ARTICLE

Peer Reviewed

Ethical Suicide Prevention in an Artificial Intelligence Driven Society

Martin Orr MB DBA, Auckland University of Technology, New Zealand
Kirsten van Kessel PhD, Auckland University of Technology, New Zealand
David Parry PhD, Murdoch University, Australia

Objective

To explore the drivers, objectives and challenges for ethical artificial intelligence-based suicide prevention in a technologically evolving clinical and societal context.

Method

Narrative selected literature review.

Results

There is an ethical need for more effective suicide prevention that is aware of the limitations of prediction. AI research needs to recognise the importance of clinical formulation and risk management and timely therapeutic engagement in suicide prevention.

The AI transformation of society may change relationships and roles and alter suicide determinants and prevention.

AI may contribute to suicide prevention by facilitating novel research and practice and building the understanding of suicide typology, lived experience, determinants, risk signal models, intervention effectiveness and priority. AI may transform suicide prevention, through enabling timely personal to societal level responses, and leveraging technologies from social media and natural language processing to intelligent digital clinicians and guardians and self-driving cars. AI is increasingly used to enhance engagement and sense, model and shape conscious thought, experience and action patterns at a population and personal level. AI can both protect or harm and guard against or help propagate cyber trauma and cyberbullying. The capacity to architect reality and modulate consciousness may have positive and negative consequences on mental health and suicide prevention and a need for psyche cybersecurity. The potential use of AI in assisted suicide also helps highlight key ethical issues that clinicians and society face.

Conclusions

To optimise benefit and minimise algorithmic harm, clinicians have an ethical duty to be involved in the sociocultural and contextually aware co-design and governance of AI based suicide prevention.

Key words

Suicide prevention, ethics, research, artificial intelligence, social media, co-design

Introduction

There has been a significant focus on the potential prediction capabilities of Artificial Intelligence (AI) in suicide prevention (Lejeune et al.,2022). This paper aims to put that discussion in a broader ethical, societal, health and risk-benefit analysis context and portent the need for clinician involvement in AI co-design and governance. The limitations of prediction will be highlighted, and an argument made for the broader impacts AI could have on suicide and suicide prevention.

Suicide is integrally linked with society and society is rapidly being transformed by digital technology, particularly AI. Suicide is influenced by social determinants, and societal disruptions and transformations from AI could impact on suicide rates both negatively and positively (Mueller et al., 2021).

The science and practice of suicide prediction and prevention is complex, contentious, and limited. Assessing and addressing suicidal ideation is a core aspect of clinical practice. However suicidal ideation is a moderately weak signal for completed suicide with limited predictive power (Large, 2018; Kessler et al., 2020; Mann et al., 2005; Pisani et al., 2016; Nock et al., 2019; Woodford et al., 2019).

Legally, a finding of suicide typically requires an intentional and knowing act; however, it can be challenging to piece together the lived experience, signals, chronology and nature of what at times is mixed, changing and ambiguous intent and limited, even to the individual, knowing (Tait & Carpenter, 2015).

There is an increasing interest in AI being applied to prediction models in healthcare and providing human behavioural insights (de Hond et al.,2022; Dehghan et al.,2022).

An AI research focus on suicide risk prediction may afford some benefits in better risk stratification for risk management and care prioritisation. However, there is a need to recognise the limitations and potential harms of prediction of human behaviour in the prevention of suicide as a complex low base rate event.

AI is an evolving concept that concerns the intelligence of machines or computer algorithms. AI has a focus on pattern recognition, modelling and continuous learning. AI may assist with the suicide prevention need for novel richer data sources, theory, research and practice. AI can identify, analyse, integrate, and model multiple signals utilising a broad range of sources that may include behavioural,

neurophysiological, and language data (Resnik et al., 2020).

The scalability, leverage and reach of AI can amplify both benefit and risk. Although social media platforms are introducing AI driven suicide prevention, the area has many complex ethical, potential bias, clinical effectiveness, explicability, acceptability, medicolegal and safety issues. These require further research to inform practice (Floridi & Cowls, 2019; Binz & Schulz, 2023).

The aim of this paper is to highlight some of the strategically significant issues related to where AI and suicide prevention might be going, and the importance of ethical co-design and governance for the journey.

The paper highlights the potential benefits but also the need to recognise the limitations of natural and artificial intelligence and guard against both iatrogenic and algorithmic harm. Explicable and reflexive, evidence and ethical based, sociocultural and contextually aware algorithmic modelling and governance will be essential.

The paper provides an overview of suicide, and the typical quoted statistics and influences on our current theory and daily practice. It will discuss the historical and evolving ethical, legal and societal context of suicide. A range of AI technology and decision science concepts across the suicide prevention continuum are reviewed. These include machine learning and deep neural networks, societal and global level interventions, modelling of lived experience, new types of digital practice and practitioner, and AI's capacity to engage, analyse, influence and architect consciousness and reality. The central argument in this paper is that mental health clinicians have a duty and expertise to help co-design ethical AI based suicide prevention.

Methodology

This study used a narrative selected literature review (Ferrari 2015; Baumeister & Leary, 1997). The literature review involved the search of Google Scholar, PsycINFO, Cochrane Library and Medline focusing on relevance, citation weight and recency, and following up on references in key identified sources. The search was limited to articles published in English. Relevant references were identified from 1878 to 2023. Search terms included: suicide prediction and prevention; suicide and society, stigma, trauma, culture; AI and society, ethics, bias and discrimination, co-design and governance; digital mental health, social media, AI, machine learning, computational linguistics, natural language

processing (NLP) and suicide prevention; assisted suicide; dual use dilemma, cyberwarfare, cyberbullying and cyber trauma, threat detection and social cybersecurity; intelligent digital agents, guardians, theory of mind and therapeutic alliance. The broad range of search terms and wide date range reflected the intent to give a broad critical perspective of the changing societal context of suicide and the many complex intricate factors that require consideration in relation to the potential impact of AI.

Suicide prevention challenges and need

Suicide prevention needs new theory, ethical research methods and practice. Globally each year an estimated 15.4/100000 adults will kill themselves, and 20 times this number will make an attempt (World Health Organization, 2014). The cross-national lifetime prevalence of suicidal ideation is 9.2% (Nock et al., 2008a). A systematic review of psychological autopsy studies of suicide found 90% had an identifiable mental disorder and 38% co-morbid mental disorder and substance abuse. (Cavanagh et al., 2003). There is a strong link between depression and suicide (Harris & Barraclough, 1997). In the year before suicide 3/4 may have visited a primary care provider and 1/3 have visited mental health services (Luoma et al., 2002). There is evidence, at least for some, of progression of suicide risk from ideation to plans to action. There are also demographic and diagnostic correlations that can be linked to both risk of, and speed of progression from, ideation to action (Nock et al., 2008b).

