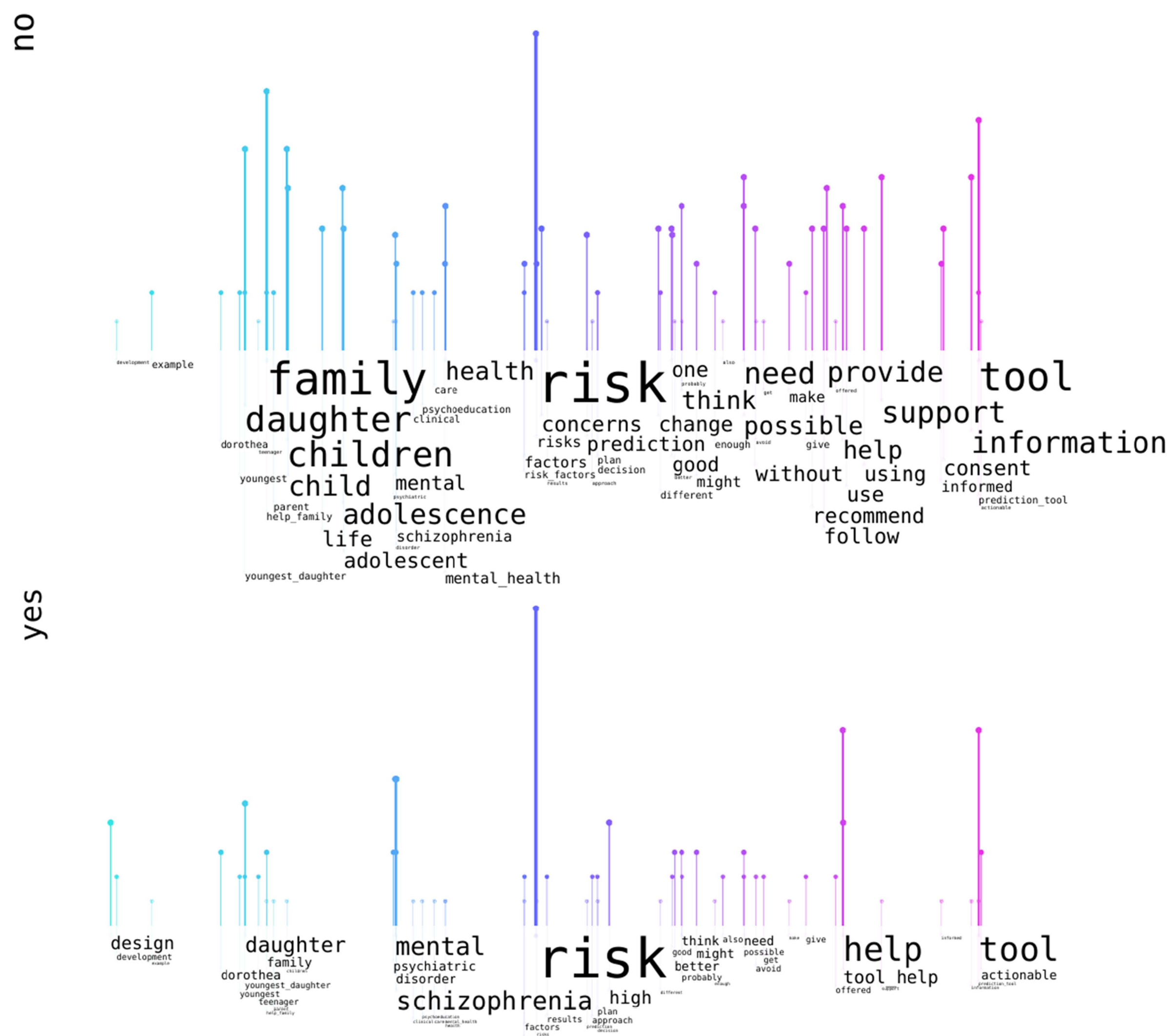

Figure 4. Word rain visualisation of qualitative responses.

