ANNALES ACADEMIAE MEDICAE SILESIENSIS

ACCEPTED MANUSCRIPT

Ann. Acad. Med. Siles. (Online) 2026; DOI: 10.18794/aams/219514

Review

The use of artificial intelligence in suicide prevention: opportunities, limitations, and clinical implications – A narrative review

Dorota Turska-Czyż[1], Błażej Pilarski[2], Robert Pudlo[1]

[1]Department of Psychoprophylaxis, Faculty of Medical Sciences in Zabrze, Medical University of Silesia, Katowice, Poland
[2]Doctoral School, Department of Psychoprophylaxis, Faculty of Medical Sciences in Zabrze, Medical University of Silesia, Katowice, Poland

Address for correspondence:
mgr Dorota Turska-Czyż
Zakład Psychoprofilaktyki
Wydział Nauk Medycznych w Zabrzu ŚUM
ul. Pyskowicka 49, 42-612 Tarnowskie Góry
tel. +48 32 285 43 58
e-mail: dturska@sum.edu.pl

Received: 29.01.2026, Revised: 03.03.2026, Accepted: 16.03.2026, Published: May 2026

Publisher: Medical University of Silesia, Katowice, Poland

ABSTRACT

Introduction: The aim of this study is to review scientific research on the application of artificial intelligence (AI) in crisis interventions related to suicidal risk. In recent years, there has been a growing interest in using AI technologies to support mental health care and suicide prevention.

Methodology: The analysis was based on a literature review of studies available in the PubMed (n = 64) and Cochrane (n = 0) databases. The analysis encompassed the period spanning 2022 to the end of 2025 with the use of predefined keywords, including the terms “artificial intelligence” and “suicide crisis”. Redundant publications, along with articles that failed to satisfy the predefined thematic inclusion criteria, were systematically excluded (for example: review articles, expert opinions, and information on ongoing calls for research funding, conference reports). A total of 37 articles meeting the inclusion criteria were identified.

Results: The reviewed publications highlight the potential of AI-based tools in early detection of suicidal risk, assessment, and provision of psychological support. AI models can assist professionals in monitoring, risk assessment, and delivering immediate support to individuals in crisis. However, studies emphasize limitations regarding algorithm reliability and ethical concerns, particularly the possibility of misinterpretation of sensitive data.

Conclusions: Findings indicate the growing potential of AI in suicide prevention. Further research is necessary to ensure the effectiveness, reliability, and ethical use of AI technologies in supporting people experiencing suicidal crises.

KEYWORDS

artificial intelligence, crisis intervention, suicidal risk, suicide prevention

INTRODUCTION

Global and national burden of suicide

Suicide remains one of the most serious public health challenges worldwide. From both a clinical and epidemiological perspective, predicting suicidal behaviour with sufficient accuracy and consistency to enable effective prevention continues to be difficult [1]. The scale of the problem is well illustrated by data from Poland. According to the National Police Headquarters, approximately 13 people in Poland die by suicide every day (11 of whom are men). It should be clearly emphasised that between 2023 and 2024, the most visible decline (over the last 10 years) in the number of suicides in Poland was observed (7.4%), that is, 388 fewer suicides. Statistics collected by the Polish National Police Headquarters show that every suicide is preceded by more

than 10 suicide attempts. It should be emphasized that police statistics on suicides may be underestimated, as a proportion of deaths of this nature may be classified under other causes of death, such as poisoning, cardiorespiratory failure, or traffic accidents. In the case of people under the age of 19, this number increases to over 16 [2,3], and data from Canada (province of Quebec) indicate that for every suicide death there are between 25 and 30 suicide attempts, many of which result in hospitalisation or urgent psychiatric intervention [4]. On a global scale, there are nearly 20 suicide attempts for every suicide death. Recent estimates suggest that each suicide may affect a large number of people within the deceased's social network [5,6]. Furthermore, suicide bereavement has been recognized as a potentially traumatic experience associated with significant psychological distress and adverse mental health outcomes [7,8,9].

Social, clinical, and systemic aspects of suicide

Suicide affects not only the individual who dies but also family members, friends, colleagues, and other people connected to the deceased [10]. Numerous studies indicate that mental disorders, particularly depression, bipolar disorder, schizophrenia, and substance use disorders, are among the most important risk factors associated with suicidal behavior [11,12,13,14].

Suicide prevention constitutes a public health priority of global importance [15]. It is estimated that around 800,000 people take their own lives each year. Despite the largely preventable nature of suicide, factors such as limited healthcare system capacity, varying quality of services provided, and numerous barriers to accessing support significantly hinder progress in reducing suicide rates. Research shows that in more than 80% of cases, shortcomings can be identified in the availability and quality of primary psychiatric care and addiction treatment [4]. In this sense, suicide is not only the result of an individual's suffering but also the outcome of broader programmatic, systemic, and social determinants [16,17,18].

Theoretical and clinical perspectives on suicidal behaviour

Contemporary analyses of the aetiology of suicidal behaviour emphasise that the decision to attempt to take one's own life cannot be reduced to a single dominant factor. It is a multidimensional phenomenon resulting from the interplay of biological, psychological, social, and existential determinants. The literature highlights that a person contemplating suicide is in a deep psychological crisis [19], referred to as a suicidal crisis, which constitutes one of the most complex and burdensome states within the area of mental health [20,21]. A suicidal crisis is an emotional and psychological state in which an individual experiences extreme psychological tension, and the resolution of this crisis, according to the person themselves, seems possible only through committing a suicidal act [22]. It is a specific form of psychological crisis in which intensifying

suffering, a sense of loss of meaning, and a lack of effective coping strategies lead to thoughts of taking one's own life. This phenomenon has the character of a process, encompassing the gradual increase of emotional tension and progressive narrowing of cognitive and emotional functioning [23]. Suicidological research emphasises that its essence is not the desire for death but the need to end unbearable suffering [5,24]. From a clinical perspective, this state is associated with profound helplessness, loss of hope, and an imbalance between psychological resources and environmental demands, which may lead to life-threatening behaviours [25]. Understanding these mechanisms enables more accurate recognition of moments of highest risk and identification of protective factors supporting the patient's return to a sense of control and safety [26,27].

