Health Justice, Resource Allocation, and Age in Suicide Risk Assessment

James Hart (Corresponding Author)
James.Hart@ethox.ox.ac.uk
Postdoctoral Research Fellow
Ethox Centre, University of Oxford, Oxford, UK

Sapfo Lignou
Sapfo.Lignou@ethox.ox.ac.uk
Senior Researcher in Bioethics
Ethox Centre, University of Oxford, Oxford, UK

Seena Fazel
Seena.Fazel@psych.ox.ac.uk
Director, Centre for Suicide Research and Professor of Forensic Psychiatry
Department of Psychiatry, University of Oxford, Oxford, UK
& Oxford Health NHS Foundation Trust, Oxford, UK

Abstract

Suicide risk assessment tools and prediction models aim to assess individual suicide risk in order to assist clinical decision-making and improve health outcomes. Whilst these tools have proliferated in recent years, concerns and criticisms of them have also increased and NICE recommended against their use. In this paper, we discuss these tools from the novel perspective of health justice. Firstly, we argue that, in the context of limited resources, having tools to assist in identifying the highest risk individuals may help us to prioritise the worst-off and distribute resources more fairly. Secondly, we argue that, where there are known inequalities in suicide risk and related outcomes between groups, not prioritising according to risk threatens to indirectly discriminate against higher risk groups. Particularly salient is the case of older populations who have much higher risks of suicide after self-harm than younger populations, and who already experience inequalities in access to mental health care. We conclude by making some recommendations for future research.

Introduction

Suicide, defined as a death caused by intentional self-directed injury, is a major, yet preventable, disease burden globally with significant accompanying economic costs. [1] [2] [3] Thus, it is being increasingly recognised as a major public health priority.[4] Suicide risk assessment plays a significant role in suicide prevention and is a central part of individual medical care. In general hospital emergency departments, this is mainly done for all admissions for self-harm incidents, such as overdoses. Such assessments are routine as suicide rates are considerably elevated after self-harm, with increased odds estimated to be ten-fold.[5] Furthermore, longitudinal studies have reported that around 4-8% die from suicide within five years after presenting to hospital with self-harm. [6] [7] In mental health, these assessments inform decisions about treatment, admission to psychiatric hospital, and the frequency and intensity of follow-up care. Many mental disorders are associated with substantially higher risks of suicide compared to general population comparisons[8] especially after discharge from inpatient settings (100 times the general population rate)[9] [10]. Risk is further elevated after recent suicidal behaviour, with significant heterogeneity across disorders – especially for schizophrenia-spectrum disorders (with elevated odds around nine-fold of suicide), other nonorganic psychosis (around seven-fold increased odds) and depression (three-fold increased odds)[11].

Risk assessment is typically done using clinical judgement, which can be augmented by the use of structured tools. These tools have historically taken the form of symptom checklists,[12] which were not developed for the purposes of risk assessment, and hence their accuracy and usefulness has been questioned.[13] Another potential challenge, in addition to the use of these checklists, is that they are used for all suicide outcomes (ideation, self-harm, and suicide mortality) whereas the research evidence suggests that risk markers are different. One clear example is the effects of age and sex at birth – these are different for suicidal ideation and repeat self-harm (higher risks for younger women[14]) than suicide (with elevated risks for older men[15]) – suggesting that risk

assessments for suicide and for self-harm can be differentiated, so they can be sensitive to these factors.

Despite the problems with older checklists, there are advantages to structured tools. This includes consistency between and within services, which has been highlighted as a major problem.[16] More importantly is that clinical decision-making varies widely[17], and may not incorporate evidence about risk factors, their relative weights, overemphasising predisposing factors[18], and how they might interact. Furthermore, general evidence suggests that structured tools may help mitigate human biases, especially recency bias and availability of information. [19]

More evidence-based risk assessment tools have proliferated in recent years, using prediction models to capture a wider array of these considerations. Unlike prior checklists, prediction models combine data from multiple sources using robust statistical methods to generate more precise probability scores. Proponents of these tools argue that these models may thus be more accurate and reliable than previous structured tools whilst improving consistency between and within services. [20] Moreover, if they are more accurate and reliable, proponents argue that they may allow better targeting of limited resources[21] and improve the appropriateness, efficacy and efficiency of treatments.[22]

Nevertheless, concerns remain about the use of these models, with critics pointing to their low accuracy as indicated by positive predictive value (PPV), the use of arbitrary thresholds, and the current paucity of evidence as to their effectiveness in clinical settings.[23] [24] [25] The National Institute for Health and Care Excellence (NICE), the UK public body in charge of health technology assessment, agreed with the critics, and in their otherwise comprehensive, helpful, and evidence-based self-harm guidelines recommended against the use of these tools, raising concerns that tools might be used to inappropriately deny care. “The [NICE] committee agreed based on their experience that *denying care on this basis* could lead to repeat self-harm, distress, and lower service

user satisfaction, and therefore tools and scales should not be used for this purpose either. [Our Italics]"[26]

In what follows, we re-appraise these tools from the novel perspective of health justice and resource allocation. We suggest that NICE overlook some of the justice implications of these tools for care prioritisation, that their judgement should be revisited, and that more evidence ought to be gathered about the clinical effectiveness of these tools in practice.

The Context of Resource Allocation

The NICE committee, in their quote above, is right that the denial of care based on a suicide risk assessment could lead to various patient harms. However, this concern must be understood within a broader context. The denial of care *on any basis* could lead to adverse patient outcomes. There is no reason to think that denial of care on the basis of suicide risk assessment tools would be worse for an individual patient than denial of care on the basis of clinician judgement, risk (or safety) formulation or any other basis. Where universal interventions ensuring a standard level of care are possible and appropriate, these should be provided to all patients without risk stratification. Yet, given the context of limited healthcare resources it is inevitable that some care will be denied. Higher resource treatments cannot be provided to everyone, and waiting lists functionally deny care to many when they most need it. Thus, concerns about these tools must be understood in the context that all triage systems will inevitably misclassify some patients and consume resources. The relevant questions are not therefore whether structured tools and instruments are flawed, but if they are worse than existing alternatives at identifying which individuals should be denied care, or whether they increase the amount of care that is denied.

Unfortunately, mental healthcare resources in the UK and many other countries are already at capacity and so many who could benefit from treatment are not receiving it, are placed on long waiting lists, or receive sub-optimal care packages.[27] [28] In this context, it seems unlikely that risk

assessment tools will increase the amount of care that is denied. Thus, whilst NICE are right to highlight limitations of current tools and the existing evidence base up to 2022, they seem to overlook this vital context in their assessment of whether and when these tools may be useful. If NICE decide that everyone should receive funding for personalised treatment (determined without the use of risk assessment tools), then this is probably appropriate. However, such an assessment is no guide for decision makers when they are in the difficult position to need to prioritise patients for treatment.