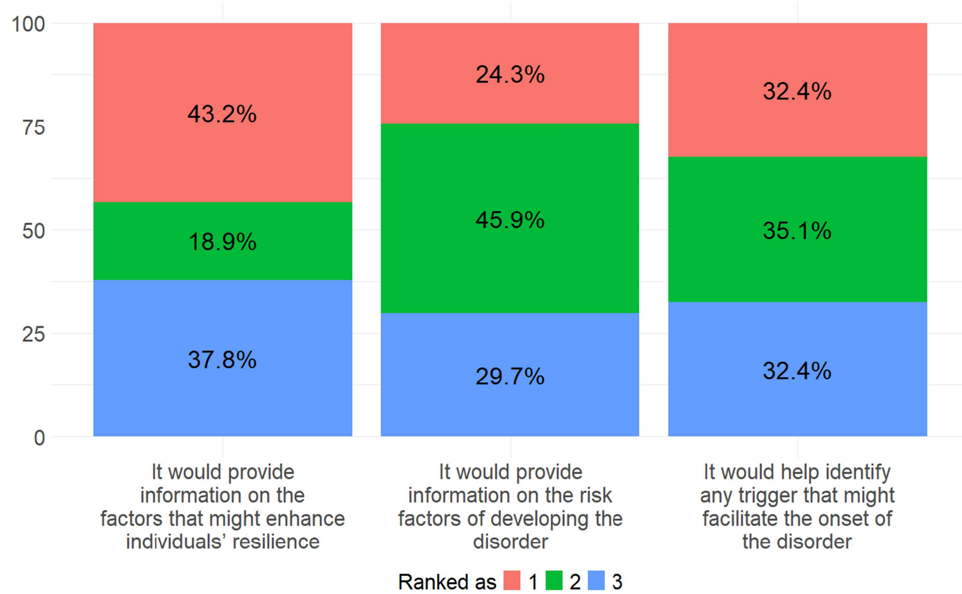

Figure 5. Ranking of three different outcomes that the clinical use of the prediction tool would provide.

employ the prediction tool described above?" with 90% accuracy in this sample. Those who agreed with the claims "Knowing one's risk of developing severe mental illness later in life would allow individuals to better plan their lives in the long term, that is, to be more in control" (Q1) or "Being able to predict one's risk of severe mental illness later in life would allow individuals to enhance their resilience" (Q2) more often answered positively to the question, i.e. they said that they would suggest to this family to use the tool. Those who agreed with "I would not be comfortable using such a prediction tool if it relied on algorithms I do not understand and whose outcome I cannot explain to a patient"(Q10), "I have the necessary skills for communicating risks efficiently" (Q23), "I am uncertain about my own accountability regarding the prognosis based on the prediction tools" (Q18) more often answered negatively.

We also tested whether a participant's age or clinical experience predicts their answers to the survey questions. There were no significant relationships between participants' clinical experience and their answers. In two cases, there was a statistically significant relationship between age and the answers. First, there was a statistically significant difference in age between those who agreed (strongly agree/agree) and those who disagreed (strongly disagree/disagree) with the claim "Clinical use of such tools would provide relief to people who are at risk, and also their family members, because they would be better informed about familial risk." (Q6) Those who disagreed with this claim are younger (45.4 ± 10.4 vs 53.6 ± 12.9) The t-test p-value is 0.0174. Second, there is a marginally statistically significant difference in age between those who agreed (strongly agree/agree) and those who disagreed (strongly disagree/disagree) with the claim "Being able to predict one's risk of severe mental illness later in life would harm those individuals at risk by increasing potential anxiety and diminishing self-esteem."(Q8) Those who disagreed are older (54.3 ± 13.7 vs 47.4 ± 13.0) The t-test p-value is 0.05.