The presuicidal syndrome

In the diagnosis of suicide risk, it is important to make a clear distinction between suicidal thoughts and preparatory behaviours [18,28] since each of these phenomena reflects a different degree of crisis advancement and different implications for clinical intervention and AI-based support systems. The concept of the presuicidal syndrome by Ringel [29] represents one of the key frameworks for understanding the process leading to suicidal behaviour. Ringel [29] described characteristic psychological mechanisms preceding a suicidal act, emphasising that suicide is not a sudden reaction but a prolonged process of accumulating psychological injuries and the gradual weakening of defence mechanisms. In his view, suicidal behaviour is an extreme form of aggression directed against oneself and appears in a state of emotional constriction and increasing tension. The author distinguished three phases of this process, which usually develop sequentially, though they may also coexist. The first phase of the presuicidal syndrome involves the progressive narrowing of psychological functioning, encompassing thinking, emotions, and relationships, leading to rigid, pessimistic perceptions of reality and loss of meaning. In this state, intense negative emotions appear, alternatives in life situations are not seen, previously held values are devalued, and there is a gradual withdrawal from social contact. In the second phase of the presuicidal syndrome, suppressed aggression begins to turn inward, taking the form of intensified psychological or physical self-aggression, which serves as an apparent mechanism of control over increasing tension. The final stage involves suicidal fantasies that bring temporary relief and are accompanied by ambivalence between the desire to live and the imagined end of suffering, indicating advanced psychological disintegration. As Ringel [29] noted, suicidal thoughts constitute the initial phase of the crisis, in which the person considers ending their life without necessarily taking specific actions; they may take the form of general reflections on death or detailed plans and often dominate the person's awareness, leading to an increasing sense of hopelessness and social withdrawal. Preparatory behaviours signify the transition to the phase of actively preparing

to carry out a suicidal intention and include actions such as collecting means (e.g., medication, dangerous objects), writing farewell letters, organising personal or financial matters, and taking other specific steps towards executing the plan [30,31]. The presence of such behaviours is associated with a significant increase in immediate risk and requires urgent, often immediate, therapeutic or crisis intervention [28]. The distinction between thoughts and preparatory behaviours is also crucial for the design of AI-based models, as linguistic and behavioural signals associated with suicidal thoughts may be widespread and not always lead to action, whereas preparatory indicators usually signal a higher probability of crisis escalation [32]. In this context, empirical investigations employing speech analysis within crisis helpline interactions have demonstrated that distinct categories of crisis are associated with markedly divergent patterns of linguistic expression and acoustic characteristics. These differences are particularly salient with respect to emotional intensity and speech organisation, thereby enabling deep learning-based models to reliably discriminate between varying levels of suicide risk and to detect indicators signalling a progression from suicidal ideation to the preparatory phase [33]. The authors proposed a multi-task deep learning framework grounded in a Bidirectional Long Short-Term Memory (BiLSTM) architecture augmented by an attention mechanism, which integrates speech-based acoustic parameters with psychologically salient dimensions. In contrast to generative language models (GenAI/LLMs), which are principally designed to analyse and generate textual content in response to user prompts, the proposed multi-task speech-analytic model functions as a specialised predictive instrument. Its primary objective is the identification of crisis-related signals derived from acoustic features, rather than the facilitation of dialogic interaction. Consequently, support systems should recognise not only the presence of suicidal thoughts but also identify markers of preparatory behaviours and automatically refer such cases for rapid evaluation by clinical staff in accordance with predefined safety protocols [34]. An illustrative example of this approach is the system deployed within the Open Up in Hong Kong, which incorporates a suite of intelligent components – referred to as six “Smart Models” – grounded in artificial intelligence and natural language processing [35]. This system enables consultants to categorise conversational themes dynamically and to continuously evaluate a user’s level of suicide risk (low, moderate, high, or crisis). Crucially, the assessment extends beyond verbal expressions of suicidal ideation to encompass indicators of preparatory behaviour, including the formulation of plans, access to means, functional capacity, and the availability of social support. Risk levels are updated in real time and visually communicated to consultants, who retain the capacity to adjust the assessment and to implement appropriate responses in line with established safeguarding and clinical governance protocols, such as referral to specialised mental health

services. Comparable technological solutions have been adopted within the context of the 988 Suicide & Crisis Lifeline in the United States, where machine learning models are employed to automatically analyse intervention calls with respect to core components of suicide risk assessment [34].

Suicide risk assessment tools

In Polish clinical practice, reliable diagnostic tools are available, including [36]:

Columbia–Suicide Severity Rating Scale (C-SSRS) is one of the most widely used tools for assessing suicide risk and is widely utilised in both clinical practice and research projects. The scale, developed by the team at Columbia University, has been recognised in the United States as a reliable and effective screening instrument, as confirmed by numerous institutional recommendations [30]. Thanks to its high reliability, clear structure, and short administration time, it has found application in many therapeutic settings. The tool enables the assessment of the intensity of suicidal thoughts and behaviours in individuals over 12 years of age and has been approved by the FDA (U.S. Food and Drug Administration) [37] as suitable for monitoring suicidal risk in clinical research. As emphasised by Mokros et al. [38], a shorter screening version is available, whose predictive validity has been empirically confirmed. The scale currently exists in two forms: a short screening version (culturally adapted to Polish conditions) and an extended interview version allowing for an in-depth analysis of suicidal content.

Tool for Assessment of Suicide Risk (TASR), developed by Kutcher and Chehil [39], is a tool designed for the assessment of suicide risk and is also available in a Polish adaptation. In a clinical case-control study, Nowak and Pawełczyk [36] confirmed its diagnostic usefulness in everyday clinical practice.

The Reasons for Living Inventory (RFL), developed by M.M. Linehan, is used to assess psychological resources and the reasons for which a person chooses to live, which constitutes an important component in evaluating protective factors in the context of suicide risk. In Poland, the RFL-48 version for adults has been adapted, retaining the psychometric properties of the original, including high internal reliability. The factorial structure of the Polish adaptation corresponds to the original version and other international validations, encompassing six factors described by analogous test items. The tool also demonstrates satisfactory theoretical validity, confirming its usefulness in clinical assessment. Despite minor limitations, the Polish version of the RFL-48 may be used both in research and in screening therapeutic activities aimed at identifying individual motives that protect against suicide [40].

The Revised Suicidal Behaviours Questionnaire (SBQ-R) is a four-item self-report tool used to assess the severity of suicidal thoughts and behaviours. It was developed as a shortened version of Linehan's earlier, more extensive questionnaire, and its psychometric properties have been confirmed in numerous international studies. The scale also has a Polish adaptation [41].

Technological innovations in suicide prevention

The World Health Organization (WHO) has set a target of reducing the global suicide rate by one-third by 2030 [42]. Furthermore, it recommends the use of modern technologies in preventive efforts [15]. The literature suggests that the application of artificial intelligence (AI) within the domain of suicide crisis intervention encompasses a diverse array of solutions that do not constitute a uniform category with respect to either functional characteristics or potential clinical utility. Conversational artificial intelligence (CAI) represents an expanding class of digital interventions in mental health, referring to systems that employ natural language processing to engage interactively with users. These systems most commonly manifest as commercially available psychotherapeutic chatbots (e.g., Woebot, Wysa), whose responses are formed by specific therapeutic paradigms, such as cognitive behavioural therapy (CBT) or acceptance and commitment therapy (ACT) [43]. Although low-cost, continuously available AI-driven chatbots facilitate rapid and scalable access to psychological support, they are subject to substantial limitations. These include variability in response accuracy, a constrained capacity for nuanced empathic attunement, and the potential for inappropriate handling of highly sensitive personal data. Such limitations raise significant ethical and clinical concerns and underscore the conclusion that these technologies cannot serve as substitutes for direct engagement with a qualified mental health professional. Self-help applications and online mental health platforms constitute a distinct category of digital interventions. They operate as autonomous tools designed to facilitate mental health self-monitoring and the acquisition of coping strategies; however, their clinical impact is constrained by persistently low levels of sustained user engagement and the absence of embedded mechanisms for acute crisis response. Social media platforms, by contrast, may foster peer support and enable the early identification of crisis-related signals, yet they simultaneously entail substantial risks. These include exposure to potentially triggering material, the dissemination of misinformation, and violations of user privacy [44].