These research findings and statistics suggest many potential opportunities for prevention. However, translating this statistical knowledge into preventing suicides presents a multitude of clinical and ethical challenges from the individual assessment to national and international strategy level. These statistics and the resultant focus on identifying suicidal ideation and treating mental illness are not without limitations and are subject to varying degrees of methodological and reporting validity and regional, cultural, sex and age relevancy. Of note is the concern that psychological autopsy studies may be flawed and limited in international application. Research from predominantly European and North American contexts may not be as relevant in countries such as China and India where there is a higher proportion of suicide deaths without a psychiatric diagnosis, and different conceptualisations of mental illness and social determinants (Hjelmeland et al., 2012; Pridmore, 2015; Milner et al., 2013).

Although the absolute numbers are devastating, suicide remains a low base rate complex event making prediction and the ability to demonstrate benefit from an intervention challenging. The time

period for intervention may be brief and not align with engagement or contact with clinical services (McHugh et al., 2019a; Blanchard & Farber, 2020; McHugh et al., 2019b; Simon et al., 2001).

Suicide prevention programs, particularly those focusing on restricting means, facilitating engagement and follow up, and recognising and treating mental illness (especially depression) have had some success (Turecki & Brent, 2016). However, what works in what context and at what time and in what combination is still unclear and it is difficult to demonstrate a sustained significant impact on predictive ability and mortality for many trialled interventions that appear to have good face validity (Schlichthorst et al., 2020; Nock et al., 2019; Franklin et al., 2017).

There is a need to more meaningfully understand different types and pathways to suicide, to timely prioritise and target limited resources, and ethically enhance the humanity of our approach for all those caught up in a traumatic vulnerable process.

Greater capability is required to research those who never present to psychiatric services, who do not identify themselves as suicidal nor identify as mentally ill; there is a need to understand better their lived experience, the metaphors and concepts they use to describe how suicide develops and how it is decided and acted on as an option. AI machine learning approaches particularly in the area of computational linguistics may assist in building a richer, timelier, research picture of suicide and addressing some of these challenges (Resnik et al., 2020).

Machine learning and Modelling

Computational linguistics is the study of language using computational methods and theoretical models. Machine learning is a branch of AI involving the development and training of algorithms in key feature and pattern recognition to facilitate future decision making and prediction. Natural Language Processing (NLP) is an applied subfield of computational linguistics. Neural network and deep learning-based NLP methods are the current predominant machine learning approaches to computational linguistic research and application in suicide prevention (Resnik et al., 2020). NLP focuses on pattern recognition of the mathematical probabilistic vector relationship between words in different contexts. Pattern recognition around words and language in context is central to mental health practice. Language is a method to gain insight and create a model of another's consciousness, intent and experience. NLP can be applied to social media data with a view to modelling and detecting signals of suicide and other mental health issues and assisting with clinical formulation and the triaging

and prioritisation of care (Resnik et al., 2020; Orr et al., 2022).

Pattern recognition, learning and modelling are core to AI. Any deficits or bias in the underlying model of practice, or societal perspective can be learned by the AI. Diagnostic, risk categorical and treatment models derived from international data may have transcultural limitations particularly when trying to understand the consciousness, intent and experience of another (Bredström, 2019; La Roche et al., 2015; Durie, 2017; Hatcher et al., 2017). To ethically maximise benefit and minimise harm, models built and utilised with AI need to be interpreted and formulated with a sociocultural contextual awareness (Orr et al., 2022).

AI has increasing strengths in the analysis and generation of language and influencing human attention and engagement (Binz & Schulz, 2023; van Dis et al., 2023). There is increasing capacity for AI to be pre-trained on wide arrays of data where a mathematical probabalistic understanding of the relationships and logic of language is developed. Pre-trained learning and chatbot language generation capacity can then be transferred and fine-tuned to specific tasks through varying levels of supervision, weighing and integration of additional contextual data. These AI strengths could enhance clinical formulation (where weighing and integration of multiple data sources is central) and the capacity for customised timely trauma informed therapeutic engagement (Yeager & Benight, 2022). Enhanced clinical and risk management formulation and timely therapeutic engagement may be of more significant suicide prevention utility than a static risk prediction categorisation that doesn't recognise the importance of timing, dynamic complexity, ambiguity and context (Orr et al.,2022; Hawton et al.,2022).

AI can amplify bias, including gender and cultural bias, by picking up and learning what is happening in the current system and replicating that language and practice on the assumption that being a common pattern it is what is desired. There is a need to detect and protect against inherent bias in the learning dataset (on which the AI algorithm is developed and trained), being amplified in real world application as perceived desired and best practice. There is a need for ongoing critical analysis, debiasing and governance of the algorithm's design and outcomes including debiasing and diversity measures for the humans in the design and governance loop (Savoldi et al., 2021; Garrido-Muñoz et al.,2021).

AI and the Suicide Prevention Continuum

AI and social media platforms are potential cost-effective tools for unmet need across the suicide prevention continuum. They may assist with timely early intervention, and improving integration, reach, education, research and mental health workforce capacity and augmentation (Paterson et al., 2018).

However, there are fundamental questions about the degree suicide prediction can be improved by AI computational or statistical processes and how such predictions may meaningfully, ethically and sustainably contribute to suicide prevention. There are major limitations and clinical and ethical risks in trying to absolutely predict and categorise complex multifactorial emergent low base rate human behavioural events that have a high magnitude of adverse consequence if that prediction is incorrect. There are issues both around acting or failing to act. There are issues around having the sustainable capacity, motivation and resource mix to act, particularly when a calculated risk refers to some undefined time in the future. However, AI may be able to assist in identifying what key potential contextual risk factors are for an individual or community; assist in the triage and prioritisation of attention for that individual or community; and do this at a timely speed and scale that exceeds human capabilities (Resnik et al., 2020; Shing et al., 2020).

AI could contribute to the broader suicide prevention continuum that extends beyond prediction to wellness and resilience promotion, detection of suicidal ideation and behaviours and targeting of selected interventions, and postvention support for grieving families and communities (Ministry of Health, 2019).

Intervention triggers and vigilance

The concept of suicide crisis syndrome highlights the challenges and limitations of utilising voiced suicidal ideation as the primary risk signal and trigger for suicide prevention interventions. Suicide crisis syndrome is characterised by the perception of an unbearable or inescapable situation, with related affective and cognitive dysregulation, hyperarousal, and social withdrawal (Rogers et al.,2017). The emerging related literature emphasises that many people who suicide never report suicidal ideation and there may be a relatively short window between an individual feeling overwhelmed and the emergence of thoughts and behaviour leading to death. (Galynker et al., 2017; Obegi, 2019; Schuck et al., 2019; Yaseen et al., 2019).

Clinical services being typically arranged around discrete, brief, spaced and rationed contacts have (even with the best triage and emergency support services) limited capacity to respond to the dynamic complex environments of human lives where the briefest change in timepoint may have a significant impact on risk and emergent outcome. It is possible that clinical services could benefit from the timeliness and multi-signal vigilance that AI can provide, but the degree to which such technology should be adopted must be subject to ethical risk-benefit analysis.