Discussion

To our knowledge, this is the first study to examine child and adolescent psychiatrists' attitudes towards the ethical considerations involved in the potential clinical use of AI-based tools for predicting severe mental disorders in children and adolescents. A study that is closest to the present research is a paper by Pereira et al. on child and adolescent psychiatrists' knowledge, attitudes, and experiences regarding polygenic risk scores (Pereira et al., 2022). Another empirical study that explores a not-too-far territory is Salm et al.'s paper on attitudes on use of genetic tests among neurologists and psychiatrists (Salm et al., 2014). However, both these studies focus only on genetic testing, which, in comparison to the novel prediction tool described in the present study, utilises only genetic information and does not rely on AI. Therefore, these previous studies cannot be generalised to this novel technology.

The psychiatrists in our sample in general, agree that clinical use of a novel prediction tool has the potential to bring several benefits, namely, that it would enable at-risk individuals to start interventions earlier and, therefore, either prevent or postpone the onset of the disease and allow healthcare professionals to improve the quality of care. This, of course, is not surprising as these medical benefits are amongst the central reasons why the tool is being developed in the first place. Regarding non-medical benefits, the participants in our study were slightly less enthusiastic but still generally positive. For example, over half of the participants supported the idea that it would give individuals and families more control over their lives. In general, the psychiatrists in this study expressed views on the clinical benefits of predictive tools that closely align with the perceived benefits individuals associate with psychiatric genetic testing in children (similar to perceived benefits in adults) (Meiser et al., 2020, pp. 282–283).

It is noteworthy that several empirical studies indicate that many people believe that genetic testing might decrease social stigma by validating mental illness as a physical illness (Meiser et al., 2020). The participants in our study, however, are more sceptical about the idea that clinical use of prediction tools might decrease stigmatisation. More than half (53.1%) of respondents disagreed with that claim; nevertheless, 25.9% of participants found the idea plausible. At the same time, about half of the participants (48.1%) agree that being able to predict one's risk of mental illness would increase stigmatisation, and 28.4% disagree with that. Two things appear somewhat unexpected. First, if one looks into literature on the concerns raised by predictive genetic testing in psychiatry, stigmatisation of individuals at risk and their family members appears as one of the most often mentioned harms (Brannan et al., 2019; DeLisi & Bertisch, 2006; Meiser et al., 2005; Ruhrmann et al., 2012; Wilde et al., 2010) and yet respondents in our study seem to be more concerned about other issues, such as opacity of algorithms, discrimination and privacy, than stigmatisation. Second, given that concerns about stigmatisation have so far been highly relevant in the context of psychiatric genetic testing, it may seem puzzling that a substantial proportion of participants (roughly 20%) in our study expressed a neutral view on both claims related to the issue (Q5, Q9). There are several explanations that one can offer for these findings. First, one might speculate that given the emphasis on AI in the survey, the participants were to some extent nudged to pay more attention to AI-related concerns. Second, participants might have felt that they are professionally trained to mitigate stigma effectively and therefore, this did not strike them as an important concern in the context. Finally, an alternative interpretation of these findings – one that does not contradict the previous suggestions but rather complements them – is supported by the literature on stigmatisation. Stigma is a complex phenomenon made up of several elements – blame, perceptions of dangerousness and social distance. When a psychiatric condition is understood via a biogenetic explanation, these elements can be affected in different ways. So Kvaale et al. (Kvaale et al., 2013; Kvaale et al., 2013) provide empirical evidence that that is exactly the case. People who rely on biogenetic explanations of mental disorders are less likely to blame affected individuals for their condition. Still, at the same time, they view them as more dangerous and want to distance themselves from them. Therefore, biogenetic explanations of psychiatric disorders provide what they call "mixed blessings" (Haslam & Kvaale, 2015). This, we think, might explain why some respondents in our study indicate that clinical use of novel prediction tools will decrease stigmatisation, why others hold the opposite view (while the proportion of those is lower than one might expect) and, finally, why a considerable proportion of participants are not sure (as some participants might be torn both ways). How the actual clinical use will affect stigmatisation is, of course, impossible to tell, but there is some empirical evidence that suggests that concerns about increasing some aspects of stigmatisation might be justified (Buchman et al., 2024).