One of the technology-based methods is crisis support provided through chat and text messaging. Nakamura et al. [45] indicate that crisis interventions conducted via chat result in positive experiences [46] and beneficial effects on mental health [47] and reduction in suicidal ideation [48]. In turn, crisis support based on text messaging shows a promising impact on reducing suicidal

thoughts within a month after consultation. Nevertheless, a low percentage of users maintain dialogue with a consultant, while the rest either stop responding during the conversation or withdraw completely. The percentage of incomplete conversations remains a serious issue – it amounts to 30% in chat support and 56% in text messaging support [49]. Empirical findings reveal a comparable pattern, with studies indicating that approximately one third of individuals engaging with the Australian Lifeline service expressed opposition to the incorporation of technological solutions and automated processes. Moreover, nearly half of respondents reported that they would be less inclined to seek assistance under such circumstances. These findings suggest that concerns regarding the potential erosion of human contact may exert a substantial influence on individuals' willingness to initiate or sustain engagement with crisis support services [49].

Aim of the review

The aim of this narrative review is to present the current state of knowledge regarding the use of artificial intelligence in suicide prevention and crisis intervention. Special focus is placed on the potential of AI-based technologies in the early detection of suicide risk, monitoring warning signals, and supporting clinical decision-making. The review also discusses the limitations, ethical concerns, and practical implications associated with the use of artificial intelligence in mental health care.

METHODOLOGY

This narrative review was based on a search of the PubMed (n = 64) and Cochrane (n = 0) databases. The search covered publications from 2022 to the end of 2025 and used predefined keywords, including "artificial intelligence" and "suicide crisis". Titles and abstracts were initially screened to assess relevance, followed by full-text evaluation of eligible publications. The selection process was conducted independently by two researchers with expertise in clinical psychology and psychiatry, and disagreements were resolved through discussion. Ultimately, 37 articles meeting the inclusion criteria were included in the final analysis. The review focused on studies addressing the use of artificial intelligence in suicide prevention, suicide risk assessment, crisis intervention, and related ethical and clinical considerations.

ARTIFICIAL INTELLIGENCE IN SUICIDE PREVENTION: CONCEPTUAL FOUNDATIONS AND EMERGING RESEARCH DIRECTIONS

Rationale and conceptual framework

Given these realities, it is becoming increasingly clear that it is necessary to search for new methods of risk identification and support for intervention efforts. In recent years, interest has

been growing in the potential of artificial intelligence (AI), particularly large language models (LLMs), which can support the process of early threat detection, monitoring warning signals, and developing innovative tools for both clinicians and institutions involved in suicide prevention [50].

Pilot studies: Upstream approach using contextual advertising

An example of such initiatives may be a pilot study aimed at presenting the possibility of applying "upstream" approaches in suicide prevention using search engine–based contextual advertising (click-through rate of 4.25%) [51]. Activities were directed towards Internet users searching for keywords related to areas of life characterised by increased suicide risk, such as the perinatal period, domestic violence, addictions, depression, belonging to sexual minorities, and experiencing child abuse. In five problem areas, excluding addictions, the mental health status of target individuals was considered serious, which indicates that information campaigns conducted through contextual advertising can accelerate access to support and increase the effectiveness of directing supportive information to people with severe mental health problems. In the case of addictions, the intervention was effective among people struggling with gambling and alcohol abuse, whereas reaching people addicted to psychoactive substances proved difficult within the adopted methodology. Comparable studies employing machine learning and natural language processing techniques to detect markers of suicidal ideation within social media content – specifically the platform X/Twitter – were conducted by Allam et al. [52]. Their predictive model, grounded in the analysis of linguistic patterns and affective content expressed in publicly available posts, demonstrated robust performance, achieving an overall accuracy of 85%, a precision of 88%, and a sensitivity of 83%. A related application of advanced natural language processing methodologies was reported by Wang et al. [53], who analysed descriptions of circumstances surrounding death drawn from police reports and other sources. Their large-scale study, encompassing over 260,000 suicide cases, illustrated the utility of such approaches for the identification and characterisation of suicidal crises. This methodological paradigm has been further extended to encompass analyses of visual data disseminated via social media platforms. One notable study employed machine learning techniques - specifically the CLIP algorithm (Contrastive Language–Image Pre-training) - to examine photographic content (177,220 images) shared on Facebook by users who concurrently completed a standardised suicide risk assessment instrument. The findings indicated that specific visual characteristics, including the expression of negative affect, diminished indicators of social belonging, and imagery suggestive of isolation, were significantly associated with elevated suicide risk [54]. This trajectory of research is further exemplified by a prospective investigation conducted by Mansoor and Ansari [55], in which the application of advanced artificial intelligence models integrating linguistic analysis with patterns of

social media behaviour facilitated the early detection of mental health crises, including suicidal ideation. Notably, the model achieved an accuracy of 89% and demonstrated an average predictive lead time exceeding seven days relative to expert clinical assessment.

Limitations of traditional assessment and analytical potential of AI

Traditional methods of risk assessment: clinical interview, questionnaires, psychometric assessment, have limitations. They require knowledge of the specific nature of suicidal crisis, are point-based, require the presence of a specialist, and do not allow for continuous monitoring. The development of AI creates an opportunity to overcome these barriers, enabling the analysis of large datasets, early detection of warning signals, and the implementation of scalable forms of support. Within this context, the integration of artificial intelligence and virtual reality (VR) into the training of clinical personnel facilitates the development of immersive and adaptive learning environments that enhance competencies in risk identification and decision-making under crisis conditions, even in settings characterised by limited human and professional resources [56,57]. Empirical research examining the attitudes of mental health professionals suggests that AI-based tools are commonly regarded as valuable early warning mechanisms – “red flags”, with the potential to augment initial patient assessment, inform treatment planning, and optimise referral pathways to appropriate forms of care. These perceived benefits appear particularly salient in the context of systemic pressures and capacity constraints within youth mental health services, notwithstanding persistent concerns relating to financial cost, accountability, and the security of sensitive data [58]. Concurrently, other investigations have demonstrated that contemporary large language models, including GPT-3.5 and GPT-4.5, are capable of processing extensive and richly descriptive clinical documentation generated by healthcare professionals and of identifying multiple, co-occurring indicators of suicide risk. Such models enable the systematic organisation of fragmented clinical information relating to suicidal ideation, preparatory behaviours, prior attempts, or self-harm, thereby supporting more rapid identification of individuals at elevated risk and facilitating the planning of appropriate clinical interventions. Notably, these models permit the transformation of unstructured narrative entries within electronic health records into structured datasets, which may be leveraged for the early detection of high-risk patients, clinical decision support, and the development of evidence-based intervention strategies at both individual and system-wide levels [59]. Nevertheless, evidence also indicates that these models are susceptible to systematic errors, including the overestimation of suicide risk or the misclassification of suicidal ideation as suicide attempts. Such limitations underscore the necessity of expert clinical interpretation and reinforce the imperative that artificial intelligence be

employed as an adjunct to, rather than a substitute for, specialist assessment and professional judgement.