Similarly, other calls to focus on person-centred approaches (which involve building a 'therapeutic alliance' with the patient to inform assessment and collaboratively develop a care plan based on their individual needs)[29] may be appropriate for deciding on treatment options for patients, but person-centred approaches do not speak to issues of resource allocation and patient prioritisation. These issues are by nature not about best individual care, but about how to distribute care amongst individuals. It is uncomfortable to think about distributing or rationing mental health care, but the reality is that systems already do so and they do not have enough resources not to.

Resource allocation decisions, especially at more local levels, are not just about what care to fund, how much to fund, or who is eligible for treatment. The core issue is that, due to resource constraints, not all patients can receive the best treatment. In this context, patient prioritisation is unavoidable and we already prioritise care based on assessments of need and risk. Except these assessments are made independently by clinicians and clinical services. In this context, it is implausible to think that risk assessment tools would increase the amount of care that is denied.

What remains is an important distinction between two questions about risk assessment tools. One is whether risk assessment tools are good at determining best treatment options for an individual. The other is about whether using risk tools to assist patient prioritisation is better than current priority-setting decision-making. NICE answer the first question in the negative, but the equally important

second question has been overlooked. If the answer to it is yes, then the tools perhaps ought to be used in those contexts.

Different approaches might thus be used for different purposes: patient centred assessment for care plans or for therapeutic purposes, and risk assessment tools for patient prioritisation. These tools have been evaluated for their predictive accuracy but not specifically for their priority-setting effectiveness (i.e., how they make use of limited resources). Importantly, this will require evaluating the effectiveness of current processes using the same measures, rather than merely assuming that low predictive accuracy and sensitivity in risk assessment tools means they would be worse than current methods. We consider the issue of accuracy in more detail in the penultimate section.

Moreover, we also need to consider several fairness and inequalities issues when thinking about this second question. We now highlight some of these issues and the ways in which risk assessment models may be better suited at addressing them than clinical judgement alone.

Age-Related Inequalities and Indirect Discrimination

The first justice issues that NICE overlooks are those regarding health inequalities. We know that there are significant health inequalities when it comes to suicide and suicide risk. There are significant variations in suicide rate according to social characteristics, including age, sex and deprivation.[30] These factors interact in complex ways, complicating the justice issues, but for the purposes of this paper we have decided to focus on just age. Age is a particularly good exemplar because it is easily, routinely and reliably measured, and because disparities in suicide risks are substantial. Whilst suicide is a leading cause of mortality in younger populations, mortality from suicide remains substantially higher in older populations. [31 32] Despite this, older people are less likely to be referred to specialist services, suggesting that current clinician-led priority-setting is ineffective and unjust.[33 34 35] NICE themselves recognise that older people have higher rates of suicide after self-harm than younger people, and advise that healthcare professionals take this into account when carrying out a psychosocial assessment.[36] However, this is the limit of NICE's recognition of this

substantial inequality. There is no further suggestion about how this inequality should inform decision-making, particularly patient prioritisation. In our opinion, this might amount to indirect discrimination.

According to the UK Equality Act 2010: "Indirect discrimination occurs when a policy which applies in the same way for everybody has an effect which particularly disadvantages people with a protected characteristic."[37] For instance, using the same BMI threshold for weight classification and thus for accessing related care may indirectly discriminate against people from certain ethnic minority backgrounds for whom cardiometabolic risk is higher at lower BMIs. For this reason, NICE recommend that lower BMI thresholds are used for people from those backgrounds.[38] (Another plausible example of indirect discrimination is the lack of prostate cancer screening in the UK where there is a particularly high prevalence of prostate cancer in older men.)[39]

If groups with different suicide risk profiles are treated the same, or given equal priority, then this too might amount to indirect discrimination. Assuming that resources are limited and it is not possible to provide gold-standard treatment for everyone, treating all groups the same could disadvantage those at higher risk. Where older people have significantly higher risks of suicide after self-harm than younger people, then providing or denying care at similar rates will disadvantage the older population. This is clearly the case after self-harm where the rates of suicide mortality in the 12 months after presenting to English emergency departments in men and women aged 55 and over are 1874 and 1282 per 100,000 person-years respectively, compared with 309 in men and 111 in women aged 15-24. This equates to a 6-fold higher risk for older versus younger men and an 11-fold difference in risk between older and younger women (and a 17-fold difference between younger women and older men).[40] If we do not acknowledge these underlying risks, some younger individuals with lower risks will get treatment whilst other older individuals with higher risks will not.

In theory, risk assessment instruments are not necessary to prevent indirect discrimination, as long as these disparities in risks are accommodated in clinical judgements and priority-setting procedures. But

the reality is that, in practice, these disparities are unlikely to feature in clinical decision-making consistently or to the full extent of their impact on the groups effected. In fact, evidence shows that clinicians are susceptible to ageist stereotypes, demonstrating a bias against older patients, with older patients less likely to receive correct diagnosis and treatment compared to younger patients.[41] [42] Expecting clinicians to internalise and accurately incorporate older patient's higher suicide risks into their assessments in a way that fully compensates for their disadvantage, thus seems exceedingly unlikely to achieve equality. Conversely, risk assessment tools might provide a powerful way in which clinicians can consistently ensure the underlying suicide risk for particular groups is appropriately accounted for in all decision making.

Nevertheless, indirect discrimination can be legally justified by policymakers if they have good reasons to do so. For instance, if it were decided that any differentiation in care would increase stigma around mental health in older people, then this could be a legitimate reason not to differentiate. However, no such reasons seem to have been given by decision-makers, and we think none are likely to be strong enough, given the extent of the inequalities. In the specific case of NICE, their Guidance Equality Impact Assessments did not consider how these equalities issues might affect priority-setting.[43] In our opinion, this is a failure in procedure, one that plausibly amounts to a procedural injustice. Procedural justice requires that decision makers consider all relevant considerations, including health inequalities, when making priority-setting decisions.[44] [45] Failure to do so is unfair because it ignores the reasons that favour treating certain patients, thus disadvantaging them in the decision-making process. In this case, by not considering whether older patients should be prioritised care based on their higher suicide risk, NICE pre-emptively disadvantages older populations in the competition for limited mental health care resources.

Importantly, there is also legal precedent on this issue. In 2007, Eisai, a pharmaceutical company, took NICE to court accusing them of indirect discrimination.[46] [47] The case, *Eisai vs NICE*, concerned a drug, Aricept (donepezil), which was shown to have mild benefits to mental function in Alzheimer's

Dementia. Due to the mild benefit, NICE deemed that the drug should only be funded for patients with 'moderate severity' Alzheimer's, determined by cognitive test scores. However, Eisai argued that using cognitive test scores in this manner would indirectly discriminate against individuals whose test scores might be lower for reasons unrelated to their dementia. In particular, individuals with learning difficulties and non-native English speakers.