Whole of Societal Impact

In a digital society, how we communicate and relate will continue to change and how we communicate and experience our distress and seek support and solace may alter. Individuals and family are increasingly turning to social media to access interventions and crowd source advice (Thorn et al., 2020; Latha et al.,2020; Garg, 2023). Technologies may shape and structure the fabric of society and potentially the function and expression of suicide and related interventions.

The dystopian fear of artificial general intelligence and dominance of humanity may not be imminent (Mullins, 2022). However, there is the possibility of significant human redundancy resulting from AI empowered destruction of industries and roles. This may be relevant to a consideration of potential social determinants of suicide. The hope is that this may be creative destruction with better opportunities for meaningful work and activity, rights protection and collaborative innovation emerging (Borsci et al., 2022; Feijóo et al., 2020). Still, there is the realistic fear that some may be left behind and lose the sense of control, competency, and connectedness in one's world that is important for a positive self-concept and mental wellbeing.

Suicide prevention programs are increasingly recognising the need for sociological insights and whole of society approaches, where everyone can play a role in self-care and care of their community (Mueller et al., 2021; Hadlaczky et al., 2014; Bond et al.,2021). There is increasing international interest in how family and community members can play a role in emotional, mental and suicide prevention first aid for those in crisis (Armstrong et al., 2020; Lu et al., 2020) This includes the role of mobile applications and online groups and design of programs that may best facilitate bystander intervention (Gilat & Shahar, 2009; Brown et al., 2020; Hill et al.,2020).

Expanding from a clinical service to societal focus allows for more protective layers and opportunities for intervention at more time points. Expanding suicide prevention investment to leveraging the whole

community can harness the vigilance, reach and integration advantages of AI based digital transformation. However, it is important to recognise the potential limitations of citizen based mental health and suicide prevention first aid and online support type programs and the need for ongoing education, governance and research to improve the safety and quality of these interventions (Lavis & Winter 2020; Forthal et al., 2022).

Ethical Co-design and AI

AI is changing the nature of research in terms of how data is generated and utilised to timely sense, analyse and shape human decisions and behaviour. Relatively little of this research occurs within traditional ethical frameworks but instead as a function of neuromarketing particularly in relation to social media. Digital technology companies are already applying these techniques to suicide prevention (Marks, 2019).

Researching complex, distressing, nonlinear low base rate events, in vulnerable populations with a limited theoretical, definitional, and investigative toolkit base is pragmatically and ethically difficult. Required are ethical processes that are facilitative of good practice and protect a vulnerable community, but not so prohibitive they deter good research and fail to grapple with the societal suicide prevention reality and need (Andriessen et al., 2019; Hom et al., 2017; Lakeman & FitzGerald, 2009; Leenaars et al., 2002).

Psychiatry's focus on the unidentified and undertreated illness model as a predominant target for suicide prevention has had benefit. However, it also has limitations particularly if there is a failure to appreciate the broad influence of the social fabric in which we live and in which mental illness, if present, is woven (Pridmore, 2015; Sara, 2015).

When considering the application of AI to a domain such as suicide prevention, it is important to consider the conscious and subconscious biases that may remain inherent in a society as the machine may learn from these and consider them a desired norm of the collective consciousness. Through history there has been a spectrum of cultural, societal and religious perceptions and responses to suicide from acceptance to commiseration to commendation to condemnation. In the 'Anglo-Western' world suicide has been considered a crime against canon and common law, an infringement of moral and property law, and a crime (Felo de Se) requiring punishment beyond death (Chang, 2018). Historically this condemnation approach was associated with a perceived imperative for clinicians and

coroners to declare someone thought to have died by suicide as mentally ill or non-compos mentis to protect the individual and their family (Huggard, 1878; Chang, 2018). This significant societal stigma may have adversely impacted on research and intervention and blunted the critical search for broader social determinants.

Durkheim's concepts of different types and motivations for suicide can be criticized for limited methods and analysis (Mueller et al., 2021). However, the fundamental idea that humans exist within a social fabric and are influenced by their perception of their place, connections and role within that fabric and the forces that hold it together or rip it apart remains a sound premise. There may be different and changing reasons, social determinants and typologies for suicide and AI enabled research may be able to better timely categorise, trend and define these (Maskill et al., 2005; Clapperton et al., 2020; John et al., 2020; Martin et al., 2020).

Co-design and personalisation are important for digital mental health service development (Fleming et al., 2019; Hickie et al., 2019; Thorn et al., 2020). Co-design in AI empowered suicide prevention requires the working together of clinicians, researchers, communities, consumers, and digital media companies.

Floridi and Cowls (2019), posit 5 core ethical principles for AI design and deployment: autonomy, justice, beneficence, non-maleficence and explicability. Explicability integrates concepts of intelligibility and accountability. The design and deployment of AI should require an understanding of how the algorithm functions and who is accountable for the ongoing quality monitoring, development, and fit for purpose governance (Floridi & Cowls, 2019; Price et al., 2019).

Digital agents, cars, consciousness, cybersecurity and assisted suicide

This section addresses selected AI emergent developments with significant ethical and societal aspects and their potential impact on suicide prevention. These include digital clinicians or guardians and the therapeutic alliance, self-driving cars, consciousness, psyche cybersecurity and assisted suicide.

Digital agents, empathy and therapeutic alliance

Digital transformation is about creating and empowering new types of practice, practitioner and contexts for care. Artificial and natural intelligence have the capacity to synergically improve the

understanding and performance of each other including in the key areas of decision making, theory of mind, consciousness, and self-awareness. These areas are particularly important to the development of intelligent digital agents. Intelligent digital agents can take many forms, including chatbots, virtual assistants, companions and guardians. These digital agents may learn from interactions and provide increasingly more accurate, timely and personalised responses, that sustainably engage and enhance the user experience.

Trust, empathy and the perception of mutual goals and values is central to the therapeutic alliance and engagement. There is significant interest in the concept of an AI empowered digital therapeutic alliance and digital empathy and the impact of varying types of anthropomorphism, linguistic patterns, engagement contexts and technologies (Tremain, 2020; Lederman & D'Alfonso, 2021; Christian et al., 2021; Gaume et al., 2019). An ability to collaboratively personalise a constantly available and vigilant, empathic AI empowered digital clinician, coach or guardian and customise multiple characteristics from age, sex, gender, ethnicity, appearance, manner, voice, skill set and level of anonymity may be a trusted preference or acceptable for many (Sharma et al., 2023).