RESULTS AND THE POTENTIAL APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN SUICIDE CRISIS INTERVENTION

The role of artificial intelligence in early detection and crisis support

Research indicates that artificial intelligence (AI) can enhance early detection of warning signs through the analysis of language, crisis conversation content, and online activity. Models based on Natural Language Processing (NLP) and deep learning are capable of identifying subtle linguistic patterns associated with suicidal ideation [1]. Large Language Models (LLMs), such as GPT and Gemini, have been identified as suitable foundations for chatbots designed to interact with individuals in crisis [60]. The literature highlights that such systems have the potential to improve the accessibility and scalability of crisis support, particularly among youth populations and in regions with limited access to mental health professionals [61] as well as in populations particularly vulnerable to inequalities in access to care, such as members of the LGBTQ community, for whom generative conversational tools may provide a complementary source of mental health support, however current evidence indicates a need for further research into their efficacy and safety [62].

Projects such as Vickybot [63] and tools based on Explainable AI (XAI) [4] demonstrate that artificial intelligence can assist both clinicians and public health systems by providing data that facilitates early diagnosis and intervention planning. Similar conclusions were reached by Campbell et al. [60] who compared different chatbot versions across two phases of research: in the first phase, Google Bard/Gemini, Bing Copilot (Microsoft), ChatGPT-3.5 (OpenAI), and Claude (Anthropic); whereas in the second phase, Gemini (Google), Claude 2 (Anthropic), Grok 2 (xAI), Mistral AI, and Meta AI (Meta Platforms). The findings revealed an improvement in the depth and accuracy of responses between the two phases. Responses generated in the second phase were more comprehensive and contextually appropriate, covering a wider range of suicide-related aspects (e.g., warning signs, lethality assessment, available resources, and support strategies). Additionally, these chatbots more consistently emphasised the crisis helpline “988” in the United States. Another illustrative example is provided by the study conducted by Lee et al. [64], which evaluated the capacity of GPT-4 to detect both current and prospective suicidal ideation using data obtained from patients’ initial interactions with a telemental health platform. In certain domains, the AI model demonstrated a level of performance comparable to that of experienced clinicians, including psychologists and psychiatrists, particularly with respect to sensitivity in the

detection of suicidal ideation. These findings are further contextualised by the work of Van Meter et al. [65], which suggests that, notwithstanding the progressively sophisticated and supportive nature of responses generated by generative AI tools, the content produced remains largely conservative and seldom incorporates elements of evidence-based interventions. The authors interpret this tendency as a deliberate strategy to mitigate the risks associated with responding to highly sensitive queries concerning suicide. The objective of the study was to evaluate the quality of generative AI (GenAI) tools' responses to suicide-related queries by analysing the outputs of five systems – ChatGPT 3.5, GPT-4, the health data protection – compliant version of GPT-4, Gemini, and Bing Copilot - across ten distinct prompts. The responses were systematically coded according to seven criteria, encompassing the provision of supportive content, inclusion of crisis helpline information, adherence to evidence-based intervention guidelines, and the presence of potentially harmful material. The analysis revealed that the majority of responses were supportive in nature (79%); however, only 24% incorporated a crisis helpline number, and a mere 4% referenced evidence-based suicide prevention interventions. Instances of harmful content were infrequent (5%) and were observed exclusively within responses generated by Bing Copilot.

AI-based support tools and cultural contexts

AI-based tools such as chatbots, virtual therapists (artificial volunteers – AV), and emotion recognition systems [66], represent innovative approaches to addressing mental health challenges among Indian youth. However, their effectiveness remains variable and strongly context-dependent. Chatbots and virtual therapists provide immediate, relatively low-cost, and anonymous access to psychological support, thereby reducing barriers associated with stigma and therapy costs. Tools such as Wysa and Your Dost have gained significant popularity in India [67] and show promising outcomes, particularly in urban areas, where more than 50% of young users report feeling supported following interaction with AI-based mental health systems [68]. The research conducted within Inuit communities in Canada (2024) underscores that artificial intelligence has the potential to contribute to suicide prevention; however, its effectiveness is contingent upon culturally sensitive adaptation, active collaboration with community leaders, and a thorough consideration of the distinctive sociocultural characteristics of the community in question. Ayer et al. [69] reached analogous conclusions, emphasising that the utility of AI in mitigating suicidal behaviour is realised only when such technologies are specifically tailored to the contextual features of the target environment and are integrated within a broader social and institutional framework. These findings indicate that the application of artificial intelligence in mental health care requires consideration of sociocultural, linguistic, and ethical contexts to support its appropriate and effective use.

Online communication, bereavement, and meaning-making

In the study by Cipolletta et al. [70], conducted among individuals bereaved by suicide (from 48 hours to 10 years post-loss) who used real-time chat services, findings revealed that anonymity and accessibility make such chats a valuable form of initial contact. They enable users to obtain relevant information and begin the process of meaning-making in the aftermath of suicide loss. Participants used chat communication as a safe space in which they could express socially discouraged thoughts and emotions, as well as engage in reconstructing the events and motivations underlying the suicide.

Limitations and risks

In the study by Greaves and Colucci [71], it was demonstrated that artificial intelligence could serve a complementary rather than substitutive role in providing human support, offering a form of AV – a distinct and more accessible channel of assistance for individuals who prefer, or might benefit more from, interactions devoid of direct human contact. The study employed thematic analysis to interpret thirteen semi-structured qualitative interviews conducted with volunteers from a British crisis service. Three main areas of concern were identified: the perceived rigidity and lack of authenticity in AI's responses, and the potential risk of dehumanising users of such chat systems. These concerns align with known factors contributing to suicidal behaviour, such as feelings of rejection and entrapment.

The present study explores the potential of artificial intelligence as an alternative to human interaction in crisis situations, particularly in terms of AI's capacity to function as a conversational partner compared to human volunteers (who are trained in active listening, demonstrating empathy and compassion, risk assessment, safety planning, and de-escalation techniques). The authors argue that AI could be utilised in chat-based support; however, in situations involving crisis escalation, such conversations should always take place under human supervision.