Further, regarding the concerns that are more directly AI-related, the participants in our sample share the concerns that are discussed in the literature. AI-based predictive models rely on large and diverse sets of data, and for this reason, privacy and data safety have been in focus of many authors who write about the topic (Ahmed & Hens, 2022; Fusar-Poli et al., 2022; Gillett & Saunders, 2019; Lane et al., 2020; Manchia et al., 2020). Around 71% of our participants expressed concerns about threats to privacy and confidentiality of their patients (Q12), and even more (92.6%) were concerned that potential leakage might cause discrimination (Q11). Another issue that has been widely discussed in the literature is the opacity of the ML algorithms. Fusar-Poli et al. point out that "high complexity" of so-called "black-box" multimodal clinical prediction models does not allow for "backtracking of the key patterns that produced a specific prediction"

(Fusar-Poli et al., 2022, p. 22). This lack of transparency means that physicians using the tool may be unable to understand the rationale behind its outcomes, which in turn makes it difficult for them to assess and justify the decisions generated by the model (Martinez-Martin et al., 2018, p. 806). Furthermore, some authors have argued that this apparent epistemic problem gives rise to an ethical issue, as it jeopardises physicians' ability to adequately inform patients and undermines patients' capacity to provide informed consent (Lane & Broome, 2022, p. 173). It is, therefore, suggested that the developers of the prediction models should avoid “black-box” systems and instead concentrate on developing “transparent, glass-box” alternatives (Fusar-Poli et al., 2022, p. 22). Not everybody agrees, however, that opacity is such a critical problem. For example, Wang et al. have argued that black-box models are not considerably different from some clinical decisions or treatments in medicine, where physicians do not have complete biological or clinical knowledge: “Clinicians cannot always explain why they arrived at a particular diagnosis. Many effective drugs were in widespread use for decades before their mechanism of action was understood. It is still not clear how electroconvulsive therapy or selective serotonin reuptake inhibitors work” (Wang et al., 2020, p. 59). In this light, the demands for transparent ML algorithms may seem unjustified, as that would mean holding AI to a higher standard than other technologies that are already widely used in clinical practice. On the other hand, one can still argue that in the case of AI the stakes are high and the demands, therefore, are reasonable. Especially, if there is a possibility to achieve comparable results by using transparent alternatives. Nevertheless, it is also important how the issue is perceived by relevant stakeholders. Our study indicates that clinicians are concerned about their inability to understand the model's algorithms, which in turn hinders their capacity to explain its outputs to patients (Q10). There is no way to determine how stable that view may prove to be upon reflection, but it certainly should not be ignored. Moreover, it is important to determine whether potential patients and their families share this concern; however, as far as we can tell, no empirical data is currently available on this question.

In general, the psychiatrists in our survey expressed agreement with the concerns they were asked about (Q8–Q15). More than half of them agreed that the issues mentioned in the survey are relevant concerns (except the claim regarding stigmatisation, Q9). In addition, in their free-text responses, some participants expressed concerns about the quality of the datasets the tool would rely on, as well as the readiness of existing health care systems to provide the necessary resources to make early prediction of mental disorders practically useful. Similarly, the majority of psychiatrists in our study agreed that the predictive tool should be used only in situations where the information is actionable – that is, when it can be used to prevent the disorder, delay its onset, or mitigate its symptoms. When presented with a vignette describing a family that might potentially benefit from the use of a predictive tool, participants were divided on whether the tool should be used in the given case. Likewise, when asked to rank possible outcomes of using the predictive tool, participants did not express a clear preference for any of the options provided. Nevertheless, although opinions on recommending the tool to families were evenly split, the justifications provided in the free-text responses reflected contrasting priorities: those in favour emphasised the value of knowledge, opportunities for prevention, and strengthening resilience, whereas those opposed pointed to insufficient justification, resource constraints, and risks of medicalisation and negative self-perception.