AI empowered guardian type technology is already diffusely deployed within the digital ecosystem. Social media platforms and search engines are continually analyzing for threats and potential high risk words or topics like suicidal ideation and methods, self-harm and potential drug overdose and will offer a suicide prevention result (SPR) in response. Similarly, chatbots and digital assistants can recognize trigger words and sentiment in written and spoken word and recommend seeking professional help or calling crisis services. SPRs give advice on self-care and resources to call for support with the aim that at least some will take the option of clicking and following through on seeking help and support (Borge et al., 2021; Broer, 2022). There may be significant diversity and equity issues in the nature of SPR response, with variation in the provision of SPR depending on country and language and the availability of local services (Scherr et al., 2022). Future digital guardians focused on suicide prevention may have a spectrum of roles including monitoring and alerting, education and coaching, and therapeutic interventions depending on the personalised clinical formulation of the individual and their context and needs (Orr et al., 2022).

The ability to design the look and characteristics of AI technology preferred or desirable to an individual is already a rapidly developing aspect of chatbot companion AI. Such personalised technology can quickly learn preferences and perceived needs and patterns of behaviour from millions of users and

utilise a range of contextualised targeted calibrated techniques to keep a user engaged (Hakim et al., 2019; Xie et al., 2022). Digital AI agents raise significant ethical issues from concerns about privacy, security, bias, personal autonomy, coercive intrusion and psychological addiction and manipulation to clinical effectiveness, safety, and responsibility (Bhatia-Lin et al., 2019). Digital agents may have challenges in building a therapeutic alliance and handling the complexity, anger and pain of people who are desolate, desperate, disengaged or overwhelmed by crisis (Tong et al., 2022). However, for some individuals and social systems, digital care agents may be the major timely or preferred option for care and could make significant contributions to diversity and equity of service provision and augmenting engagement and accessible intervention (Torok et al., 2020).

Self-driving cars

An AI development that may have an immediate and transferrable technology impact on suicide is self-driving cars. AI empowered autonomous cars can improve mobility and access to meaningful employment and social connection. They may have a reduced ability to crash, reducing the head injury and suicide by car accident rate. Medical contact for a traumatic brain injury is significantly associated with an increased suicide risk and at least 2% of fatal car accidents are thought to be related to suicide (Madsen et al., 2018; Pompili et al., 2012). The real time image analysis and autonomous process innovations developed for self-driving cars could assist with the development of virtual clinicians and guardians and improving personalised care delivery. Self-driving cars also illustrate some of the ethical challenges that AI affords e.g., who the car saves if a driver is intentionally trying to kill themselves but to avert the danger may potentially harm others.

AI Architecting Reality and modulating consciousness

Suicide may at times be perceived as an escape from consciousness particularly when that consciousness is dominated by an escalating overwhelming sense of pain, loss, shame or entrapment (Shneidman, 1993).

Mixed realities, hybrid identities and consciousness are a significant focus of AI research primarily in relation to how AI may influence, enhance or emulate them. Digital media AI is designed to continually architect reality, draw relationships between entities, modulate consciousness and shape identity; it amplifies and diminishes attention, focus, arousal, curiosity, and action and connects the dots when

making sense of experience. While this use of technology is typically through an external device there is increasing interest in the use of implantable devices. These Cyborg psychotherapeutic and AI plugin neural symbiotic enhancements may come to be used in the neuromodulation and management of a range of disorders and the psyche including suicide prevention (Kulshreshth et al., 2019; Reinares-Lara et al., 2018). These reality architecting and consciousness modulating strategies could be used for societal and mental healthcare good or harm.

Cyber Trauma and Psyche Cybersecurity.

Cybertrauma is a significant concern for mental health and suicide prevention (Suman, 2018; Paat & Markham, 2021; Dorol & Mishara, 2021; Perwitasari & Wuryaningsih, 2022).

Cyber trauma or harm may result from targeted attacks on the psyche of individuals and society with threats, bullying, fragmentation and manipulation of their attention, motivational values and sense of purpose, hope, self and community.

AI based social media is designed to be targeting, engaging, connecting and addictive. There may be a spectrum of vulnerabilities to adverse consequences and perceptions of what may be harmful within a particular context. Social media may have a range of adverse associations with suicide by inviting cyberbullying and cyber trauma or harm and constant influences that question worth and identity, as well as pro suicide influences that offer suicide as a solution to problems, to the social isolation and sleep deprivation related to prolonged addictive use (Millington, 2017; Kato et al., 2020; Perwitasari & Wuryaningsih, 2022).

Current computational linguistic research and social media-based suicide protection responses tend to be aimed at risk, threat, aggression and cyber bullying and harm detection. There is a focus on detection, classification, and moderation of malicious, adverse and risk content, and through this the aim to sanction perpetrators and protect vulnerable users (Kumar et al., 2020).

AI's capacity for sensing and shaping of the psyche represents a significant dual use ethical dilemma with potential for both good and harm at the individual and societal level (Butorac et al., 2021). An AI enhanced digital theory of mind and formulation capacity, situational awareness and therapeutic alliance may contribute to guardian protective functions and effective trauma informed, empowering, growth focused care (Chen et al., 2022). However, the same capacities could also be utilized to

increase the vulnerability to cyber trauma, disempowerment, control, manipulation and degradation of the psyche.

There is increasing recognition of the potential for cyber trauma, control, manipulation and harm to occur not just at the individual level, via individual cyberbullies or bad actors, but to occur at a larger organised societal scale (Carley, 2020). State actors may engage in societal control, and hybrid or cyberwarfare, including disinformation campaigns, algorithmic imperialism, and the related surveillance, repression, and political influencing of whole communities in how they think, feel and act. Computational propaganda and disinformation can be utilised for population radicalisation and demoralisation. There is a capacity to make a calibrated targeted attack on the mental state and morale of a broad range of groups within a society using trolls, bots, and deep fake technology. With a potential influence on mental wellbeing and the social determinants of suicide, cybersecurity developments to protect against such attacks on the human psyche and society will require a co-design approach with mental health stakeholders and experts (Beskow & Carley, 2019; DiResta, 2018; Massing-Schaffer & Nesi, 2020).

Assisted Suicide.

Attitudes, definitions, legal and clinical models and practices of euthanasia and physician assisted suicide continue to evolve throughout the world and be associated with significant complexity and contention (Calati et al., 2021; Emanuel et al., 2016). Evolving concepts and practices often leverage and integrate the evolving technology in a society and it can be envisaged that, as well as suicide prevention, AI could be used in assisted suicide and euthanasia. Assisted and legally sanctioned suicide has the inherent belief that some forms of suicide are rational. If something is rational, an artificial intelligence algorithm may assist with the optimal application of logic to that process. In state sanctioned assisted suicide, euthanasia, or medical assistance in dying, clinicians participate in the decision making with a view to relieving suffering. Capacity for competent, conscious judgement that is not being inappropriately or unduly influenced by coercion or mental or neurological disorder must be determined to be present. In keeping with the motivating force that may drive some forms of unassisted suicide is the understanding that death is the best or only way to timely relieve the conscious experience of suffering or pain that could not or should not be borne.