NICE conceded that their guidance could be discriminatory in theory, but argued that in practice clinicians were expected to use their own judgements when using the guidelines. NICE claimed that this flexibility in subsidiarity would ensure 'atypical' groups like these would still be eligible for funding, and thus there would be no discriminatory impact. However, the High Court disagreed, ruling that NICE's approach was unsatisfactory and discriminatory. Justice Dobbs concluded that NICE did not sufficiently consider their positive duties to promote equality and eliminate discrimination and instead "[left] it to others to sort out in the hope and expectation that they would".[48]

There are strong parallels between the indirect discrimination found in Eisai vs NICE and the current guidance regarding suicide risk assessment tools. Whilst NICE does make clear that mental health professionals should recognise the higher rates of suicide amongst older people after self-harm, this is the limit of their consideration of this important age-related disparity, seemingly outsourcing responsibility to clinicians once again. Importantly, they neglect to consider how their own policies might help address these inequalities and whether risk assessment tools may play a role in ensuring older individuals are not disadvantaged in priority-setting decisions.

Fairness in clinical patient prioritisation

Lastly, there are some wider fairness considerations for patient prioritisation that are worth touching on. For a patient prioritisation process to be fair requires a number of features. Particular important are: sensitivity to relevant considerations, lack of bias, consistency, transparency and

accountability.[49] For each of these there is good reason to think that decisions informed by risk assessment tools may fare better than clinical judgement alone.

Firstly, any fair patient prioritisation process should be sensitive to relevant considerations and, importantly, be insensitive to irrelevant considerations. For instance, patient prioritisation should be sensitive to a patient's risk of harm and ability to benefit from treatment, but obviously should not be sensitive to a patient's looks or character. To be clear, risk-assessment tools do not capture every relevant prioritisation consideration (for instance, risk-assessment tools do not tell us directly about an individual's capacity to benefit from treatment), and so they should not by themselves decide prioritisation decisions. Nevertheless, in so far as we ought to prioritise patients primarily by need, suicide risk assessments reflect a dimension of need that should not be discounted. This is especially true when we know that risks between different subgroups can differ by as much as 17 times.

Secondly, we have good reason to think that clinicians are susceptible to a large number of cognitive and implicit biases that might affect how they assess suicide risk and prioritise care.[50 51] Some, like base rate fallacies, are failures to appropriately account for evidence, whilst others, like unconscious biases and stereotypes, are more clearly issues of justice. The intersection of multiple social markers for suicide risk may make these biases worse and adjusting for them especially challenging. Of course, risk-assessment tools are not going to be able to eliminate all of these biases as the tools themselves are developed using historical data and thus may reflect past inequalities. Nevertheless, risk-assessment tools will not be susceptible to the same array of biases that clinicians are, and they can be more closely inspected to reduce biases. Because new tools are based on combinations of risk markers (in multivariable models), they may also provide particularly helpful insight about the relative contribution of different social markers. Thus, if they are used to inform, guide and augment patient prioritisation decisions rather than replace clinical judgement, it is plausible that they may reduce overall bias and improve fairness in decision making.

In addition, tools can be developed using statistical methods to adjust for potential disparities or to not include clearly more biased predictors (e.g., ethnicity). Here proxies for these excluded variables, such as income or postcode, might also be problematic, and if so, they require testing to investigate any systematic over, or under, estimation of risks. Thus, as tools improve, we might reasonably expect better subgroup specific predictive performance – as has occurred in cardiovascular risk assessment[52] – likely reducing bias.

Relatedly, risk-assessment tools ought to improve the consistency of decision-making. Fairness requires at least some level of consistency between decisions. If some patients are being provided or denied care according to some criteria, and others are being provided or denied care according to different criteria, we end up with inevitable disparities in care; equally situated people may receive very different risk formulations, treatments, and prioritisation. Clearly some clinical judgement will always be necessary and we should accept some differences in judgement that arise from appropriate subsidiarity. Nevertheless, this scope is not unlimited and where risk-assessment tools can improve consistency of decision-making between clinicians by providing information (based on a uniform analysis of underlying risk factors) upon which clinicians can base their decisions, fairness is also likely to be improved.

Lastly, fairness requires that decision-making processes are transparent and can be held accountable. If an individual is denied care, provided with lower standard care, or put on a long waiting list because of a patient prioritisation decision, the reasons why should be available to them, and if there are issues with the way that decisions are made, there should be ways to challenge and improve them. Clinicians may give reasons for their judgements, but it will always be hard to inspect exactly how they came to those decisions and to hold clinicians to account; they are, in some sense, a black-box. This will be especially true for a clinician's assessments of risk for low prevalence outcomes, where inaccurate risk predictions can easily be missed. By contrast, risk assessment tools can, at least in principle, have their judgements more easily inspected (though this may not be true

in the case of AI models – which raise further interesting ethical questions). Importantly, risk-assessment models should thus also be capable of being improved over time, in a way that clinical judgement by itself is unlikely to.

Criticisms

Having considered these tools from a justice perspective, it is worth briefly examining some of the main criticisms of the tools. There are two primary concerns, both of which seem operative in NICE's judgement. First, critics question the statistical accuracy and reliability of these tools in predicting suicides. Second, there are concerns that prediction models may categorise individuals by risk using arbitrary thresholds and imply that an individual's risk is static rather than dynamic. We argue that neither of these concerns undermines the case for their use, especially in the context of resource allocation.

Accuracy

At the crux of the issue are questions about the predictive accuracy of these models. Many large studies and meta-analyses have shown that our ability to predict suicide remains poor. Prediction tools have been shown to have low measures of classification, including positive predictive values (PPV) and sensitivity.[53 54 55 56 57] In other words, most individuals classified as high risk do not go on to attempt suicide, and most suicides occur in those individuals classified as lower risk ('the low-risk paradox'). Moreover, the association between single clinical risk factors and suicide is weak, with low certainty of evidence and significant heterogeneity between studies. [58 59 60 61 62] Thus, models can give wrong predictions, particularly if based on very few risk markers, and so on the face of it, seem to be ineffective. However, the plausible seeming leap from low accuracy to ineffectiveness comes too quickly.

