



OPEN ACCESS

EDITED AND REVIEWED BY
Arch Mainous,
University of Florida, United States

*CORRESPONDENCE
Urška Smrke
✉ urska.smrke@um.si

†These authors have contributed equally to this work

RECEIVED 29 January 2026
REVISED 29 January 2026
ACCEPTED 30 January 2026
PUBLISHED 13 February 2026

CITATION

Smrke U, Klén R, Mlakar I, Mulej Bratec S and Levkovich I (2026) Editorial: AI with insight: explainable approaches to mental health screening and diagnostic tools in healthcare.
Front. Med. 13:1798999.
doi: 10.3389/fmed.2026.1798999

COPYRIGHT

© 2026 Smrke, Klén, Mlakar, Mulej Bratec and Levkovich. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: AI with insight: explainable approaches to mental health screening and diagnostic tools in healthcare

Urška Smrke¹, Riku Klén^{2*}, Izidor Mlakar¹, Satja Mulej Bratec^{3,4†} and Inbar Levkovich^{5†}

¹Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia, ²Turku PET Centre, University of Turku and Turku University Hospital, Turku, Finland, ³Department of Psychology, Faculty of Arts, University of Maribor, Maribor, Slovenia, ⁴Faculty of Medicine, University of Maribor, Maribor, Slovenia, ⁵Faculty of Education, Tel Hai College, Upper Galilee, Israel

KEYWORDS

artificial intelligence, depression, diagnosing, explainable artificial intelligence, mental health, monitoring, screening

Editorial on the Research Topic

[AI with insight: explainable approaches to mental health screening and diagnostic tools in healthcare](#)

Mental health conditions are often underrecognized and overlooked in healthcare, especially in primary healthcare (1). Primary care physicians frequently encounter patients presenting with mental health issues, either as standalone conditions or comorbidities to physical illnesses, placing them in a pivotal role for early detection and management. Screening for psychological disorders has been recognized as of great importance (2), however, several barriers to its implementation persist, including lack of staff, time constraints, accessibility restrictions and availability of services (3). Additionally, quality of self-reporting of internal states can be diminished due to reconstruction of memories, inattentive responding, social desirability, and reliance on cognitive heuristics (4–7).

With advancements in big data, artificial intelligence (AI) and technology, exciting new possibilities are arising in the screening for and supporting diagnosing of mental health conditions (8–11). In primary care, explainable artificial intelligence (XAI) offers solutions that can streamline mental health evaluations by providing accessible and interpretable tools that align with physicians' workflows (12, 13). Such approaches have the potential to overcome certain barriers to identifying persons with mental health disorders, as they can be cost-effective, efficient, and can provide an unintrusive assessment that can also support regular monitoring of persons at risk (14–16).

Especially promising are explainable, evidence-based approaches that translate existing knowledge into technology-supported screening and diagnosis of mental health conditions (12, 16). XAI tools can empower primary care providers by demystifying complex algorithms and enabling them to make informed decisions based on AI-generated insights (17). Additionally, these tools can facilitate interdisciplinary collaboration between primary care physicians and mental health specialists, bridging the gap in holistic care (10, 14).

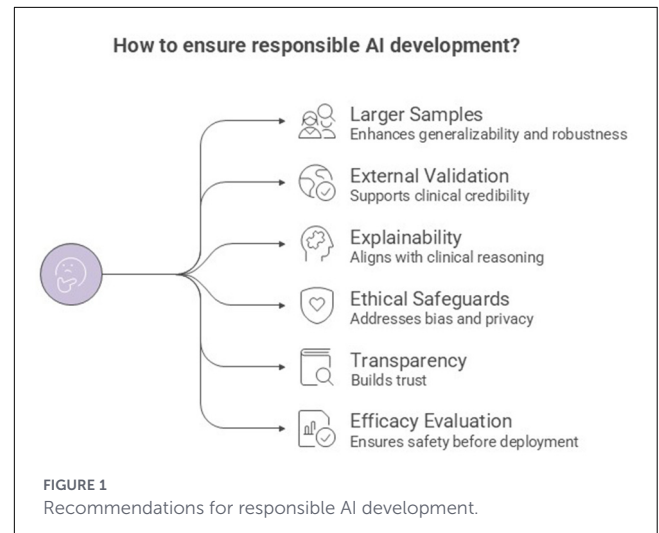
This Research Topic brings together 11 multidisciplinary contributions that explore XAI in the realm of mental health and healthcare. Next, we briefly present each article and highlight how they together move the field from a “black-box” promise toward theoretically and clinically grounded, interpretable tools.