Our study has several limitations. The sample is rather small – only 81 participants completed the survey. Low response rates in surveys of professionals and health care workers are a common problem (Asch et al., 2000; Melnyk et al., 2012; Stedman et al., 2019). According to a meta-analysis conducted by Cho et al. in 2013, response rate has been on a decline for several decades (Cho et al., 2013). This makes the study susceptible to non-response bias. However, Kellerman and Herold have argued that non-response bias in physicians' surveys may not be as crucial as in surveys of the general population because physicians as a group are more homogeneous (Kellerman & Herold, 2001, p. 65). Further, it is worth mentioning that due to the cross-sectional study design, it is impossible to determine causal relationships between the variables. Another important limitation is that the study reflects a single stakeholder perspective, as only psychiatrists were surveyed. The absence of views of other groups such as health care professionals, parents, caregivers, ethicists, or policymakers limits the completeness of the ethical discussion and should be considered a priority for future research.

It is also worth noting that a significant share of participants selected “neither agree nor disagree” when responding to certain claims (Q6, Q14, Q16, Q17, Q18, Q22, Q23). This, of course, might simply indicate that, given the novelty of the issue, participants did not yet hold a definite point of view. However, it should not be ruled out that their uncertainty may, to some extent, reflect a failure on our part to formulate the issue more clearly. It is impossible to tell which of the interpretations is correct. If the latter reading is accurate, it should be considered an important limitation of the study.

Conclusion

This article presents findings from a survey exploring child and adolescent psychiatrists' views on ethical concerns surrounding the potential clinical use of novel prediction tools for severe mental disorders. The majority of respondents agreed that the most significant potential benefits of such tools include enabling healthcare professionals to intervene earlier, improving the quality of care, and helping families build resilience. Participants also identified several potential harms they found particularly concerning, including violations of privacy, discrimination, and the lack of explainability in ML algorithms. In comparison, a notably smaller proportion expressed concern about the risk of stigmatisation. The findings related to stigmatisation suggest that, although it is a widely discussed concern in the context of psychiatric genetic testing, participants in our study appear more attuned to AI-specific risks such as algorithmic opacity and privacy. The mixed and often neutral responses regarding stigma may reflect the complex and ambivalent nature of biogenetic explanations, which can simultaneously reduce blame while reinforcing perceptions of danger and social distance. Furthermore, several respondents raised additional concerns about the underfunding of healthcare systems, suggesting that patients may not receive the necessary care even if predictive tools are available. This worry, in turn, according to the participants, casts doubt on whether the clinical use of such tools—under current conditions—would truly deliver the anticipated benefits.

Acknowledgements

Previous versions of this study were presented at the Nordic-Baltic Network for Philosophy of Medicine in Vilnius, June 9–10, 2025, and at the 35th European Conference on Philosophy of Medicine and Health Care in Manchester, August 13–16, 2025. We would like to thank the audiences at these events for their helpful comments, which significantly improved the paper. We are also grateful to Neeltje van Haren, Lisanne van Houtum, and Greta Miknevičiūtė for their assistance in designing the survey. In addition, we extend our thanks to Karen Schlaegel from ESCAP for her help in reaching participants. Finally, we are very grateful to all the participants for their time; without them, this study would not have been possible.

Author contributions

CRediT: **Ivars Neiders:** Conceptualization, Data curation, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing; **Jekaterina Kalēja:** Conceptualization, Investigation, Methodology, Writing – review & editing; **Ilze Mileiko:** Conceptualization, Investigation, Methodology, Writing – review & editing; **Inese Poļaka:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – review & editing; **Signe Mežinska:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – review & editing.

Funding

This research has received funding from the FAMILY project under the European Union's Horizon Europe research and innovation programme (grant number 101057529), the Swiss State Secretariat for Education, Research and Innovation (SERI) and the UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee.