It is worth mentioning that the accuracy of these tools is improving all the time. For instance, the recent OxSATS model has high levels of discrimination (as reported by AUC statistic) and good

calibration (difference between observed and predicted risk estimates).[63] Secondly, it is important to note that for rare outcomes, like suicide, the PPV of any predictive model will inevitably be low (although it is threshold-dependent). Though low prevalence, and thus low absolute risk of suicide, does itself further complicate any analysis of the usefulness of these tools. Similarly, the low sensitivity of some models can partially be explained by the chosen threshold (although the OxSATS model had a high sensitivity of 82% at the prespecified cut-off). Because of this, neither low PPV nor low sensitivity need necessarily indicate an ineffective tool. Depending on the context and use, a model with these features may still be helpful.

More important however, is the comparison. A prediction model may not be that accurate, but if it is more accurate than its alternatives, we should reasonably expect it to be more effective than its alternatives (see Box 1). A prediction model may still be better at picking out higher risk individuals, than clinical judgement or suicide risk screening questions alone.[64] [65] [66] The accuracy of clinical judgement alone has been reviewed using sensitivity and PPV for future self-harm. Sensitivity was reported at 31%, i.e., 31% of patients who had a repeat self-harm episode were identified by clinicians as high risk.[67] Interestingly, this is lower than many structured tools, and considerably worse than modern tools. For example, a recent risk tool for repeat self-harm (OxSET) reported sensitivities of 52-57% in external validations based on a prespecified cut-off.[68]

Box 1

The accuracy of any predictive model or tool cannot determine its usefulness without consideration of its use case. To see how a low accuracy tool can still be effective consider the following toy example:

There is a population of 1000 with Disease X. 1% of this population (10 individuals) will die without treatment. But there is only enough treatment for 10% (100 individuals).

There is a low accuracy test available. The test only has 30% sensitivity, and a positive predictive value of 3%. This means that if the test is used to assign treatment, three lives will be saved in expectation (i.e., the weighted average of all possible outcomes is three lives saved). This seems low, but if treatment is assigned randomly, only one life will be saved in expectation.

In the context of low prevalence outcomes and resource constraints, low sensitivity tools can thus still prove effective.

However, as NICE notes there is a lack of evidence as to the clinical effectiveness of these tools in practice, with risk assessments only compared to each other.[69] Importantly, this is a lack of evidence and not evidence of lack. Despite this, NICE did not use this opportunity to recommend more research on their effectiveness in prioritisation and their comparison to alternative approaches.

Instead, they recommended, based on their own professional experience, that collaborative risk formulations should be taken between the mental health professional and the patient. Unlike tools such as the ASQ[70] and SAFE-T[71] used in the US, risk formulations are not intended as screening or triage tools, but rather intended to help form an understanding of an individual's risks and difficulties to inform their treatment plan. This appropriately emphasises shared, patient-orientated, and multidisciplinary assessment for the development of care plans, but overlooks questions of prioritisation and triage. Thus, it presents a false dichotomy between the use of risk assessment tools and risk formulations whilst also assuming that risk formulations would prove more effective than risk assessments. Yet, as they themselves note, we do not have evidence of their relative effectiveness; the lack of evidence in this case cuts both ways. Without comparative evidence, we cannot conclude that one approach better promotes patient safety or fairness compared to the other. If risk formulations are themselves poor guides to prioritisation, then the low accuracy of risk models is no mark against them.

Moreover, context and use-case will, again, matter in determining their effectiveness. It is possible, and in our view plausible, that risk formulations are more helpful for informing treatment plans, whilst prediction models may be more effective in priority-setting. Nor do these models need to replace clinician-led decision-making. Instead, they may complement or augment decision-making, and there is significant scope to determine how much reliance clinicians should put in the models when making decisions.[72] Importantly, this would not be unprecedented. Whilst we need to be cautious with comparisons with physical health due to differences in complexity and in the quality and nature of predictors, predictive models in other areas of medicine have demonstrated their

value in supporting clinical decision-making under uncertainty. For example, in cardiovascular medicine, risk scores have been shown to improve decision-making and increase treatment access by guiding prescribing decisions more effectively.[73] This suggests that, rather than dismissing predictive models due to concerns over accuracy, we should consider their potential utility in improving resource allocation and clinical decision-making.

Thresholds

The other main criticism focuses on how many models categorise individuals into high, medium, or low risk, based on arbitrary thresholds.[74] [75] Thus, individuals whose risk scores are similar but fall either side of a threshold will be categorised differently and could consequently receive different treatment. Moreover, there are worries these categorisations can imply that someone's risk is static rather than dynamic, and so discourage clinicians from being sensitive to an individual's presenting problems.[76] This may especially be the case if risk assessments are based on subjective questionnaires of individuals' feelings and symptoms.

These criticisms may be accurate for less well-developed tools and those based on checklists, however they imply that all models must only be classifiers (i.e., have cut-offs) and will be used in these particular ways. Risk tools in the rest of medicine have been robustly developed following the Framingham, SCORE, QRISK or STENO cardiovascular risk models, diabetes risk and also cancer mortality risk scores, where they provide a probability score rather than categorisation.[77] [78] [79] [80] [81] [82] [83] Probability scores are not going to be subject to the same arbitrariness and fine-grained distinctions in analysis and treatment. For instance, someone who scores 9.9% in the QRISK will be considered for treatment, even if current guidelines say you should consider a statin for individuals scoring above 10%.

This approach should also mitigate against the assumption of static risk, especially if these models are not used to replace clinical judgement but to complement and inform clinical judgement. Moreover, models need not be based on subjective questionnaires, but on more objective risk

markers that will not be sensitive to misreporting nor overemphasise patient experiences and attitudes at a particular time slice. These factors ought to be seen as base-rate factors; they provide background information on the outcomes of similarly situated patients, upon which clinicians can anchor their broader assessments. This might require more nuanced use of these tools than in current practice, with particular emphasis on not using them in isolation, but questions about implementation should not be mistaken for flaws in the tools themselves. Whilst there is some early evidence on the feasibility of these tools in practice[84 85 86], further research on how these tools can be used in practice to enhance rather than replace clinical decision-making is needed.

Box 2

Criticism 1 - Accuracy:

- Most who are considered high-risk do not attempt suicide (low positive predictive value);
- The majority of suicides occur in those considered low-risk (the "low-risk paradox").

Response:

- Low positive predictive values are expected in low prevalence outcomes and cannot determine usefulness without determining how a tool is used;
- Accuracy has not been compared to clinician judgement alone or 'risk formulations' – predictive models may be more accurate relative to other options;
- These criticisms are dependent on tools using risk categorisation (such as low or high risk), which recent evidence-based tools avoid.

Criticism 2 – Thresholds:

- Arbitrary thresholds may categorise similar risk individuals into different categories;
- They may also imply that individual's risk is static rather than dynamic and discourage sensitivity to patient's presenting needs.

Response:

- Risk models can be developed without thresholds and without using simplistic categories, using probability scores instead;
- Tools can be used to anchor and inform clinician judgement, rather than replace judgement with arbitrary categorisation;
- Clinical judgement can pay insufficient weight to historical and predisposing factors, with over-reliance on recent events.