Several contributions in this Research Topic demonstrate how Generative AI (GenAI) and XAI can augment suicide prevention and AI-mediated support. Two studies evaluate a GenAI driven Question, Persuade, and Refer (QPR) suicide prevention simulation training aimed at mental health gatekeepers. [Haber et al.](#) provide support for its reliability, unbiasedness and ability to provide nuanced in-depth feedback to those in training, validating this approach as a valuable tool for advancing complex crisis intervention skills. In another study, [Levkovich et al.](#) evaluate the same simulation training from the perspective of trainees, finding large gains in their self-efficacy after training and generally favorable attitudes toward it. Within the suicide prevention domain, [Grimland et al.](#) focus on identifying suicide risk patterns via explainable Natural Language Processing (NLP) model. The study revealed several theoretical insights and highlighted the potential of AI-driven tools to support crisis counselors in real-time triage.

A second cluster focuses on interpretable models for individual-level screening of mental health disorders. [Mekulu et al.](#) presented a transparent four-feature speech model for depression screening, trained on brief conversational segments. Their model, optimized for deployment in resource-constrained settings, showed moderate discriminative performance and high sensitivity, and the article itself provides many insights on how semantic content can be used in screening, providing interpretable and clinically relevant information. [Nozaki et al.](#) extended explainable behavioral screening into the cognitive domain by applying machine learning to the self-administered rapid task to identify individuals presenting cognitive decline. Their approach suggests a non-invasive, resource efficient strategy in the field. In the cognitive domain, two articles focused on screening and classifying Alzheimer’s disease, both harnessing imaging data collection techniques. [Reddy et al.](#) presented self-attention-based vision used to predict Alzheimer’s disease via magnetic resonance imaging (MRI) images with promising results, presenting an important step toward robust and clinically applicable models in this area. [Slimi et al.](#) present a hybrid convolutional neural network-spiking neural network aimed at classifying Alzheimer’s disease stages. This study suggests possibilities to improve early detection of Alzheimer’s disease in a computationally efficient and biologically relevant approach.

The third topic explored AI approaches beyond detection and classification and provides a critical perspective on the MIT-OpenAI RCT study that explored AI-supported therapy. [Ophir et al.](#) provide a broad perspective on the evaluation of early evidence in this domain and advise on several points in the rapidly developing area of AI-supported therapy.

Finally, three articles provide conceptual and methodological scaffolding for XAI in explainable mental health. [Taskynbayeva and Gutoreva](#) present a systematic review of anxiety-prediction machine learning (ML) models synthesizing the evidence from 19 studies. Findings offer support to the effectiveness of ML for



early anxiety detection. Next, [Močnik et al.](#) present a review of 24 review articles that focused on multimodal cues of mood, anxiety or borderline personality disorders, and provides a valuable synthesis of observable speech, language, facial, physiological, and digital-behavioral markers of these conditions, positioning them as promising for exploitation in developing XAI algorithms for early detection and monitoring of said conditions. Finally, [Yang et al.](#) propose a novel XAI framework for economic mental-health time-series forecasting. This promising framework illustrates how XAI can support not only individual-level screening but also policy-relevant monitoring of mental health trends.

Across the contributions to this Research Topic, several priorities emerge for advancing XAI approaches in mental health care (Figure 1). There is a strong consensus on the need to move beyond small and homogeneous samples toward larger, richer datasets to improve the generalizability and robustness of the models. Future research should prioritize external validation, and tightly align explainability with clinical reasoning, to transfer high accuracy rates into clinically trustworthy deployment, while not forgetting crucial ethical aspects, such as bias, cultural blind spots and erosion of human connections.

The articles demonstrate critical progress in embedding transparency, interpretability, and clinical grounding within AI systems for screening and diagnosing mental health. Across suicide prevention, depression and cognitive screening, dementia imaging, anxiety prediction, and population-level forecasting, they demonstrate that model transparency and performance can coexist. As such, the XAI approaches have a promising potential to build on these foundations and substantially support healthcare services to detect problems earlier, personalize support, and substantially improve access and quality of care.