Complex intersections of law, ethics and medical practice are often developed into best practice

decision guidelines and sanctioned protocol in both paper and electronic form. Within an assisted suicide context this process could become augmented then increasingly dominated by artificial intelligence. The related AI algorithm could evolve from an advisory to increasingly autonomous role with the means for an approved suicide delivered by the machine. Such an evolvement would afford significant moral and ethical challenges.

Evolving algorithmic ability and capacity for autonomy and how much a human could or should remain in the algorithmic loop are significant design, ethical and quality issues for AI. Dynamic and ongoing evolvement of the algorithm requires dynamic and ongoing governance. Assisted suicide and clinician involvement in assisted suicide raises major complex and contentious ethical issues, particularly if the perceived untreatable nature of, or suffering from, a mental disorder is a major factor in the assessment and decision-making process. Acting or failing to act and participate in the process could be interpreted as ethically wrong by different groups (Maher, 2017; Dembo, Schuklenk et al., 2018). For those opposed to the concept and clinician involvement, there is the ethical dilemma or risk, that even collaborating to minimise harm may be considered or reframed as consensus and support. This may afford particular difficulties if the perceived sanction and approval are considered ongoing and the practice later evolves or is actioned in ways not previously envisaged or agreed.

Assisted suicide AI algorithms would need human design and governance, raising issues of how life and death, judgement, and suffering, may be valued and operationally logically defined. There would be a need to establish who gets to design and decide the assessment process and relative weighting for each algorithmic factor. There would be a need to determine who governs process and algorithmic quality and clinical, legal, and ethical safety and ensures adequate assessment and treatment and guards against coercion and a right becoming an obligation (Gale & Barak, 2020; Kim et al., 2016; Maher, 2017).

Discussion and Conclusion

This narrative literature review has identified a significant body of work that emphasises the complexity, contention, and limitation of current suicide prevention practice. This review adds to that and highlights with an ethical focus the potential of AI for broader calibrated research and interventions at the individual to societal and national strategy level. AI will likely shape and structure the fabric of society, and inevitably to some degree the function and expression of suicide. Clinicians need to

consider their ongoing role and duty in shaping the digital society; what role they will play in architecting and governing the digital clinical services that will be provided; and how they will work with the community and the digital industry to promote ethical research and practice.

Suicide is a complex time and context dependent phenomenon. AI assisting in the expansion from a clinical service to societal focus allows for more protective layers and opportunities for intervention at more time points. It affords more opportunities for connection, support, problem solving and meaningful purpose, recognising the central role that feeling overwhelmed, useless, hopeless and alone may play in suicide.

AI could contribute to suicide risk identification, resilience development and related disorder recognition and recovery. AI could help develop a greater understanding of suicide typology, sociocultural determinants and the prioritisation and development of specific interventions in specific contexts. However, AI has the potential for algorithmic harm and bias, highlighting the importance of explicable, reflexive, contextual and socioculturally aware modelling, and governance.

This narrative literature review has limitations. AI technologies are described only at a conceptual level. The paper includes only a select literature review in the coverage of subject matter that embraces multiple rapidly growing and changing fields and ethical and philosophical dimensions.

The purpose therefore was not to set out a full map, but to draw attention to some significant issues related to the intersection of AI and suicide prevention, and the central role of ethical co-design and governance for the journey. The aim was not to attempt to provide all the answers, but to highlight that AI raises complex moral and ethical issues to be addressed collectively by mental health professionals and wider society. Although the focus was suicide, the issues raised have major implications for mental health in general and the ethical duty of clinicians to evolve with and shape the emergent digital society.

References

Andriessen, K., Reifels, L., Krysinska, K., et al. (2019). Ethical Concerns in Suicide Research: Results of an International Researcher Survey. *J Empir Res Hum Res Ethics*, 14(4), 383-394.

Armstrong, G., Sutherland, G., Pross, E., Mackinnon, A., Reavley, N., & Jorm, A. F. (2020). Talking about suicide: an uncontrolled trial of the effects of an Aboriginal and Torres Strait Islander mental health first aid program on knowledge, attitudes and intended and actual assisting actions. *PLoS one*, 15(12), e0244091.

Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of general psychology*, 1(3), 311-320.

Beskow, D. M., & Carley, K. M. (2019). Social cybersecurity: an emerging national security requirement. *Millatary Review*.

Bhatia-Lin, A., Boon-Dooley, A., Roberts, M.K., et al. (2019). Ethical and regulatory considerations for using social media platforms to locate and track research participants. *The American Journal of Bioethics*, 19(6), 47-61.

Binz, M., & Schulz, E. (2023). Using cognitive psychology to understand GPT-3. *Proceedings of the National Academy of Sciences*, *120*(6), e2218523120.

Blanchard, M., & Farber, B. A. (2020). "It is never okay to talk about suicide": Patients' reasons for concealing suicidal ideation in psychotherapy. *Psychotherapy Research*, 30(1), 124-136.

Bond, K. S., Cottrill, F. A., Mackinnon, A., Morgan, A. J., Kelly, C. M., Armstrong, G., ... & Jorm, A. F. (2021). Effects of the Mental Health First Aid for the suicidal person course on beliefs about suicide, stigmatising attitudes, confidence to help, and intended and actual helping actions: an evaluation. *International journal of mental health systems*, *15*(1), 1-12.

Borge, O., Cosgrove, V., Grossman, S., Perkins, S., & Van Meter, A. (2021). How Search Engines Handle Suicide Queries. *Journal of Online Trust and Safety*, 1(1).

Broer, T. (2022). The Googlization of Health: Invasiveness and corporate responsibility in media discourses on Facebook's algorithmic programme for suicide prevention. *Social Science & Medicine*, 306, 115131.

Brown, K., Toombs, M., Nasir, B., Kisely, S., Ranmuthugala, G., Brennan-Olsen, S. L., ... & Hides, L. (2020). How can mobile applications support suicide prevention gatekeepers in Australian Indigenous communities?. *Social Science & Medicine*, *258*, 113015.

Borsci, S., Lehtola, V. V., Nex, F., Yang, M. Y., Augustijn, E. W., Bagheriye, L., ... & Zurita-Milla, R. (2022). Embedding artificial intelligence in society: looking beyond the EU AI master plan using the culture cycle. *AI & society*, 1-20.

Bredström, A. (2019). Culture and context in mental health diagnosing: Scrutinizing the DSM-5 revision. *Journal of Medical Humanities*, 40, 347-363.

Butorac, I., Lentzos, F., & Aicardi, C. (2021). Gray Matters: Exploring Technologists' Perceptions of Dual-Use Potentiality in Emerging Neurotechnology Applications. *Health security*, *19*(4), 424-430.

Calati, R., Olie, E., Dassa, D., Gramaglia, C., Guillaume, S., Madeddu, F., & Courtet, P. (2021). Euthanasia and assisted suicide in psychiatric patients: A systematic review of the literature. *Journal of Psychiatric Research*, 135, 153-173.