Conclusion

To be clear, none of what we say in this paper is decisive in favour of these tools, our suggestions instead focus on how NICE can improve their decision-making regarding them (see Box 3 for a summary of our suggestions). We have outlined reasons why these tools may be effective at assisting mental health priority setting and reducing bias (particularly against older individuals) in priority setting, but these theoretical considerations remain largely untested, (though there is some positive early qualitative data[87]). However, this lack of evidence cuts both ways, and so any determinative ruling that these tools should never be used is also premature. Instead, at this juncture, the focus should be on investigating the issues and use cases identified in this paper. Where there is uncertainty, further research and cautious implementation is important rather than categorical exclusion.

In particular, further research could determine the effectiveness (including cost effectiveness) of these tools for patient prioritisation, including at the population level. Analysis of these tools has largely focused on managing individual patients at the clinician level, but other contexts where priority-setting decisions are made should also be included. For instance, could these tools be helpful at the operational level, for determining where to focus staff and resources within, or across, inpatient settings?

Importantly, these tools should also be compared against current clinical practice and clinician judgement alone, rather than - as is mostly the case currently - only against each other. Bodies like NICE should also ensure that evaluation standards for predictive tools are consistent with those applied to other methods. Further investigation into how best to integrate these tools within existing care pathways may also be warranted. Lastly, tools should be tested to see whether they increase (or decrease) the amount of denied care, and if they work as instruments of justice to reduce bias against underdiagnosed and undertreated groups such as the elderly.

Box 3

NICE Decision-making:	Our Suggestions:
Decision primarily made based on the expertise and experience of the committee.	Decision should be made on the basis of a comprehensive body of evidence, which considers predictive performance, acceptability, feasibility and utility, and be proportionate to that evidence.
Systematic review informing NICE decision was limited to impact studies.	Systematic review should be more comprehensive and consider predictive performance and feasibility studies.[88]
Acknowledged a lack of evidence, but did not recommend more research to gather more evidence.	Committee should recommend research to gather more evidence on areas where there is uncertainty.
NICE acknowledged that older people have higher rates of suicide following self-harm, and indicated clinicians should take this inequality into account in decision-making.	NICE should consider its own positive duties to address inequalities in suicide rates, including considering whether and how risk-assessment tools may impact on inequalities.
Focused on non-comparative assessments of tools accuracy and effectiveness.	Should compare accuracy and impact of tools with current practice and alternatives (e.g., relying on unstructured subjective clinical risk assessment).
Assumed risk assessment tools would lead to more inappropriate denial of care.	Committee should consider evidence of risk-assessment tools impact on denial of care, and effectiveness in triage. It should recommend research if there is a lack of evidence, and not assume greater inappropriate denial of care.
Assumed that risk assessment tools use arbitrary thresholds.	Should acknowledge that tools need not use thresholds and that such thresholds, if applied, will vary depending on setting and service.
Assumed that tools would replace or worsen clinician judgement.	Committee should consider how tools can complement and inform clinician judgement.
Risk formulations presented as alternative to risk assessments.	Risk formulations and assessments should not be considered as mutually exclusive options.

Author Contributions

JH and SF contributed equally to the initial conceptualisation of the manuscript. JH wrote the original draft. JH, SL and SF reviewed and edited subsequent drafts.

Competing Interests Statement

SF was part of the research team that developed and validated the OxSET and OxSATS models cited in this article and has co-authored papers on them. JH and SL have no competing interests.

[1] GBD 2021 Suicide Collaborators. Global, regional, and national burden of suicide, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet Public Health*. 2025;10(3):e189-e202. doi:10.1016/S2468-2667(25)00006-4.

[2] Peterson C, Haileyesus T, Stone DM. Economic cost of U.S. suicide and nonfatal self-harm. *Am J Prev Med*. 2024;67(1):129-133. doi:10.1016/j.amepre.2024.03.002.

[3] Samaritans. *The economic cost of suicide in the UK*. March 2024. https://www.samaritans.org/documents/2538/The_economic_cost_of_suicide_in_the_UK_-_web.pdf (accessed September 1st, 2025).

[4] World Health Organization. *Suicide*. March 2025. https://www.who.int/news-room/fact-sheets/detail/suicide (accessed September 1st, 2025).

[5] Favril L, Yu R, Geddes JR, Fazel S. Individual-level risk factors for suicide mortality in the general population: an umbrella review. *Lancet Public Health*. 2023;8(11):e868–e877. doi:10.1016/S2468-2667(23)00207-4.

[6] Demesmaeker A, Chazard E, Hoang A, Vaiva G, Amad A. Suicide mortality after a nonfatal suicide attempt: a systematic review and meta-analysis. *Aust N Z J Psychiatry*. 2021;56(6):603–616. doi:10.1177/00048674211043455.

[7] Olfson M, Wall M, Wang S, Crystal S, Gerhard T, Blanco C. Suicide following deliberate self-harm. *Am J Psychiatry*. 2017;174(8):765–74. https://doi.org/10.1176/appi.ajp.2017.16111288

[8] Fu XL, Qian Y, Jin XH, Yu HR, Wu H, Du L, Chen HL, Shi YQ. Suicide rates among people with serious mental illness: a systematic review and meta-analysis. *Psychol Med.* 2023;53(2):351-361. doi:10.1017/S0033291721001549

[9] Chung DT, Ryan CJ, Hadzi-Pavlovic D, Singh SP, Stanton C, Large MM. Suicide rates after discharge from psychiatric facilities: a systematic review and meta-analysis. *JAMA Psychiatry.* 2017;74(7):694–702. doi:10.1001/jamapsychiatry.2017.1044

[10] Chung D, Hadzi-Pavlovic D, Wang M, et al. Meta-analysis of suicide rates in the first week and the first month after psychiatric hospitalisation. *BMJ Open*. 2019;9:e023883. doi:10.1136/bmjopen-2018-023883.

[11] Haglund A, Lysell H, Larsson H, Lichtenstein P, Runeson B. Suicide immediately after discharge from psychiatric inpatient care: a cohort study of nearly 2.9 million discharges. *J Clin Psychiatry*. 2019;80(2):18m12172. doi:10.4088/JCP.18m12172.

[12] Green KL, Brown GK, Jager-Hyman S, Cha J, Steer RA, Beck AT. The predictive validity of the Beck Depression Inventory suicide item. *J Clin Psychiatry*. 2015;76(12):1683–1686. doi:10.4088/JCP.14m09391.

[13] Fortune S, Hetrick S. Suicide risk assessments: why are we still relying on these a decade after the evidence showed they perform poorly? *Aust N Z J Psychiatry*. 2022;56(12):1529–34. doi:10.1177/00048674221107316.