A key limitation remains AI's inability to exhibit genuine empathy and contextual understanding, which are inherently human domains. Research indicates that language models may produce overly general or even hazardous responses – as evidenced by an analysis of ChatGPT conversations, which in certain contexts provided information on means of self-harm [72]. Similar conclusions were drawn by Pichowicz et al. [73], who found that none of the examined agents (ChatGPT, Gemini, Llama, DeepSeek and Lechat) met the basic criteria for an adequate response (as evaluated according to the suicide crisis criteria outlined in the C-SSRS). In total, 51.72% of chatbots met the relaxed criteria for a marginal response, while 48.28% were assessed as entirely inadequate. The most frequent errors included a failure to provide information about emergency

helplines and deficiencies in understanding the user's contextual statements. Comparable limitations were identified in the study conducted by Heston [74], which assessed ChatGPT 3.5's responses to simulated crisis scenarios, including both scenarios modelled on escalating symptomatology as measured by the PHQ-9 (Patient Health Questionnaire) and generalised crisis vignettes. The study found that only two of the evaluated systems provided a referral to a crisis hotline, and that conversations were interrupted with explicit guidance for help only at points corresponding to the highest risk. Parallel observations emerge from a systematic review of randomised controlled trials examining the application of artificial intelligence in suicide prevention [75]. While AI tools have demonstrated potential to enhance short-term assessment, facilitate differentiation of suicide risk levels, support intervention processes, and personalise treatment selection, the overall certainty of the evidence remains limited. This limitation is primarily attributable to small sample sizes and inadequate statistical power to detect rare but clinically critical outcomes, including suicide attempts and completed suicides. Ohu et al. [76] identify a range of irregularities that contravene foundational principles of safe clinical practice, including the ethical tenet of non-maleficence, the establishment and maintenance of a therapeutic alliance, and the application of evidence-based interventions. The authors contend that AI models are inherently deficient in emotional intelligence, contextual attunement, and ethical accountability - capacities that are central to the professional responsibilities of mental health practitioners. Furthermore, they observe that the models' propensity to respond in a conciliatory or excessively non-directive manner – driven by optimisation objectives related to model training rather than clinical efficacy – further undermines their therapeutic utility.

These findings underscore the necessity for robust ethical frameworks, clinical oversight, and rigorous regulation. Another pressing issue concerns the potential for AI to foster unhealthy dependency relationships. Critics highlight those commercial applications such as Replika or Character.AI may exacerbate loneliness by simulating false empathy and generating relationships that become a source of emotional addiction for users. Consequently, particularly vulnerable groups – such as adolescents – may face additional psychological risks [77]. Issues related to data privacy and the burden of surveillance constitute a further dimension of these challenges – as many as 65% of respondents reported fear of personal data misuse. Simultaneously, 72% of adolescents experience sleep disturbances associated with excessive screen exposure, which significantly and negatively affects their mental well-being [68].

SAFETY AND CLINICAL OVERSIGHT

A review of the available the empirical evidence demonstrates a paucity of studies directly addressing safety and clinical oversight in the context of artificial intelligence applications in suicide prevention. The limited body of identified research indicates that AI-based systems should be designed and implemented solely as adjuncts to clinical decision-making rather than as substitutes for professional judgment, with clearly defined escalation mechanisms to qualified specialists in situations of increased risk. The literature further suggests that users of GenAI systems are frequently insufficiently informed about system limitations, data storage practices, and the legal implications of liability, while most existing systems lack formal mechanisms for clinical auditing or structured feedback from licensed professionals. Additionally, these systems may blur the boundary between psychological support and clinical interventions, thereby giving rise to unresolved questions regarding accountability in cases of harm. Accordingly, there is a clear need for the development of regulatory frameworks that delineate the attribution of legal responsibility among technology developers, platform operators, clinicians, and end users [78].

DISCUSSION

An increasing number of studies indicate the usefulness of analysing digital traces in monitoring suicide risk. For instance, research conducted in South Korea revealed a significant association between the frequency of internet searches related to suicide and the actual number of deaths, particularly among women [79]. These findings confirm that the analysis of online behaviour can serve as an early warning tool and a component of suicide surveillance systems. Conclusions drawn by researchers in the field of digital epidemiology suggest that the integration of data from search engines and social media can complement traditional public health monitoring methods; however, this requires both methodological caution and appropriate legal regulation.

An analysis of databases such as PubMed and Google Scholar indicates that the potential of AI in suicide prevention is substantial, though its limitations are equally significant. Machine learning models demonstrate high accuracy in risk classification in retrospective studies [80,81], yet in real-world settings, issues often arise concerning the reliability and interpretability of results. It is important to emphasise the role of psychological factors, such as feelings of hopelessness and rejection which, as shown by Grimland et al. [80] play a key role in suicide risk prediction and can be effectively detected by language-based algorithms. On the other hand, experiences such as the analysis of ChatGPT [72] demonstrate that without appropriate safeguards, AI may not only fail to assist but could in fact cause harm. Furthermore, methods such as Mind Genomics and Explainable AI hold considerable promise, as they enable not only risk prediction but also an understanding of

the psychological mechanisms underlying suicidal behaviours. This allows for the development of more personalised yet transparent interventions.

Finally, while AI-based mental health tools offer potential avenues for support, the lack of sufficient personalisation and limited user trust considerably reduce their effectiveness. This is reflected in the finding that only 20% of young people consider such tools reliable [68,69]. Meanwhile, studies conducted among Inuit communities in Canada (2024) emphasise that AI can support suicide prevention, but only when culturally adapted and developed in collaboration with local leaders, respecting the specific characteristics of each community. In a cross-sectional online survey (participants aged 18–69 years from the database of a Japanese internet research company), approximately 20% of individuals exhibiting symptoms of depression and suicidal ideation reported a willingness to seek help, even if it were provided by artificial intelligence [45]. Conversely, Purtle et al. [82] found that within the American crisis hotline 988, the majority of individuals still preferred telephone communication (70.1%) rather than text or chat. This suggests that, in crisis situations, voice contact with another human being is perceived by many as more authentic and secure.

Findings from studies in digital epidemiology [79] indicate that integrating data from search engines and social media can complement traditional methods of public health monitoring; however, such integration necessitates both methodological caution and appropriate legal frameworks [78]. Holmes et al. [83] stress that maintaining human oversight is crucial for the responsible development and application of generative artificial intelligence-both through human-in-the-loop testing and independent expert validation. The authors suggest that high-quality training data are essential for further model improvement and for minimising the risk of unintended biases. Moreover, regulatory policies addressing ethical issues; particularly those concerning privacy and safety in the use of social media data are indispensable.

Artificial intelligence has the potential to facilitate not only the identification of suicide risk but also the strategic planning of preventive interventions at the population level. By enhancing insights into the multifactorial determinants of suicidal behaviour, such technologies may contribute to an increased sense of security among key stakeholders, including educational personnel and caregivers of children and adolescents [69].

CONCLUSIONS

1. The integration of AI with traditional clinical practice creates the opportunity to develop holistic forms of care, in which modern technology enhances professional competence without replacing the key therapeutic elements of empathy and sensitivity.

2. Socio-demographic differences among individuals at risk of suicide highlight the necessity for precise interventions, which may be supported by AI-based tools designed to improve the identification of high-risk groups.

3. The development of modern technologies must proceed through close collaboration within interdisciplinary teams comprising mental health professionals, suicidologists, programmers, as well as legal and ethical experts, in order to ensure safety, adequacy, and the responsible use of AI tools.

4. Potential applications of AI also include the advancement of educational domains, such as training for individuals serving as crisis intervention workers. Moreover, AI can support clinicians through “clinician copilot” models (AI-assisted diagnostic and therapeutic processes) and contribute to the modernisation of medical education and training for healthcare students.

5. The implementation of AI can help reduce existing barriers to accessing care, particularly among individuals who avoid contact with professionals due to shame, fear of stigmatisation, or organisational difficulties. Furthermore, AI can provide support for groups with elevated or heterogeneous levels of risk.

