BMJ Health & Care Informatics

Enhancing trust in clinical decision support systems: a framework for developers

Caroline Jones [1], James Thornton [2], Jeremy C Wyatt [3]

To cite: Jones C, Thornton J, Wyatt JC. Enhancing trust in clinical decision support systems: a framework for developers. *BMJ Health Care Inform* 2021;**28**:e100247. doi:10.1136/bmjhci-2020-100247

Received 02 October 2020
Revised 23 December 2020
Accepted 15 January 2021

[1]Hillary Rodham Clinton School of Law, Swansea University, Swansea, Wales, UK
[2]Law School, Nottingham Trent University, Nottingham, Nottinghamshire, UK
[3]Wessex Institute, University of Southampton, Southampton, UK

Correspondence to
Dr Caroline Jones;
caroline.jones@swansea.ac.uk

INTRODUCTION

Systematic reviews show that clinical decision support systems (CDSSs) can improve the quality of clinical decisions and healthcare processes[1] and patient outcomes[2]; although caution has been expressed as to balancing the risks of using CDSSs (eg, alert fatigue) when only small or moderate improvements to patient care have been shown.[3] Yet, despite the potential benefits, studies indicate that uptake of these tools in clinical practice is generally low due to a range of factors.[4–7] The well-funded National Health Service (NHS) PRODIGY programme is an example of a carefully developed CDSS - commissioned by the Department of Health to support GPs - which failed to influence clinical practice or patient outcomes, with low uptake by clinicians in a large-scale trial.[8] A subsequent qualitative study revealed that, among other issues—such as the timing of the advice—trust was an issue: 'I don't trust … practising medicine like that … I do not want to find myself in front of a defence meeting, in front of a service tribunal, a court, defending myself on the basis of a trial of computer guidelines' [quote from GP].[9]

Another qualitative study exploring factors hindering CDSSs' uptake in hospital settings found that clinicians perceive that CDSSs 'may reduce their professional autonomy or may be used against them in the event of medical-legal controversies'.[10] Thus, CDSSs may be 'perceived as limiting, rather than supplementing, physicians' competencies, expertise and critical thinking', as opposed to a working tool to augment professional competence and encourage interdisciplinary working in healthcare settings.[10] Similarly, a recent survey carried out by the Royal College of Physicians revealed that senior physicians had serious concerns about using CDSSs in clinical practice, with trust and trustworthiness being key issues (see examples below).[11]

Trust is an important foundation for relationships between the developers of information systems and users, and is a contemporary concern for policymakers. It has, for example, been highlighted in the House of Lords Select Committee on Artificial Intelligence (AI) report[12]; the Topol Review[13]; a number of European Commission communications,[14–16] reports[17–20] and most recently a White Paper on AI[21]; and investigated in the context of knowledge systems, for example, for Wikipedia.[22] Although it is an important concept, it is not always defined; rather, its meaning may be inferred. For example, the House of Lords Select Committee used the phrase 'public trust' eight times,[12] but the core concern appeared to be about confidence over the use of patient data, rather than patient perceptions regarding the efficacy (or otherwise) of the AI in question. Such documents appear to take an implicit or one-directional approach to what is meant by 'trust'.

Notably, the Guidelines of the High-Level Expert Group on AI outline seven key requirements that might make AI systems more trustworthy[17]; whereas the White Paper focuses on fostering an 'ecosystem of trust' through the development of a clear European regulatory framework with a risk-based approach.[21] Therefore, in keeping with the drive for promoting clinical adoption of AI and CDSSs while minimising the potential risks,[13] here we apply Onora O'Neill's[23 24] multidirectional trust and trustworthiness framework[25] to explore key issues underlying clinician (doctor, nurse or therapist) trust in and the use (or non-use) of AI and CDSS tools for advising them about patient management, and the implications for CDSS developers. In doing so, we do not seek to examine particular existing CDSSs' merits and flaws in-depth, nor do we address the merits of the deployment process itself. Rather, we focus

on generic issues of trust that clinicians report having about CDSSs' properties, and on improving clinician trust in the use and outputs of CDSSs that have already been deployed.