[14] Fazel S, Vazquez-Montes MDLA, Lagerberg T, Molero Y, Walker J, Sharpe M, et al. Risk of repeat self-harm among individuals presenting to healthcare services: development and validation of a clinical risk assessment model (OxSET). *BMJ Ment Health*. 2024;27:e301180. doi:10.1136/bmjment-2024-301180.

[15] Geulayov G, Casey D, Bale L, Brand F, Clements C, Farooq B, Kapur N, Ness J, Waters K, Tsiachristas A, Hawton K. Suicide following presentation to hospital for non-fatal self-harm in the Multicentre Study of Self-harm: a long-term follow-up study. *Lancet Psychiatry*. 2019;6(12):1021–30. doi:10.1016/S2215-0366(19)30402-X.

[16] Quinlivan L, Cooper J, Steeg S, et al. Scales for predicting risk following self-harm: an observational study in 32 hospitals in England. *BMJ Open.* 2014;4(2):e004732. doi:10.1136/bmjopen-2013-004732
[17] Berman NC, Stark A, Cooperman A, Wilhelm S, Cohen IG. Effect of patient and therapist factors on suicide risk assessment. *Death Stud*. 2015;39(7):433–41. doi:10.1080/07481187.2014.958630.
[18] Pease JL, Forster JE, Davidson CL, Holliman BD, Genco E, Brenner LA. How Veterans Health Administration suicide prevention coordinators assess suicide risk. *Clin Psychol Psychother*. 2017;24:401–10. doi:10.1002/cpp.2011.
[19] Mamede S, van Gog T, van den Berge K, et al. Effect of availability bias and reflective reasoning on diagnostic accuracy among internal medicine residents. *JAMA*. 2010;304(11):1198–1203. doi:10.1001/jama.2010.1276
[20] Whiting D, Fazel S. How accurate are suicide risk prediction models? Asking the right questions for clinical practice. *Evid Based Ment Health*. 2019;22(3):125–8. doi:10.1136/ebmental-2019-300102.
[21] Botchway S, Tsiachristas A, Pollard J, Fazel S. Cost-effectiveness of implementing a suicide prediction tool (OxMIS) in severe mental illness: economic modeling study. *Eur Psychiatry*. 2022;66(1):e6. doi:10.1192/j.eurpsy.2022.2354.
[22] Kessler RC, Bossarte RM, Luedtke A, Zaslavsky AM, Zubizarreta JR. Suicide prediction models: a critical review of recent research with recommendations for the way forward. *Mol Psychiatry*. 2020;25(1):168–79. doi:10.1038/s41380-019-0531-0.
[23] Simpson SA, Goans C, Loh R, Ryall K, Middleton MCA, Dalton A. Suicidal ideation is insensitive to suicide risk after emergency department discharge: performance characteristics of the Columbia-Suicide Severity Rating Scale Screener. *Acad Emerg Med*. 2021;28:621–9. doi:10.1111/acem.14198.
[24] Hawton K, Lascelles K, Pitman A, Gilbert S, Silverman M. Assessment of suicide risk in mental health practice: shifting from prediction to therapeutic assessment, formulation, and risk management. *Lancet Psychiatry*. 2022;9(11):922–928. doi:10.1016/S2215-0366(22)00232-2. PMID: 35952701.
[25] Runeson B, Odeberg J, Pettersson A, Edbom T, Jildevik Adamsson I, Waern M. Instruments for the assessment of suicide risk: a systematic review evaluating the certainty of the evidence. *PLoS ONE.* 2017;12(7):e0180292. doi:10.1371/journal.pone.0180292
[26] National Institute for Health and Care Excellence. Self-harm: assessment, management and preventing recurrence: [G] Evidence review for risk assessment and formulation [NG225]. 2022. https://www.nice.org.uk/guidance/ng225/evidence (accessed May 21, 2025). p.12-13
[27] Care Quality Commission. *The state of health care and adult social care in England 2023/24*. October 2024. https://www.cqc.org.uk/sites/default/files/2024-10/20241025_stateofcare2324_print_0.pdf (accessed Sept 1, 2025).
[28] Endale T, Qureshi O, Ryan GK, et al. Barriers and drivers to capacity-building in global mental health projects. *Int J Ment Health Syst*. 2020;14:89. https://doi.org/10.1186/s13033-020-00420-4
[29] Hawton K, Lascelles K, Pitman A, Gilbert S, Silverman M. Assessment of suicide risk in mental health practice: shifting from prediction to therapeutic assessment, formulation, and risk management. *Lancet Psychiatry*. 2022;9(11):922–928. doi:10.1016/S2215-0366(22)00232-2. PMID: 35952701.
[30] Bailey J, Schartau P, Fazel S, Nazareth I, Petersen I. Incidence of suicide within two years of a first diagnosis of depression, anxiety, or mixed anxiety and depression: an exploratory cohort study in primary care using the Clinical Practice Research Datalink. *EClinicalMedicine*. 2025;87:103441. doi:10.1016/j.eclinm.2025.103441.
[31] GBD 2021 Suicide Collaborators. Global, regional, and national burden of suicide, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet Public Health*. 2025;10(3):e189-e202. doi:10.1016/S2468-2667(25)00006-4.
[32] Danechi S. Suicide statistics. Commons Library Research Briefing No 07749. 8 September 2025. UK Parliament House of Commons. https://researchbriefings.files.parliament.uk/documents/CBP-7749/CBP-7749.pdf#page=10.09 (accessed October 1st, 2025).
[33] Morgan C, Webb RT, Carr MJ, Kontopantelis E, Chew-Graham CA, Kapur N, Ashcroft DM. Self-harm in a primary care cohort of older people: incidence, clinical management, and risk of suicide and other causes of death. *Lancet Psychiatry*. 2018;5(11):905–912. doi:10.1016/S2215-0366(18)30348-1. PMID: 30337211; PMCID: PMC6203699.
[34] The Royal College of Psychiatrists. Suffering in silence: age inequality in older people's mental health care. 2018. https://www.rcpsych.ac.uk/improving-care/campaigning-for-better-mental-health-policy/college-reports/2018-college-reports/cr221 (accessed May 21, 2025).
[35] Gao, M., Wong, N.M.L., Lin, C. et al. Multimodal brain connectome-based prediction of suicide risk in people with late-life depression. *Nat. Mental Health* 1, 100–113 (2023). https://doi.org/10.1038/s44220-022-00007-7