6. Systems capable of analysing emotional and linguistic signals may assist diagnosticians by providing data on the patient’s current emotional state, thereby enhancing the accuracy of clinical assessments and enabling more personalised intervention planning.

REFERENCES

1. Esmaeilzadeh P. Decoding the cry for help: AI’s emerging role in suicide risk assessment. AI Ethics. 2025;5(5):4645–4679. doi: 10.1007/s43681-025-00758-w.

2. Makara-Studzińska M. Przyczyny prób samobójczych u młodzieży w wieku 14–18 lat. Psychiatria. 2013;10(2):76–83.

3. Malec N. Zachowania samobójcze dzieci i młodzieży w Polsce w latach 2013–2022 na podstawie statystyki policyjnej. In: Hołyst B [ed.]. Zapobieganie samobójstwom. Vol. 3. Zachowania suicydalne dzieci i młodzieży. Difin. Warszawa 2023, p. 37–75.

4. Gholi Zadeh Kharrat F, Gagne C, Lesage A, Gariépy G, Pelletier JF, Brousseau-Paradis C, et al. Explainable artificial intelligence models for predicting risk of suicide using health administrative data in Quebec. PLoS One. 2024;19(4):e0301117. doi: 10.1371/journal.pone.0301117.

5. Brown HL, Selbe SM, Flesaker M, Rosellini AJ, Maple M, Gradus JL, et al. The impact of relationship type and closeness on mental health following suicide loss. Suicide Life Threat Behav. 2024;54(3):479–488. doi: 10.1111/sltb.13063.

6. Ruczaj J. Wpływ doświadczenia samobójczej śmierci bliskiego na dzieci i młodzież. Znaczenie podejmowania działań postwencyjnych. Seminare. Poszukiwania Naukowe. 2025;46(1):111–125. doi: 10.21852/sem.1842.

7. Czabański A. Zachowania samobójcze w rodzinach. Społeczne konsekwencje zjawiska. Teologia i Moralność. 2020;14(2(26)):47–57. doi: 10.14746/TIM.2019.26.2.4.

8. Czabański A. Samobójstwa na torach kolejowych jako przykład zaburzenia bezpieczeństwa publicznego. Zesz Nauk. 2012;24:145–155.

9. Pater D. Trauma i ból po stracie dziecka – implikacje terapeutyczne. Prison Syst Rev. 2024;124:249–265. doi: 10.52694/ThPSR.124.12.

10. Whitebrook J, Lafarge C, Churchyard JS. A suicide bereavement model: based on a meta-ethnography of the experiences of adult suicide loss survivors. Front Public Health. 2025;13:1596961. doi: 10.3389/fpubh.2025.1596961.

11. Barrigon ML, Cegla-Schvartzman F. Sex, Gender, and Suicidal Behavior. Curr Top Behav Neurosci. 2020;46:89–115. doi: 10.1007/7854_2020_165.

12. Macellaro M, Cafaro R, Kelly CM, Ostacher MJ, Dell'Osso B, Lyu J, et al. The influence of depressive and manic symptoms on suicidal ideation in mixed mood states. Int J Bipolar Disord. 2025;13(1):23. doi: 10.1186/s40345-025-00390-x.

13. Athey A, Shaff J, Kahn G, Brodie K, Ryan TC, Sawyer H, et al. Association of substance use with suicide mortality: An updated systematic review and meta-analysis. Drug Alcohol Depend Rep. 2024;14:100310. doi: 10.1016/j.dadr.2024.100310.

14. Ventriglio A, Gentile A, Bonfitto I, Stella E, Mari M, Steardo L, et al. Suicide in the Early Stage of Schizophrenia. Front Psychiatry. 2016;7:116. doi: 10.3389/fpsyt.2016.00116.

15. Preventing suicide: A global imperative. World Health Organization, 17 August 2014 [online] https://www.who.int/publications/i/item/9789241564779 [accessed on 25 may 2026].

16. Hutchinson E, Scott L, Choukas-Bradley S, Silk J. Interpersonal risk factors for suicide in daily life among young people: A review of intensive longitudinal studies. Dev Psychopathol. 2025;37(4):2196–2216. doi: 10.1017/S0954579424001810.

17. Darvishi N, Farhadi M, Poorolajal J. The role of social support in preventing suicidal ideations and behaviors: A systematic review and meta-analysis. J Res Health Sci. 2024;24(2):e00609. doi: 10.34172/jrhs.2024.144.

18. Rezaei Z, Mohammadi S, Aghaei A, Pouragha H, Latifi A, Keshavarz-Mohammadi N. Assessment of risk factors for suicidal behavior: results from the Tehran University of Medical Sciences Employees' Cohort study. Front Public Health. 2023;11:1180250. doi: 10.3389/fpubh.2023.1180250.

19. Baryshnikov I, Isometsä E. Psychological pain and suicidal behavior: A review. Front Psychiatry. 2022;13:981353. doi: 10.3389/fpsyt.2022.981353.

20. Kijanka R, Rak M, Białoń P, Dudek M. Gdy śmierć przychodzi w milczeniu – analiza prób samobójczych w Polsce w latach 2013–2017. Na Ratunek. 2018;4:60–64

21. Galynker I, Bloch-Elkouby S, Cohen LJ. Suicide crisis syndrome: a specific diagnosis to aid suicide prevention. World Psychiatry. 2024;23(3):362–363. doi: 10.1002/wps.21229.

22. Romaszko K. Samobójstwa i zachowania samobójcze jako przedmiot zainteresowań nauk medycznych: definicja, podział i przegląd biomarkerów zachowań samobójczych. Część I. Psychiatryczne i neuroobrazowe markery zachowań samobójczych. Studia Med. / Med Stud. 2020;36(4):316–327. doi: 10.5114/ms.2020.102327.

23. Domańska JM. Teoretyczne aspekty zjawiska samobójstw. Pedagogika. 2017;26(2):187–201. doi: 10.16926/p.2017.26.36.

24. Pompili M. On mental pain and suicide risk in modern psychiatry. Ann Gen Psychiatry. 2024;23(1):6. doi: 10.1186/s12991-024-00490-5.

25. Stepień K. Structure and dynamics of suicidal behaviour in Poland. Intern Sec. 2023;15(2):57–71. doi: 10.5604/01.3001.0054.4558.

26. Lew B, Chistopolskaya K, Osman A, Huen JMY, Abu Talib M, Leung ANM. Meaning in life as a protective factor against suicidal tendencies in Chinese university students. BMC Psychiatry. 2020;20(1):73. doi: 10.1186/s12888-020-02485-4.

27. Voros V, Tenyi T, Nagy A, Fekete S, Osvath P. Crisis concept re-loaded? The recently described suicide-specific syndromes may help to better understand suicidal behavior and assess imminent suicide risk more effectively. Front Psychiatry. 2021;12:598923. doi: 10.3389/fpsyt.2021.598923.

28. Bayliss LT, Christensen S, Lamont-Mills A, du Plessis C. Suicide capability within the ideation-to-action framework: A systematic scoping review. PLoS One. 2022;17(10):e0276070. doi: 10.1371/journal.pone.0276070.