Carley, K. M. (2020). Social cybersecurity: an emerging science. *Computational and mathematical organization theory*, *26*(4), 365-381.

Cavanagh, J. T., Carson, A. J., Sharpe, M., et al. (2003). Psychological autopsy studies of suicide: a systematic review. *Psychol Med*, 33(3), 395-405.

Chang, H. Y. (2018). A Brief History of Anglo-Western Suicide: From Legal Wrong to Civil Right. *SUL Rev*,

46, 150.

Chen, J. X., McDonald, A., Zou, Y., Tseng, E., Roundy, K. A., Tamersoy, A., ... & Dell, N. (2022, April). Trauma-informed computing: Towards safer technology experiences for all. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1-20).

Christian, C., Barzilai, E., Nyman, J., & Negri, A. (2021). Assessing key linguistic dimensions of ruptures in the therapeutic alliance. *Journal of Psycholinguistic Research*, *50*, 143-153.

Clapperton, A., Bugeja, L., Newstead, S., et al. (2020). Identifying Typologies of Persons Who Died by Suicide: Characterizing Suicide in Victoria, Australia. *Arch Suicide Res*, 24(1), 18-33.

de Hond, A. A., Leeuwenberg, A. M., Hooft, L., Kant, I. M., Nijman, S. W., van Os, H. J., ... & Moons, K. G. (2022). Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. *NPJ digital medicine*, 5(1), 2.

Dehghan, P., Alashwal, H., & Moustafa, A. A. (2022). Applications of machine learning to behavioral sciences: focus on categorical data. *Discover Psychology*, 2(1), 22.

Dembo, J., Schuklenk, U., & Reggler, J. (2018). "For their own good": A response to popular arguments against permitting medical assistance in dying (MAID) where mental illness is the sole underlying condition. *The Canadian Journal of Psychiatry*, *63*(7), 451-456.

DiResta, R. (2018) Computational propaganda: If you make it trend, you make it true. *The Yale Review*, 106(4), 12-29.

Dorol, O., & Mishara, B. L. (2021). Systematic review of risk and protective factors for suicidal and self-harm behaviors among children and adolescents involved with cyberbullying. *Preventive medicine*, *152*, 106684.

Durie, M. (2017). Indigenous suicide: the turamarama declaration. *Journal of Indigenous Wellbeing*, 2(2), 59-67.

Emanuel, E. J., Onwuteaka-Philipsen, B. D., Urwin, J. W., & Cohen, J. (2016). Attitudes and practices of euthanasia and physician-assisted suicide in the United States, Canada, and Europe. *Jama*, 316(1), 79-90.

Feijóo, C., Kwon, Y., Bauer, J. M., Bohlin, E., Howell, B., Jain, R., ... & Xia, J. (2020). Harnessing artificial intelligence (AI) to increase wellbeing for all: The case for a new technology diplomacy. *Telecommunications Policy*, 44(6), 101988.

Ferrari, R. (2015). Writing narrative style literature reviews. *Medical writing*, *24*(4), 230-235.

Fleming, T., Merry, S., Stasiak, K., et al. (2019). The importance of user segmentation for designing digital therapy for adolescent mental health: findings from scoping processes. *JMIR mental health*, 6(5), e12656.

Floridi, L. & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1).

Forthal, S., Sadowska, K., Pike, K. M., Balachander, M., Jacobsson, K., & Hermosilla, S. (2022). Mental health first aid: A systematic review of trainee behavior and recipient mental health outcomes. *Psychiatric*

services, *73*(4), 439-446.

Franklin, J. C., Ribeiro, J. D., Fox, K. R., et al. (2017). Risk factors for suicidal thoughts and behaviors: a meta-analysis of 50 years of research. *Psychological bulletin*, 143(2), 187.

Gale, C., & Barak, Y. (2020). Euthanasia, medically assisted dying or assisted suicide: time for psychiatrists to say no. *Australasian Psychiatry*, 28(2), 160-163.

Galynker, I., Yaseen, Z.S., Cohen, A., et al. (2017). Prediction of suicidal behavior in high risk psychiatric patients using an assessment of acute suicidal state: the suicide crisis inventory. *Depression and anxiety*, 34(2), 147-158.

Garg, M. (2023). Mental Health Analysis in Social Media Posts: A Survey. *Archives of Computational Methods in Engineering*, 30, 1819-1842.

Garrido-Muñoz, I., Montejo-Ráez, A., Martínez-Santiago, F., & Ureña-López, L. A. (2021). A Survey on Bias in Deep NLP. *Applied Sciences*, 11(7), 3184.

Gaume, J., Hallgren, K. A., Clair, C., Schmid Mast, M., Carrard, V., & Atkins, D. C. (2019). Modeling empathy as synchrony in clinician and patient vocally encoded emotional arousal: A failure to replicate. *Journal of Counseling Psychology*, *66*(3), 341.

Gilat, I., & Shahar, G. (2009). Suicide prevention by online support groups: An action theory-based model of emotional first aid. *Archives of Suicide Research*, *13*(1), 52-63.

Hadlaczky, G., Hökby, S., Mkrtchian, A., Carli, V., & Wasserman, D. (2014). Mental Health First Aid is an effective public health intervention for improving knowledge, attitudes, and behaviour: A meta-analysis. *International Review of Psychiatry*, *26*(4), 467-475.

Hakim, F. Z. M., Indrayani, L. M., & Amalia, R. M. (2019, February). A dialogic analysis of compliment strategies employed by replika chatbot. *Proceedings of the third International Conference of Arts, Language and Culture* (pp. 266-271). Atlantis Press.

Hatcher, S., Crawford, A., & Coupe, N. (2017). Preventing suicide in indigenous communities. *Current Opinion in Psychiatry*, 30(1), 21-25.

Harris, E. C., & Barraclough, B. (1997). Suicide as an outcome for mental disorders. A meta-analysis. *British journal of psychiatry*, 170(3), 205-228.

Hawton, K., Lascelles, K., Pitman, A., Gilbert, S., & Silverman, M. (2022). Assessment of suicide risk in mental health practice: shifting from prediction to therapeutic assessment, formulation, and risk management. *Lancet Psychiatry*, 9(11), 922-928.

Hickie, I. B., Davenport, T. A., Burns, J. M., et al. (2019). Project Synergy: co-designing technology-enabled solutions for Australian mental health services reform. *Medical Journal of Australia*, 211, S3-S39.

Hill, K., Somerset, S., Schwarzer, R., & Chan, C. (2020). Promoting the community's ability to detect and respond to suicide risk through an online bystander intervention model-informed tool. *Crisis*, 42(3), 225-231.

Hjelmeland, H., Dieserud, G., Dyregrov, K., et al. (2012). Psychological autopsy studies as diagnostic tools: are they methodologically flawed?. *Death Studies*, 36(7), 605-626.