[36]National Institute for Health and Care Excellence. Self-harm: assessment, management and preventing recurrence: Recommendations [NG225]. 2022. Recommendation 1.5.12. https://www.nice.org.uk/guidance/ng225/chapter/Recommendations (accessed May 21, 2025).
[37] *Equality Act 2010: Explanatory Notes*. Section 3.2.2.7 – Indirect discrimination. https://www.legislation.gov.uk/ukpga/2010/15/notes/division/3/2/2/7 (Accessed: 1 September 2025)
[38] National Institute for Health and Care Excellence. Overweight and obesity management [NG246]. 2025. Recommendation 1.9.11. https://www.nice.org.uk/guidance/ng246/chapter/Identifying-and-assessing-overweight-obesity-and-central-adiposity (accessed May 21, 2025).
[39] Horton R. Age barriers in healthcare. *Erasmus Law Rev*. 2020;13(1):6. https://ssrn.com/abstract=3681320 (accessed May 21, 2025).
[40] Geulayov G, Casey D, McDonald KC, et al. Suicide following presentation to hospital for non-fatal self-harm in the Multicentre Study of Self-harm: a long-term follow-up study. *Lancet Psychiatry*. 2019;6(12):1021–1030.
[41] Swift HJ, Abrams D, Lamont RA, Drury L. The Risks of Ageism Model: how ageism and negative attitudes toward age can be a barrier to active aging. *Soc Issues Policy Rev*. 2017;11(1):195–231. doi:10.1111/sipr.1203
[42] The Royal College of Psychiatrists. Suffering in silence: age inequality in older people's mental health care. 2018. https://www.rcpsych.ac.uk/improving-care/campaigning-for-better-mental-health-policy/college-reports/2018-college-reports/cr221 (accessed May 21, 2025).
[43] National Institute for Health and Care Excellence. Self-harm: assessment, management and preventing recurrence: Equality Impact Assessments [NG225]. 2022. https://www.nice.org.uk/guidance/ng225/history (accessed May 21, 2025).
[44] Daniels N, Sabin J. Setting Limits Fairly: Learning to Share Resources for Health. Oxford: Oxford University Press; 2008.
[45]Newdick C. Can judges ration with compassion? A priority-setting rights matrix. *Health Hum Rights*. 2018;20(1):107–120.
[46] Horton R. NICE vindicated in UK's High Court. *Lancet*. 2007;370(9587):547–548.
[47] Syrett K. NICE and judicial review: enforcing 'accountability for reasonableness' through the courts? *Med Law Rev*. 2008;16(1):127–140. doi:10.1093/medlaw/fwm025.
[48] Eisai Ltd v National Institute for Health and Clinical Excellence (NICE) [2007] EWHC 1941 (Admin)
[49] Daniels N, Sabin J. Setting Limits Fairly: Learning to Share Resources for Health. Oxford: Oxford University Press; 2008.
[50] Wasserman D, Wasserman C, Hadlackzy G. Cognitive biases in suicide risk assessment. In: Wasserman D, editor. *Oxford textbook of suicidology and suicide prevention*. 2nd ed. Oxford: Oxford University Press; 2021.
[51] Merino Y, Adams L, Hall WJ. Implicit bias and mental health professionals: priorities and directions for research. *Psychiatr Serv*. 2018;69(6):723–725. doi:10.1176/appi.ps.201700294.
[52] Wong ND. Cardiovascular risk assessment: the foundation of preventive cardiology. *Am J Prev Cardiol.* 2020;1:100008. doi:10.1016/j.ajpc.2020.100008
[53] Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, Musacchio KM, Jaroszewski AC, Chang BP, Nock MK. Risk factors for suicidal thoughts and behaviors: a meta-analysis of 50 years of research. *Psychol Bull.* 2017;143(2):187–232. doi:10.1037/bul0000084
[54] Quinlivan L, Cooper J, Meehan D, et al. Predictive accuracy of risk scales following self-harm: multicentre, prospective cohort study. *Br J Psychiatry*. 2017;210:429–36
[55] Chan MKY, Bhatti H, Meader N, Stockton S, Evans J, O'Connor RC, Kapur N, Kendall T. Predicting suicide following self-harm: systematic review of risk factors and risk scales. *Br J Psychiatry.* 2016;209(4):277–283. doi:10.1192/bjp.bp.115.170050
[56] Carter G, Milner A, McGill K, Pirkis J, Kapur N, Spittal MJ. Predicting suicidal behaviours using clinical instruments: systematic review and meta-analysis of positive predictive values for risk scales. *Br J Psychiatry.* 2017 Jun;210(6):387–395. doi:10.1192/bjp.bp.116.182717
[57] House A, Kapur N, Knipe D. Thinking about suicidal thinking. *Lancet Psychiatry*. 2020;7:997–1000.
[58] Grover C, Huber J, Brewer M, Basu A, Large M. Meta-analysis of clinical risk factors for suicide among people presenting to emergency departments and general hospitals with suicidal thoughts and behaviours. *Acta Psychiatr Scand.* 2023;148(6):491–524. doi:10.1111/acps.13620
[59] Large M, Kaneson M, Myles N, Myles H, Gunaratne P, Ryan C. Meta-Analysis of Longitudinal Cohort Studies of Suicide Risk Assessment among Psychiatric Patients: Heterogeneity in Results and Lack of Improvement over Time. *PLoS ONE.* 2016;11(6):e0156322. doi:10.1371/journal.pone.0156322
[60] Large M, Myles N, Myles H, Corderoy A, Weiser M, Davidson M, Ryan CJ. Suicide risk assessment among psychiatric inpatients: a systematic review and meta-analysis of high-risk categories. *Psychol Med.* 2017;48(7):1119–1127. doi:10.1017/S0033291717002537

[61] Runeson B, Odeberg J, Pettersson A, Edbom T, Jildevik Adamsson I, Waern M. Instruments for the assessment of suicide risk: a systematic review evaluating the certainty of the evidence. *PLoS ONE.* 2017;12(7):e0180292. doi:10.1371/journal.pone.0180292

[62] Tai A, Pincham H, Basu A, Large M. Meta-analysis of risk factors for suicide after psychiatric discharge and meta-regression of the duration of follow-up. *Aust N Z J Psychiatry.* 2025;59(8):679–691. doi:10.1177/00048674251348372

[63] Fazel S, Vazquez-Montes MDL, Molero Y, Runeson B, D'Onofrio BM, Larsson H, Lichtenstein P, Walker J, Sharpe M, Fanshawe TR. Risk of death by suicide following self-harm presentations to healthcare: development and validation of a multivariable clinical prediction rule (OxSATS). *BMJ Ment Health*. 2023;26:e300673.