29. Ringel E. The Presuicidal Syndrome. Suicide Life Threat Behav. 1976;6(3):131–149. doi: 10.1111/j.1943-278X.1976.tb00328.x.

30. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry. 2011;168(12):1266–1277. doi: 10.1176/appi.ajp.2011.10111704.

31. Motillon-Toudic C, Walter M, Séguin M, Carrier JD, Berrouiguet S, Lemey C. Social isolation and suicide risk: Literature review and perspectives. Eur Psychiatry. 2022;65(1):e65. doi: 10.1192/j.eurpsy.2022.2320.

32. Mo L, Li H, Zhu T. Exploring the suicide mechanism path of high-suicide-risk adolescents based on Weibo text analysis. Int J Environ Res Public Health. 2022;19(18):11495. doi: 10.3390/ijerph191811495.

33. Ding Z, Zhou Y, Dai AJ, Qian C, Zhong BL, Liu CL, et al. Speech based suicide risk recognition for crisis intervention hotlines using explainable multi-task learning. J Affect Disord. 2025;370:392–400. doi: 10.1016/j.jad.2024.11.022.

34. Imel ZE, Pace B, Pendergraft B, Pruett J, Tanana M, Soma CS, et al. Machine Learning-Based Evaluation of Suicide Risk Assessment in Crisis Counseling Calls. Psychiatr Serv. 2024;75(11):1068–1074. doi: 10.1176/appi.ps.20230648.

35. Yip PSF, Chan WL, Chan CS, He L, Xu Y, Chan E, et al. The Opportunities and Challenges of the First Three Years of Open Up, an Online Text-Based Counselling Service for Youth and Young Adults. Int J Environ Res Public Health. 2021;18(24):13194. doi: 10.3390/ijerph182413194.

36. Nowak MP, Pawełczyk T. Skale oceny ryzyka samobójstwa dorosłych w praktyce psychologa klinicznego i psychiatry: przegląd dostępnych narzędzi. Psychiatr Psychol Klin. 2018;18(2):180–187. doi: 10.15557/PiPK.2018.0021.

37. Gmitrowicz A. Diagnoza zachowań suicydalnych z uwzględnieniem DSM-5™ i procesu samobójczego. Psychiatr Dypl. 2014;05.

38. Mokros Ł, Rawska-Kabacińska A, Świtaj P, Wieczorek Ł, Jabłońska A, Anczewska M. Polska adaptacja i walidacja wersji przesiewowej *Skali Oceny Ryzyka Samobójstwa z Uniwersytetu Columbia* (C-SSRS). Psychiatr Pol. 2025;59(2):281–298. doi: 10.12740/PP/OnlineFirst/174591.

39. Kutcher S, Chehil S. Suicide risk management: a manual for health professionals. 2nd ed. John Wiley & Sons. Chichester 2012.

40. Siewierska J, Chodkiewicz J. Polska adaptacja *Skali Powodów do Życia* (RFL-48) M.M. Linehan i współpracowników. Psychiatr Pol. 2022;56(3):603–621. doi: 10.12740/PP/133355.

41. Chodkiewicz J, Gruszczyńska E. Polska adaptacja Zrewidowanego Kwestionariusza Zachowań Samobójczych A. Osmana i współpracowników. Psychiatr Pol. 2020;54(1):101–111. doi: 10.12740/PP/OnlineFirst/93492.

42. Comprehensive Mental Health Action Plan 2013-2030. World Health Organization, 21 September 2021 [online] https://www.who.int/publications/i/item/9789240031029 [accessed on 26 May 2026].

43. Rahsepar Meadi M, Sillekens T, Metselaar S, van Balkom A, Bernstein J, Batelaan N. Exploring the Ethical Challenges of Conversational AI in Mental Health Care: Scoping Review. JMIR Ment Health. 2025;12:e60432. doi: 10.2196/60432.

44. Sherekar P, Mehta M. Harnessing technology for hope: a systematic review of digital suicide prevention tools. Discov Ment Health. 2025;5(1):101. doi: 10.1007/s44192-025-00245-y.

45. Nakamura D, Sueki H, Ito J. AI or human support for suicide prevention? Examining help-seeking intention in suicidal crisis. [Preprint]. PsyArXiv 2025 Jun 5. doi: 10.31234/osf.io/z5wrv_v1.

46. McClellan SR, Hunt M, Olsho LEW, Dasgupta A, Chowdhury M, Sparks AC. Satisfaction and mental health outcomes associated with a large regional helpline. Community Ment Health J. 2022;58(6):1214–1224. doi: 10.1007/s10597-021-00931-5.

47. Brody C, Star A, Tran J. Chat-based hotlines for health promotion: a systematic review. Mhealth. 2020;6:36. doi: 10.21037/mhealth-2019-di-13.

48. Gould MS, Chowdhury S, Lake AM, Galfalvy H, Kleinman M, Kuchuk M, et al. National Suicide Prevention Lifeline crisis chat interventions: Evaluation of chatters' perceptions of effectiveness. Suicide Life Threat Behav. 2021;51(6):1126–1137. doi: 10.1111/sltb.12795.

49. Ma JS, O'Riordan M, Mazzer K, Batterham PJ, Bradford S, Kõlves K, et al. Consumer Perspectives on the Use of Artificial Intelligence Technology and Automation in Crisis Support Services: Mixed Methods Study. JMIR Hum Factors. 2022;9(3):e34514. doi: 10.2196/34514.

50. Macalli M, Navarro M, Orri M, Tournier M, Thiébaut R, Côté SM, et al. A machine learning approach for predicting suicidal thoughts and behaviours among college students. Sci Rep. 2021;11(1):11363. doi: 10.1038/s41598-021-90728-z.

51. Ono S, Sueki H, Nakahara T, Takahashi A, Koda A, Sakai A, et al. Upstream Approaches for Suicide Prevention Using Search-Based Advertising: A Pilot Study. Jpn Psychol Res. 2025;67(2):208–227. doi: 10.1111/jpr.12582.

52. Allam H, Davison C, Kalota F, Lazaros E, Hua D. AI-Driven Mental Health Surveillance: Identifying Suicidal Ideation Through Machine Learning Techniques. Big Data Cogn Comput. 2025;9(1):16. doi: 10.3390/bdcc9010016.

53. Wang S, Dang Y, Sun Z, Ding Y, Pathak J, Tao C, et al. An NLP approach to identify SDoH-related circumstance and suicide crisis from death investigation narratives. J Am Med Inform Assoc. 2023;30(8):1408–1417. doi: 10.1093/jamia/ocad068.

54. Badian Y, Ophir Y, Tikochinski R, Calderon N, Klomek AB, Fruchter E, et al. Social Media Images Can Predict Suicide Risk Using Interpretable Large Language-Vision Models. J Clin Psychiatry. 2023;85(1):23m14962. doi: 10.4088/JCP.23m14962.

55. Mansoor MA, Ansari KH. Early Detection of Mental Health Crises through Artifical-Intelligence-Powered Social Media Analysis: A Prospective Observational Study. J Pers Med. 2024;14(9):958. doi: 10.3390/jpm14090958.