ORCID

Signe Mežinska 0000-0002-3190-100X

References

Sharma, A., Minh Duc, N. T., Luu Lam Thang, T., Nam, N. H., Ng, S. J., Abbas, K. S., Huy, N. T., Marušić, A., Paul, C. L., Kwok, J., Karbwang, J., de Waure, C., Drummond, F. J., Kizawa, Y., Taal, E., Vermeulen, J., Lee, G. H. M., Gyedu, A., To, K. G., …

Matsui, M. (2021). A consensus-based checklist for reporting of survey studies (CROSS). *Journal of General Internal Medicine*, *36*(10), 3179–3187. https://doi.org/10.1007/s11606-021-06737-1

Ahmed, E., & Hens, K. (2022). Microbiome in precision psychiatry: An overview of the ethical challenges regarding microbiome big data and microbiome-based interventions. *AJOB Neuroscience*, *13*(4), 270–286. https://doi.org/10.1080/21507740.2021.1958096

Asch, S., Connor, S. E., Hamilton, E. G., & Fox, S. A. (2000). Problems in recruiting community-based physicians for health services research. *Journal of General Internal Medicine*, *15*(8), 591–599. https://doi.org/10.1046/j.1525-1497.2000.02329.x

Brannan, C., Foulkes, A. L., & Lázaro-Muñoz, G. (2019). Preventing discrimination based on psychiatric risk biomarkers. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *180*(2), 159–171. https://doi.org/10.1002/ajmg.b.32629

Buchman, D. Z., Imahori, D., Lo, C., Hui, K., Walker, C., Shaw, J., & Davis, K. D. (2024). The influence of using novel predictive technologies on judgments of stigma, empathy, and compassion among healthcare professionals. *AJOB Neuroscience*, *15*(1), 32–45. https://doi.org/10.1080/21507740.2023.2225470

Cho, Y. I., Johnson, T. P., & VanGeest, J. B. (2013). Enhancing surveys of health care professionals: A meta-analysis of techniques to improve response. *Evaluation and the Health Professions*, *36*(3), 382–407. https://doi.org/10.1177/0163278713496425

Colizzi, M., Lasalvia, A., & Ruggeri, M. (2020). Prevention and early intervention in youth mental health: Is it time for a multidisciplinary and trans-diagnostic model for care?*International Journal of Mental Health Systems*, *14*(1), 23. https://doi.org/10.1186/s13033-020-00356-9

DeLisi, L. E., & Bertisch, H. (2006). A preliminary comparison of the hopes of researchers, clinicians, and families for the future ethical use of genetic findings on schizophrenia. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *141B*(1), 110–115. https://doi.org/10.1002/ajmg.b.30249

Fusar-Poli, P., Correll, C. U., Arango, C., Berk, M., Patel, V., & Ioannidis, J. P. A. (2021). Preventive psychiatry: A blueprint for improving the mental health of young people. *World Psychiatry*, *20*(2), 200–221. https://doi.org/10.1002/wps.20869

Fusar-Poli, P., Manchia, M., Koutsouleris, N., Leslie, D., Woopen, C., Calkins, M. E., Dunn, M., Tourneau, C. L., Mannikko, M., Mollema, T., Oliver, D., Rietschel, M., Reininghaus, E. Z., Squassina, A., Valmaggia, L., Kessing, L. V., Vieta, E., Correll, C. U., Arango, C., & Andreassen, O. A. (2022). Ethical considerations for precision psychiatry: A roadmap for research and clinical practice. *European Neuropsychopharmacology*, *63*, 17–34. https://doi.org/10.1016/j.euroneuro.2022.08.001

Gillett, G., & Saunders, K. E. A. (2019). Remote monitoring for understanding mechanisms and prediction in psychiatry. *Current Behavioral Neuroscience Reports*, *6*(2), 51–56. https://doi.org/10.1007/s40473-019-00176-3

Hafeman, D. M., Merranko, J., Goldstein, T. R., Axelson, D., Goldstein, B. I., Monk, K., Hickey, M. B., Sakolsky, D., Diler, R., Iyengar, S., Brent, D. A., Kupfer, D. J., Kattan, M. W., & Birmaher, B. (2017). Assessment of a person-level risk calculator to predict new-onset bipolar spectrum disorder in youth at familial risk. *JAMA Psychiatry*, *74*(8), 841–847. https://doi.org/10.1001/jamapsychiatry.2017.1763

Haslam, N., & Kvaale, E. P. (2015). Biogenetic explanations of mental disorder: The mixed-blessings model. *Current Directions in Psychological Science*, *24*(5), 399–404. https://doi.org/10.1177/0963721415588082