Hom, M. A., Podlogar, M. C., Stanley, I. H., et al. (2017). Ethical Issues and Practical Challenges in Suicide Research. *Crisisx,* 38(2), 107-114.

Huggard, W. R. (1878). Felo-de-se. *J Psychol Med Ment Pathol*, 4(2), 309-319.

Kato, T.A., Shinfuku, N., & Tateno, M. (2020). Internet society, internet addiction, and pathological social withdrawal: the chicken and egg dilemma for internet addiction and hikikomori. *Current Opinion in Psychiatry,* 33(3), 264-270.

Kessler, R.C., Bossarte, R. M., Luedtke, A., et al. (2020). Suicide prediction models: a critical review of recent research with recommendations for the way forward. *Mol Psychiatry,* 25(1), 168-179.

Kim, S. Y., De Vries, R. G., & Peteet, J. R. (2016). Euthanasia and assisted suicide of patients with psychiatric disorders in the Netherlands 2011 to 2014. *JAMA Psychiatry,* 73(4), 362-368.

Kulshreshth, A., Anand, A., & Lakanpal, A. (2019). Neuralink-An Elon Musk Start-up Achieve symbiosis with Artificial Intelligence. *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS).* IEEE, 105-109.

Kumar, R., Ojha, A. K., Lahiri, B., Zampieri, M., Malmasi, S., Murdock, V., & Kadar, D. (2020, May). Proceedings of the second workshop on trolling, aggression and cyberbullying. *In proceedings of the Second Workshop on Trolling, Aggression and Cyberbullying.*

John, A., Pirkis, J., Gunnell, D., Appleby, L., & Morrissey, J. (2020). Trends in suicide during the covid-19 pandemic. *Bmj*, *371*.

Lakeman, R., & FitzGerald, M. (2009). The ethics of suicide research: The views of ethics committee members. *Crisis,* 30(1), 13-19.

Large, M. M. (2018). The role of prediction in suicide prevention. *Dialogues Clin Neurosci,* 20(3), 197-205.

La Roche, M. J., Fuentes, M. A., & Hinton, D. (2015). A cultural examination of the DSM-5: Research and clinical implications for cultural minorities. *Professional Psychology: Research and Practice, 46, 183*.

Latha, K., Meena, K. S., Pravitha, M. R., Dasgupta, M., & Chaturvedi, S. K. (2020). Effective use of social media platforms for promotion of mental health awareness. *Journal of education and health promotion*, 9.

Lavis, A., & Winter, R. (2020). Online harms or benefits? An ethnographic analysis of the positives and negatives of peer-support around self-harm on social media. *Journal of child psychology and psychiatry*, *61*(8), 842-854.

Lederman, R., & D'Alfonso, S. (2021). The Digital therapeutic alliance: Prospects and considerations. *JMIR Mental Health*, *8*(7), e31385.

Leenaars, A., Cantor, C., Connolly, J., et al. (2002). Ethical & legal issues in suicidology: International perspectives. *Archives of Suicide Research,* 6(2), 185-197.

Lejeune, A., Le Glaz, A., Perron, P., Sebti, J., Baca-Garcia, E., Walter, M., Berrouiguet, S. (2022). Artificial intelligence and suicide prevention: A systematic review. *European Psychiatry,* 1-22.

Luoma, J. B., Martin, C. E., & Pearson, J. L. (2002). Contact with mental health and primary care providers before suicide: a review of the evidence. *The American journal of psychiatry*, 159(6), 909-916.

Lu, S., Li, W., Oldenburg, B., Wang, Y., Jorm, A. F., He, Y., & Reavley, N. J. (2020). Cultural adaptation of the mental health first aid guidelines for assisting a person at risk of suicide to China: a Delphi expert consensus study. *BMC psychiatry*, *20*(1), 1-11.

Madsen, T., Erlangsen, A., Orlovska, S., et al. (2018). Association between traumatic brain injury and risk of suicide. *JAMA*, 320(6), 580-588.

Maher, J. (2017). What troubles me as a psychiatrist about the physician assisted suicide debate in Canada. *Journal of Ethics in Mental Health*, 10, 1-5.

Mann, J. J., Apter, A., Bertolote, J., et al. (2005). Suicide prevention strategies: a systematic review. *JAMA*, 294(16), 2064-2074.

Marks, M. (2019). Artificial Intelligence Based Suicide Prediction. *Yale Journal of Health Policy, Law, and Ethics, Forthcoming*, 98.

Martin, J., LaCroix, J. M., Novak, L. A., et al. (2020). Typologies of Suicide: A Critical Literature Review. *Arch Suicide Res*, 24(1), 25-40.

Maskill, C., Hodges, I., HealthSearch A, et al. (2005). Explaining patterns of suicide. *A selective review of studies examining social, economic, cultural and other population-level influences.* Ministry of Health.

Massing-Schaffer, M., & Nesi, J. (2020). Cybervictimization and suicide risk in adolescence: An integrative model of social media and suicide theories. *Adolescent Research Review*, 5(1), 49-65.

McHugh, C. M., Chun Lee, R. S., Hermens, D. F., et al. (2019a). Impulsivity in the self-harm and suicidal behavior of young people: A systematic review and meta-analysis. *J Psychiatr Res*, 116, 51-60.

McHugh, C. M., Corderoy, A., Ryan, C. J., et al. (2019b). Association between suicidal ideation and suicide: meta-analyses of odds ratios, sensitivity, specificity and positive predictive value. *BJPsych open*, 5(2).

Millington, B. (2017). Health: An optimal commodity for the attention economy. *American journal of public health*, 107(11), 1696.

Milner, A., Sveticic, J., & De Leo, D. (2013). Suicide in the absence of mental disorder? A review of psychological autopsy studies across countries. *International journal of social psychiatry*, 59(6), 545-554.

Ministry of Health. 2019. *Every Life Matters - He Tapu te Oranga o ia Tangata: Suicide Prevention Strategy 2019–2029 and Suicide Prevention Action Plan 2019–2024 for Aotearoa New Zealand*. Wellington: Ministry of Health.

Mueller, A. S., Abrutyn, S., Pescosolido, B., & Diefendorf, S. (2021). The social roots of suicide: Theorizing how the external social world matters to suicide and suicide prevention. *Frontiers in psychology*, 763.

Mullins, B. (2022). AI, Super Intelligence, and the Fear of Machines In Control. *The Cyber Defense Review*, 7(2), 67–76.

Nock, M. K., Borges, G., Bromet, E. J., et al. (2008a). Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. *Br J Psychiatry*, 192(2), 98-105.

Nock, M.K., Borges, G., Bromet, E.J., et al. (2008b). Suicide and suicidal behavior. *Epidemiol Rev*, 30(1), 133-154.

Nock, M. K., Ramirez. F., & Rankin, O. (2019). Advancing our understanding of the who, when, and why of suicide risk. *JAMA Psychiatry,* 76(1), 11-12.