56. Yıldız E. Immersive Solutions With Real-World Impact: Integrating Artificial Intelligence and Virtual Reality Into Global Suicide Prevention Training for Psychiatric Nurses. Int J Ment Health Nurs. 2025;34(5):e70154. doi: 10.1111/inm.70154.

57. Ching YC, Ho YC. AI Virtual Human-Augmented Game-Based Teaching to Enhance Emotional Intelligence in Nursing Students: Protocol for a Single-Group Pretest-Posttest Action Research Study. JMIR Res Protoc. 2025;14:e80290. doi: 10.2196/80290.

58. McKenna S, Chong MK, Poulsen A, Turner A, Gorban C, Crouse JJ, et al. Weighing Costs and Benefits of Delay and the Acceptance of Two Decision Support Tools in Mental Health Care: Scoping Study Using Quantitative and Qualitative Data. JMIR Hum Factors. 2025;12:e71678. doi: 10.2196/71678.

59. Huang M, Li Z, Hu Y, Wang W, Wen A, Lane S, et al. Multi-Label Classification with Generative AI Models in Healthcare: A Case Study of Suicidality and Risk Factors. [Preprint]. ArXiv 2025 Jul 22;arXiv:2507.17009v1.

60. Campbell LO, Babb K, Lambie GW, Hayes BG. An examination of generative AI response to suicide inquires: content analysis. JMIR Ment Health. 2025;12:e73623. doi: 10.2196/73623.

61. Li X, Chen F, Ma L. Exploring the Potential of Artificial Intelligence in Adolescent Suicide Prevention: Current Applications, Challenges, and Future Directions. Psychiatry. 2024;87(1):7–20. doi: 10.1080/00332747.2023.2291945.

62. Bragazzi NL, Crapanzano A, Converti M, Zerbetto R, Khamisy-Farah R. The Impact of Generative Conversational Artificial Intelligence on the Lesbian, Gay, Bisexual, Transgender, and Queer Community: Scoping Review. J Med Internet Res. 2023;25:e52091. doi: 10.2196/52091.

63. Anmella G, Sanabra M, Primé-Tous M, Segú X, Cavero M, Morilla I, et al. Vickybot, a chatbot for anxiety-depressive symptoms and work-related burnout in primary care and health care professionals: Development, feasibility, and potential effectiveness studies. J Med Internet Res. 2023;25:e43293. doi: 10.2196/43293.

64. Lee C, Mohebbi M, O'Callaghan E, Winsberg M. Large Language Models Versus Expert Clinicians in Crisis Prediction Among Telemental Health Patients: Comparative Study. JMIR Ment Health. 2024;11:e58129. doi: 10.2196/58129.

65. Van Meter AR, Wheaton MG, Cosgrove VE, Andreadis K, Robertson RE. The Goldilocks Zone: Finding the right balance of user and institutional risk for suicide-related generative AI queries. PLOS Digit Health. 2025;4(1):e0000711. doi: 10.1371/journal.pdig.0000711.

66. The Oxford Handbook of Affective Computing. Calvo RA, D'Mello S, Gratch J, Kappas A [ed.]. Oxford University Press. New York 2015.

67. Dey NC. Mental health of adolescents and youth in India: A critical analysis in the era of AI. SSRN 2024. doi: 10.2139/ssrn.5056368.

68. Coeckelbergh M. AI ethics. MIT Press. Cambridge (MA) 2020.

69. Ayer L, Boudreaux B, Paige JW, Holmes P, Blagg TL, Mendon-Plasek SJ. Artificial Intelligence-Based Student Activity Monitoring for Suicide Risk: Considerations for K-12 Schools, Caregivers, Government, and Technology Developers. Rand Health Q. 2024;11(2):2.

70. Cipolletta S, Entilli L, Bettio F, De Leo D. Live-chat support for people bereaved by suicide. Crisis. 2022;43(2):98–104. doi: 10.1027/0227-5910/a000759.

71. Greaves J, Colucci E. Crisis-line workers' perspectives on AI in suicide prevention: a qualitative exploration of risk and opportunity. BMC Public Health. 2025;25(1):2229. doi: 10.1186/s12889-025-23298-8.

72. Waszak PM. Chat GPT and suicide prevention – can it work? A conversation analysis. Psychiatr Psychol Klin. 2024;24(4):292–299. doi: 10.15557/PiPK.2024.0036.

73. Pichowicz W, Kotas M, Piotrowski P. Performance of mental health chatbot agents in detecting and managing suicidal ideation. Sci Rep. 2025;15(1):31652. doi: 10.1038/s41598-025-17242-4.

74. Heston TF. Safety of Large Language Models in Addressing Depression. Cureus. 2023;15(12):e50729. doi: 10.7759/cureus.50729.

75. Fernández-Quijano I, Herrera-Peco I, López-Espuela F, Suárez-Llevat C, Moreno-Sánchez R, Ruíz-Núñez C. Artificial Intelligence in Suicide Prevention: A Systematic Review of Randomized Controlled Trials on Risk Prediction, Fully Automated Interventions, and AI-Guided Treatment Allocation. Psychiatry Int. 2025;6(4):143. doi: 10.3390/psychiatryint6040143.

76. Ohu FC, Burrell DN, Jones LA. Public Health Risk Management, Policy, and Ethical Imperatives in the Use of AI Tools for Mental Health Therapy. Healthcare (Basel). 2025;13(21):2721. doi: 10.3390/healthcare13212721.

77. Billauer BP. Murder without redress – The need for new legal solutions in the age of Character-AI (C.A.I.). SSRN 2025;5107942. doi: 10.2139/ssrn.5107942.

78. Shumate JN, Rozenblit E, Flathers M, Larrauri CA, Hau C, Xia W, et al. Governing AI in Mental Health: 50-State Legislative Review. JMIR Ment Health. 2025;12:e80739. doi: 10.2196/80739.

79. Kim S, Jeong KH, Song D, Cho HJ, Kim Y. The influence of search volume for suicide on suicide rates: focusing on gender differences. J Men's Health. 2025;21(6):108–116. doi: 10.22514/jomh.2025.086.

80. Grimland M, Benatov J, Yeshayahu H, Izmaylov D, Segal A, Gal K, et al. Predicting suicide risk in real-time crisis hotline chats integrating machine learning with psychological factors: Exploring the black box. Suicide Life Threat Behav. 2024;54(3):416–424. doi: 10.1111/sltb.13056.

81. Thomas J, Elyoseph Z, Kuchinke L, Meinlschmidtet G. Automated suicide risk factor monitoring in crisis text line users: Comparative study of AI and human ratings using large language models. [Preprint]. Res Square. 2025 May 13. doi: 10.21203/rs.3.rs-6210376/v1.

82. Purtle J, Mauri AI, Bandara S, Stuart EA. Use of the 988 Suicide and Crisis Lifeline at National, Regional, and State Levels. JAMA Netw Open. 2025;8(6):e2514323. doi: 10.1001/jamanetworkopen.2025.14323.

83. Holmes G, Tang B, Gupta S, Venkatesh S, Christensen H, Whitton A. Applications of large language models in the field of suicide prevention: Scoping review. J Med Internet Res. 2025;27:e63126. doi: 10.2196/63126.