JASP - A Fresh Way to Do Statistics. (n.d.). Retrieved December 24, 2024, from: https://jasp-stats.org/

Kellerman, S. E., & Herold, J. (2001). Physician response to surveys: A review of the literature. *American Journal of Preventive Medicine*, *20*(1), 61–67. https://doi.org/10.1016/S0749-3797(00)00258-0

Koutsouleris, N., Hauser, T. U., Skvortsova, V., & De Choudhury, M. (2022). From promise to practice: Towards the realisation of AI-informed mental health care. *Lancet Digital Health*, *4*(11), e829–e840. https://doi.org/10.1016/s2589-7500(22)00153-4

Kvaale, E. P., Haslam, N., & Gottdiener, W. H. (2013). The 'side effects' of medicalization: A meta-analytic review of how biogenetic explanations affect stigma. *Clinical Psychology Review*, *33*(6), 782–794. https://doi.org/10.1016/j.cpr.2013.06.002

Kvaale, E. P., Gottdiener, W. H., & Haslam, N. (2013). Biogenetic explanations and stigma: A meta-analytic review of associations among laypeople. *Social Science & Medicine*, *96*, 95–103. https://doi.org/10.1016/j.socscimed.2013.07.017

Lane, N., & Broome, M. (2022). Towards personalised predictive psychiatry in clinical practice: An ethical perspective. *The British Journal of Psychiatry*, *220*(4), 172–174. https://doi.org/10.1192/bjp.2022.37

Lane, N. M., Hunter, S. A., & Lawrie, S. M. (2020). The benefit of foresight? An ethical evaluation of predictive testing for psychosis in clinical practice. *NeuroImage: Clinical*, *26*, 102228. https://doi.org/10.1016/j.nicl.2020.102228

Lüdecke, D. (2018). *sjPlot*—Data Visualization for Statistics in Social Science (Version 2.6.2) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.2400856

Manchia, M., Pisanu, C., Squassina, A., & Carpiniello, B. (2020). Challenges and future prospects of precision Medicine in psychiatry. *Pharmacogenomics and Personalized Medicine*, *13*, 127–140. https://doi.org/10.2147/pgpm.S198225

Martinez-Martin, N., Dunn, L. B., & Roberts, L. W. (2018). Is it ethical to use prognostic estimates from machine learning to treat psychosis?*AMA Journal of Ethics*, *20*(9), 804–811. https://doi.org/10.1001/amajethics.2018.804

Mayring, P. (2014). Qualitative content analysis: Theoretical foundation, basic procedures and software solution Klagenfurt. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-395173

Meiser, B., Mitchell, P. B., McGirr, H., Van Herten, M., & Schofield, P. R. (2005). Implications of genetic risk information in families with a high density of bipolar disorder: An exploratory study. *Social Science and Medicine*, *60*(1), 109–118. https://doi.org/10.1016/j.socscimed.2004.04.016

Meiser, B., Guo, X. Y., Putt, S., Fullerton, J. M., Schofield, P. R., Mitchell, P. B., & Yanes, T. (2020). Psychosocial implications of living with familial risk of a psychiatric disorder and attitudes to psychiatric genetic testing: A systematic review of the literature. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, *183*(5), 277–288. https://doi.org/10.1002/ajmg.b.32786

Melnyk, S. A., Page, T. J., Wu, S. J., & Burns, L. A. (2012). Would you mind completing this survey: Assessing the state of survey research in supply chain management. *Journal of Purchasing and Supply Management*, *18*(1), 35–45. https://doi.org/10.1016/j.pursup.2011.12.002

Neiders, I., Mežinska, S., & van Haren, N. E. M. (2025). Ethical and social issues in prediction of risk of severe mental illness: a scoping review and thematic analysis. BMC Psychiatry, 25. https://doi.org/10.1186/s12888-025-06949-3

Pereira, S., Muñoz, K. A., Small, B. J., Soda, T., Torgerson, L. N., Sanchez, C. E., Austin, J., Storch, E. A., & Lázaro-Muñoz, G. (2022). Psychiatric polygenic risk scores: Child and adolescent psychiatrists' knowledge, attitudes, and experiences. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *189*(7–8), 293–302. https://doi.org/10.1002/ajmg.b.32912