Author contributions

US: Writing – review & editing, Writing – original draft. RK: Writing – review & editing. IM: Writing – review & editing. SM: Writing – review & editing. IL: Writing – review & editing.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author IL declared that they were an editorial board member of Frontiers at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. During the preparation of this work

the author(s) used Anara in order to improve readability of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take full responsibility for the content of the published article. Figure 1 was produced using Napkin (<https://napkin.ai>) as a data-visualization tool.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Caruso R, Nanni MG, Riba M, Sabato S, Mitchell AJ, Croce E, et al. Depressive spectrum disorders in cancer: prevalence, risk factors and screening for depression: a critical review. *Acta Oncol.* (2017) 56:146–55. doi: 10.1080/0284186X.2016.1266090
- Knies AK, Jutagir DR, Ercolano E, Pasacreta N, Lazenby M, McCorkle R. Barriers and facilitators to implementing the commission on cancer's distress screening program standard. *Palliat Support Care.* (2019) 17:253–61. doi: 10.1017/S1478951518000378
- Granek L, Nakash O, Ariad S, Shapira S, Ben-David M. Oncologists' identification of mental health distress in cancer patients: strategies and barriers. *Eur J Cancer Care.* (2018) 27:e12835. doi: 10.1111/ecc.12835
- Sato H, Kawahara J. Selective bias in retrospective self-reports of negative mood states. *Anxiety Stress Coping.* (2011) 24:359–67. doi: 10.1080/10615806.2010.543132
- Maniaci MR, Rogge RD. Caring about carelessness: participant inattention and its effects on research. *J Res Pers.* (2014) 48:61–83. doi: 10.1016/j.jrp.2013.09.008
- Nahum M, Van Vleet TM, Sohal VS, Mirzabekov JJ, Rao VR, Wallace DL, et al. Immediate mood scaler: tracking symptoms of depression and anxiety using a novel mobile mood scale. *JMIR Mhealth Uhealth.* (2017) 5:e44. doi: 10.2196/mhealth.6544
- Robinson MA, Boies K. On the quest for quality self-report data: HEXACO and indicators of careless responding. *Can J Behav Sci.* (2021) 53:377–80. doi: 10.1037/cbs0000251
- Park Y, Park S, Lee M. Effectiveness of artificial intelligence in detecting and managing depressive disorders: systematic review. *J Affect Disord.* (2024) 361:445–56. doi: 10.1016/j.jad.2024.06.035
- Cruz-Gonzalez P, He AW, Lam EP, Ng IM Li MW, Hou R, et al. Artificial intelligence in mental health care: a systematic review of diagnosis, monitoring, and intervention applications. *Psychol Med.* (2025) 55:e18. doi: 10.1017/S0033291724003295
- Ni Y, Jia F. A scoping review of AI-driven digital interventions in mental health care: mapping applications across screening, support, monitoring, prevention, and clinical education. *Healthcare (Basel).* (2025) 13:1205. doi: 10.3390/healthcare13101205
- Levkovich I. Is artificial intelligence the next co-pilot for primary care in diagnosing and recommending treatments for depression? *Med Sci.* (2025) 13:8. doi: 10.3390/medsci13010008
- Kerz E, Zanwar S, Qiao Y, Wiechmann D. Toward explainable AI (XAI) for mental health detection based on language behavior. *Front Psychiatry.* (2023) 14:1219479. doi: 10.3389/fpsy.2023.1219479
- Muhammad D, Bendechache M. Unveiling the black box: a systematic review of Explainable Artificial Intelligence in medical image analysis. *Comput Struct Biotechnol J.* (2024) 24:542–60. doi: 10.1016/j.csbj.2024.08.005
- Wolitzky-Taylor K, LeBeau R, Arnaudova I, Barnes-Horowitz N, Gong-Guy E, Fears S, et al. A novel and integrated digitally supported system of care for depression and anxiety: findings from an open trial. *JMIR Ment Health.* (2023) 10:e46200. doi: 10.2196/46200
- Zhang Y, Stewart C, Ranjan Y, Conde P, Sankesara H, Rashid Z, et al. Large-scale digital phenotyping: Identifying depression and anxiety indicators in a general UK population with over 10,000 participants. *J Affect Disord.* (2025) 375:412–22. doi: 10.1016/j.jad.2025.01.124
- Goh YS, See QR, Vongsirimas N, Klanin-Yobas P. Artificial intelligence in diagnosing depression through behavioural cues: a diagnostic accuracy systematic review and meta-analysis. *J Clin Nurs.* (2025) 2025:17694. doi: 10.1111/jocn.17694
- Agur Cohen D, Heymann AD, Levkovich I. Partners in practice: primary care physicians define the role of artificial intelligence. *Healthcare.* (2025) 13:1972. doi: 10.3390/healthcare13161972