Obegi, J. H. (2019). Rethinking Suicidal Behavior Disorder. *Crisis*, 40(3), 209-219.

Orr, M., Van Kessel, K., & Parry, D. (2022). The ethical role of computational linguistics in digital psychological formulation and suicide prevention. In Proceedings of the Eighth Workshop on Computational Linguistics and Clinical Psychology (pp. 17-29).

Paterson, R., Durie, M., Disley, B., Rangihuna, D., Tiatia-Seath, J., & Tualamali'i, J. (2018). *He Ara Oranga: Report of the government inquiry into mental health and addiction. Wellington, New Zealand*.

Paat, Y. F., & Markham, C. (2021). Digital crime, trauma, and abuse: Internet safety and cyber risks for adolescents and emerging adults in the 21st century. *Social Work in Mental Health*, *19*(1), 18-40.

Perwitasari, D. R., & Wuryaningsih, E. W. (2022). Why did You do that to Me?: a Systematic Review of Cyberbullying Impact on Mental Health and Suicide Among Adolescents. *NurseLine Journal*, 7(1), 35-47.

Pisani, A. R., Murrie, D. C., & Silverman, M. M. (2016). Reformulating Suicide Risk Formulation: From Prediction to Prevention. *Acad Psychiatry*, 40(4), 623-629.

Price, W. N., Gerke, S.,& Cohen, I. G. (2019). Potential liability for physicians using artificial intelligence. *JAMA*, 322(18), 1765-1766.

Pompili, M., Serafini, G., Innamorati, M., et al. (2012). Car accidents as a method of suicide: a comprehensive overview. *Forensic science international*, 223(1-3), 1-9.

Pridmore, S. (2015). Mental disorder and suicide: a faulty connection. *Aust N Z J Psychiatry*, 49(1), 18-20.

Reinares-Lara, E., Olarte-Pascual, C., & Pelegrín-Borondo, J. (2018). Do you want to be a cyborg? The moderating effect of ethics on neural implant acceptance. *Computers in human behavior*, 85, 43-53.

Resnik, P., Foreman, A., Kuchuk, M., et al. (2020). Naturally occurring language as a source of evidence in suicide prevention. *Suicide and Life-Threatening Behavior*, 51(1), 88-96.

Rogers, M. L., Galynker, I., Yaseen, Z., DeFazio, K., & Joiner, T. E. (2017). An overview and comparison of two proposed suicide-specific diagnoses: Acute suicidal affective disturbance and suicide crisis syndrome. *Psychiatric Annals*, *47*(8), 416-420.

Sara, G. E. (2015). Mental disorder and suicide: a faulty connection, or a faulty argument? *Aust N Z J Psychiatry*, 49(1), 84-86.

Savoldi, B., Gaido, M., Bentivogli, L., Negri, M., & Turchi, M. (2021). Gender bias in machine translation. *Transactions of the Association for Computational Linguistics*, 9, 845-874.

Scherr, S., Arendt, F., & Haim, M. (2022). Algorithms without frontiers? How language-based algorithmic information disparities for suicide crisis information sustain digital divides over time in 17 countries. *Information, Communication & Society*, 1-17.

Schlichthorst, M., Reifels, L., Krysinska, K., et al. (2020). Trends in suicide-related research in Australia. *Int J Ment Health Syst,* 14(1), 2.

Schuck, A., Calati, R., Barzilay, S., et al. (2019). Suicide Crisis Syndrome: A review of supporting evidence for a new suicide-specific diagnosis. *Behavioral sciences & the law*, 37(3), 223-239.

Sharma, A., Lin, I. W., Miner, A. S., Atkins, D. C., & Althoff, T. (2023). Human–AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nature Machine Intelligence*, 1-12.

Shing, H-C., Resnik, P., & Oard, D. W. (2020). A Prioritization Model for Suicidality Risk Assessment. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics.* 8124-8137.

Shneidman, E. S. (1993). Commentary: Suicide as psychache. *Journal of Nervous and Mental Disease*, 181(3), 145-147.

Simon, O. R., Swann, A. C., Powell, K. E., et al. (2001). Characteristics of impulsive suicide attempts and attempters. *Suicide Life Threat Behavz*, 32(1), 49-59.

Suman, L. N. (2018). Cyber Trauma: An Overview. *Indian Journal Of Clinical Psychology*, *45*(1), 7-17.

Tait, G., & Carpenter, B. (2015). Suicide, statistics and the coroner: A comparative study of death investigations. *Journal of Sociology,* 51(3), 553-565.

Thorn, P., Hill, N. T., Lamblin, M., Teh, Z., Battersby-Coulter, R., Rice, S., ... & Robinson, J. (2020). Developing a suicide prevention social media campaign with young people (The# Chatsafe project): co-design approach. *JMIR mental health*, *7*(5), e17520.

Tong, F., Lederman, R., D'Alfonso, S., Berry, K., & Bucci, S. (2022). Digital Therapeutic Alliance With Fully Automated Mental Health Smartphone Apps: A Narrative Review. *Frontiers in Psychiatry*, 13, 819623.

Torok, M., Han, J., Baker, S., et al. (2020). Suicide prevention using self-guided digital interventions: a systematic review and meta-analysis of randomised controlled trials. *The Lancet Digital Health*, 2(1), e25-e36.

Tremain, H., McEnery, C., Fletcher, K., & Murray, G. (2020). The therapeutic alliance in digital mental health interventions for serious mental illnesses: narrative review. *JMIR mental health*, *7*(8), e17204.

Turecki, G., & Brent, D. A. (2016). Suicide and suicidal behaviour. *Lancet*, 387(10024), 1227-1239.

van Dis, E. A., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: five priorities for research. *Nature*, *614*(7947), 224-226.

Woodford, R., Spittal, M. J., Milner, A., et al. (2019). Accuracy of Clinician Predictions of Future Self-Harm: A Systematic Review and Meta-Analysis of Predictive Studies. *Suicide Life Threat Behav,* 49(1), 23-40.

World Health Organization. (2014). *Preventing suicide: A global imperative.* World Health Organization.

Yaseen, Z. S., Hawes, M., Barzilay, S., et al. (2019). Predictive validity of proposed diagnostic criteria for the suicide crisis syndrome: an acute presuicidal state. *Suicide and Life-Threatening Behavior,* 49(4), 1124-1135.

Yeager, C. M., & Benight, C. C. (2022). Engagement, Predictors, and Outcomes of a Trauma Recovery Digital Mental Health Intervention: Longitudinal Study. *JMIR Mental Health*, *9*(5), e35048.

Xie, T., & Pentina, I. (2022). Attachment Theory as a Framework to Understand Relationships with Social

Chatbots: A Case Study of Replika. Proceedings of the 55th Hawaii International Conference on System Sciences.

Acknowledgements: None

Competing Interests: None

Email address for correspondence: martinorr521@gmail.com

Date of publication: Aug 7, 2023