Raballo, A., Poletti, M., & Preti, A. (2024). Early detection and intervention through the lens of the neurodevelopmental framework: The salience of developmental years and related services. *The British Journal of Psychiatry*, *225*(6), 529–531. https://doi.org/10.1192/bjp.2024.146

Rosen, M., Betz, L. T., Schultze-Lutter, F., Chisholm, K., Haidl, T. K., Kambeitz-Ilankovic, L., Bertolino, A., Borgwardt, S., Brambilla, P., Lencer, R., Meisenzahl, E., Ruhrmann, S., Salokangas, R. K. R., Upthegrove, R., Wood, S. J., Koutsouleris, N., & Kambeitz, J. (2021). Towards clinical application of prediction models for transition to psychosis: A systematic review and external validation study in the PRONIA sample. *Neuroscience & Biobehavioral Reviews*, *125*, 478–492. https://doi.org/10.1016/j.neubiorev.2021.02.032

Ruhrmann, S., Klosterkotter, J., Bodatsch, M., Nikolaides, A., Julkowski, D., Hilboll, D., & Schultz-Lutter, F. (2012). Chances and risks of predicting psychosis. *European Archives of Psychiatry and Clinical Neuroscience*, *262*, S85–S90. https://doi.org/10.1007/s00406-012-0361-4

Salm, M., Abbate, K., Appelbaum, P., Ottman, R., Chung, W., Marder, K., Leu, C.-S., Alcalay, R., Goldman, J., Curtis, A. M., Leech, C., Taber, K. J., & Klitzman, R. (2014). Use of genetic tests among neurologists and psychiatrists: Knowledge, attitudes, behaviors, and needs for training. *Journal of Genetic Counseling*, *23*(2), 156–163. https://doi.org/10.1007/s10897-013-9624-0

Skeppstedt, M., Ahltorp, M., Kucher, K., & Lindström, M. (2024). From word clouds to word rain: Revisiting the classic word cloud to visualize climate change texts. *Information Visualization*, *23*(3), 217–238. https://doi.org/10.1177/14738716241236188

Starke, G., Clercq, E. D., Borgwardt, S., & Elger, B. S. (2021). Computing schizophrenia: Ethical challenges for machine learning in psychiatry. *Psychological Medicine*, *51*(15), 2515–2521. https://doi.org/10.1017/S0033291720001683

Stedman, R. C., Connelly, N. A., Heberlein, T. A., Decker, D. J., & Allred, S. B. (2019). The end of the (Research) world as we know it? Understanding and coping with declining response rates to mail surveys. *Society & Natural Resources*, *32*(10), 1139–1154. https://doi.org/10.1080/08941920.2019.1587127

van Houtum, L. A. E. M., Baaré, W. F. C., & Beckmann, C. F., et al. (2024). Running in the FAMILY: understanding and predicting the intergenerational transmission of mental illness. European Child & Adolescent Psychiatry, 33, 3885–3898. https://doi.org/10.1007/s00787-024-02423-9

Walker, E. F. (2002). Adolescent neurodevelopment and psychopathology. *Current Directions in Psychological Science*, *11*(1), 24–28. https://doi.org/10.1111/1467-8721.00161

Wang, F., Kaushal, R., & Khullar, D. (2020). Should health care demand interpretable artificial intelligence or accept "Black Box" Medicine?*Annals of Internal Medicine*, *172*(1), 59–60. https://doi.org/10.7326/M19-2548

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* (2nd ed.). Springer. https://doi.org/10.1007/978-3-319-24277-4

Wilde, A., Meiser, B., Mitchell, P. B., & Schofield, P. R. (2010). Public interest in predictive genetic testing, including direct-to-consumer testing, for susceptibility to major depression: Preliminary findings. *European Journal of Human Genetics*, *18*(1), 47–51. https://doi.org/10.1038/ejhg.2009.138