[64] Corke M, Mullin K, Angel-Scott H, Xia S, Large M. Meta-analysis of the strength of exploratory suicide prediction models: from clinicians to computers. *BJPsych Open.* 2021;7(1):e26. doi:10.1192/bjo.2020.162

[65] Naifeh, J.A., Edwards, E.R., Bentley, K.H. *et al.* Predicting suicide attempts among US Army soldiers using information available at the time of periodic health assessments. *Nat. Mental Health* **3**, 242–252 (2025). https://doi.org/10.1038/s44220-024-00360-9

[66] Han R, Acosta JN, Shakeri Z, Ioannidis JPA, Topol EJ, Rajpurkar P. Randomised controlled trials evaluating artificial intelligence in clinical practice: a scoping review. *Lancet Digit Health.* 2024;6(5):e367–e373. doi:10.1016/S2589-7500(24)00047-5

[67] Woodford R, Spittal MJ, Milner A, McGill K, Kapur N, Pirkis J, Mitchell A, Carter G. Accuracy of clinician predictions of future self-harm: a systematic review and meta-analysis of predictive studies. *Suicide Life Threat Behav*. 2019;49(1):23–40. doi:10.1111/sltb.12395. PMID: 28972271.

[68] Fazel S, Vazquez-Montes MDL, Lagerberg T, Molero Y, Walker J, Sharpe M, Larsson H, Runeson B, Lichtenstein P, Fanshawe TR. Risk of repeat self-harm among individuals presenting to healthcare services: development and validation of a clinical risk assessment model (OxSET). *BMJ Ment Health*. 2024;27:e301180.

[69] National Institute for Health and Care Excellence. Self-harm: assessment, management and preventing recurrence: [G] Evidence review for risk assessment and formulation [NG225]. 2022. https://www.nice.org.uk/guidance/ng225/evidence (accessed May 21, 2025). p.12-13

[70] Horowitz LM, Bridge JA, Teach SJ, et al. Ask Suicide-Screening Questions (ASQ): A Brief Instrument for the Pediatric Emergency Department. *Arch Pediatr Adolesc Med.* 2012;166(12):1170–1176. doi:10.1001/archpediatrics.2012.1276

[71] U.S. Department of Health and Human Services: Substance Abuse and Mental Health Services Administration. SAFE-T: suicide assessment five-step evaluation and triage. HHS Publication No. (SMA) 09-4432. CMHS-NSP-0193. 2024. https://library.samhsa.gov/product/safe-t-suicide-assessment-five-step-evaluation-and-triage/pep24-01-036 (accessed February 24th, 2026).

[72] Clinical decision making: more than just an algorithm. *Lancet Oncol*. 2017;18(12):1553

[73] Usher-Smith JA, Silarova B, Schuit E, et al. Impact of provision of cardiovascular disease risk estimates to healthcare professionals and patients: a systematic review. *BMJ Open*. 2015;5:e008717. doi:10.1136/bmjopen-2015-008717.

[74] Runeson B, Odeberg J, Pettersson A, Edbom T, Jildevik Adamsson I, Waern M. Instruments for the assessment of suicide risk: a systematic review evaluating the certainty of the evidence. *PLoS ONE.* 2017;12(7):e0180292. doi:10.1371/journal.pone.0180292

[75] National Institute for Health and Care Excellence. Self-harm: assessment, management and preventing recurrence: [G] Evidence review for risk assessment and formulation [NG225]. 2022. https://www.nice.org.uk/guidance/ng225/evidence (accessed May 21, 2025). p.12-13

[76] Hawton K, Lascelles K, Pitman A, Gilbert S, Silverman M. Assessment of suicide risk in mental health practice: shifting from prediction to therapeutic assessment, formulation, and risk management. *Lancet Psychiatry*. 2022;9(11):922–928. doi:10.1016/S2215-0366(22)00232-2. PMID: 35952701.

[77] Damen JA, Pajouheshnia R, Heus P, et al. Performance of the Framingham risk models and pooled cohort equations for predicting 10-year risk of cardiovascular disease: a systematic review and meta-analysis. *BMC Med*. 2019;17:109. doi:10.1186/s12916-019-1340-7.

[78] Vistisen D, Andersen GS, Hansen CS, Hulman A, Henriksen JE, Bech-Nielsen H, et al. Prediction of first cardiovascular disease event in type 1 diabetes mellitus: the steno type 1 risk engine. *Circulation*. 2016;133:1058–1066.

[79] National Institute for Health and Care Excellence. Cardiovascular disease: risk assessment and reduction, including lipid modification [NG238]. 2023. https://www.nice.org.uk/guidance/ng238 (accessed May 21, 2025).

[80] National Institute for Health and Care Excellence. Early and locally advanced breast cancer: diagnosis and management [NG101]. 2024. https://www.nice.org.uk/guidance/ng101 (accessed May 21, 2025).

[81] National Institute for Health and Care Excellence. Ovarian cancer: identifying and managing familial and genetic risk [NG241]. 2024. https://www.nice.org.uk/guidance/ng241 (accessed May 21, 2025).
[82] Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ*. 2017;357:j2099.
[83] Ferrat LA, Vehik K, Sharp SA, et al. A combined risk score enhances prediction of type 1 diabetes among susceptible children. *Nat Med*. 2020;26:1247–1255. doi:10.1038/s41591-020-0930-4.
[84] Beaudry G, Canal-Rivero M, Ou J, Matharu J, Fazel S, Yu R. Evaluating the risk of suicide and violence in severe mental illness: a feasibility study of two risk assessment tools (OxMIS and OxMIV) in general psychiatric settings. *Front Psychiatry.* 2022;13:871213. doi:10.3389/fpsyt.2022.871213
[85] Senior M, Burghart M, Yu R, Kormilitzin A, Liu Q, Vaci N, Nevado-Holgado A, Pandit S, Zlodre J, Fazel S. Identifying predictors of suicide in severe mental illness: a feasibility study of a clinical prediction rule (Oxford Mental Illness and Suicide Tool or OxMIS). *Front Psychiatry.* 2020;11:268. doi:10.3389/fpsyt.2020.00268
[86] Shaw JL, Beans JA, Noonan C, Smith JJ, Mosley M, Lillie KM, Avey JP, Ziebell R, Simon G. Validating a predictive algorithm for suicide risk with Alaska Native populations. *Suicide Life-Threat Behav.* 2022;52:696–704. doi:10.1111/sltb.12853
[87] Beaudry G, Canal-Rivero M, Ou J, Matharu J, Fazel S, Yu R. Evaluating the risk of suicide and violence in severe mental illness: a feasibility study of two risk assessment tools (OxMIS and OxMIV) in general psychiatric settings. *Front Psychiatry.* 2022;13:871213. doi:10.3389/fpsyt.2022.871213
[88] Seyedsalehi A, Fazel S. Clinical guidelines on self-harm and suicide prevention: taking uncertainty into account in the evidence base. *BMJ Ment Health* 2026;29:1–3. doi:10.1136/bmjment-2025-302420