Two points merit attention at this stage. First, O'Neill's[25] framework is favoured as—in the words of Karen Jones—O'Neill 'has done more than anyone to bring into theoretical focus the practical problem that would-be trusters face: how to align their trust with trustworthiness'.[26] Second, some nuance is required when determining who or what is being trusted. For example, Annette Baier makes clear that her own account of trust supposes:

> that the trusted is always something capable of good or ill will, and it is unclear that computers or their programs, as distinct from those who designed them, have any sort of will. But my account is easily extended to firms and professional bodies, whose human office-holders are capable of minimal goodwill, as well as of disregard and lack of concern for the human persons who trust them. It could also be extended to artificial minds, and to any human products, though there I would prefer to say that talk of trusting products like chairs is either metaphorical, or is shorthand for talk of trusting those who produced them.[27]

Similarly, Joshua James Hatherley 'reserve[es] the label of 'trust' for reciprocal relations between beings with agency'.[28] Accordingly, our focus is on the application of O'Neill's[25] framework to CDSS developers as 'trusted' agents, and measures they could adopt to become more trustworthy.

O'Neill's trust and trustworthiness framework: A summary

O'Neill notes that 'trust is valuable when placed in trustworthy agents and activities, but damaging or costly when (mis)placed in untrustworthy agents and activities'.[25] She usefully disaggregates trust into three core but related elements:

1. Trust in the truth claims made by others, such as claims about a CDSS's accuracy made by its developer. These claims are empirical, since their correctness can be tested by evaluating the CDSS.[29]
   Trust in others' commitments or reliability to do what they say they will, such as clinicians trusting a developer to maintain and update their CDSS products. This is normative: we use our understanding of the world and the actors in it to judge the plausibility of a specific commitment, such as our bank honouring its commitment to send us statements.
2. Trust in others' competence or practical expertise to meet those commitments. This is again normative: we use our knowledge of the agent in whom we place our trust and our past experience of their actions to judge their competence, such as trust in our dentist's ability to extract our tooth and the 'skill and good judgement she brings to the extraction'.[25]

This approach utilises two 'directions of fit': the empirical element (1) in one direction (does the claim 'fit' the world as it is?), and the two normative elements (2-3) in another (does the action 'fit' the claim?).[25] Relatedly, O'Neill has written on the concept of 'judgement'; drawing a distinction between judgement in terms of looking at the world and assessing how it measures up (or 'fits') against certain standards (normative), versus an initial factual judgement of what a situation is, which 'has to fit the world rather than to make the world fit or live up to' a principle (empirical).[30]

In deciding whether to trust and use a CDSS, a user is similarly also making judgements about it. O'Neill's threefold framework may therefore provide a helpful way to examine the issues in this context. In the following sections we discuss how CDSS developers can use each component of this framework to increase their trustworthiness, and conclude with suggestions on how informaticians might fruitfully apply this framework more widely to understand and improve user–developer relationships. Inevitably, this theoretical approach cannot address every potential issue, but it is used here as a means of organising diverse concerns around trust issues into a coherent framework.

TRUSTING THE TRUTH CLAIMS MADE BY DEVELOPERS

CDSS developers might assume that their users are interested in the innovative machine learning or knowledge representation method used, or how many lines of code the CDSS incorporates. However, Petkus *et al*'s[11] recent survey of the views and experience of 19 senior UK physicians representing the views of a variety of specialties provides some evidence of what a body of senior clinicians expect from CDSSs, that developers can use to shape their truth claims and build clinical trust. While this is not generalisable/representative of all clinicians it does provide a useful illustration of clinical concerns, and our intent is to demonstrate how applying O'Neill's trust/trustworthiness framework might help our understanding of how to mitigate these issues. Table 1 shows the six clinical concerns about CDSSs which scored highest in the analysis. The score combines both the participant-rated

Table 1 Concerns about CDSS quality in Petkus *et al* survey

Concerns about CDSS quality	Score
The accuracy of advice may be insufficient for clinical benefit	15.5
How extensively was clinical effectiveness of CDSS tested	15
Whether CDSSs are based on the latest evidence	14.5
CDSSs can interrupt clinical workflow or disrupt consultations	14.5
CDSSs can ignore patient preferences	12.5
Whether the CDSS output is worded clearly	11.5

CDSS, clinical decision support systems.

severity of the concern and its frequency in the responses; the maximum score on this scale was 19:[11]

The greatest concerns here relate to O'Neill's concept (or direction of fit) of empirical trust. Whether the advice provided by a CDSS is correct, has strong evidence for its clinical effectiveness from testing etc. ultimately concerns whether its advice 'fits' or matches (eg, in diagnosis) the patient's actual condition. Can it (and/or the people that designed/made it) be trusted in an empirical sense of being factually correct?

What kind of truth claims may appeal to clinicians?

Drawing on the evidence in table 1, developers should report to clinicians: the accuracy of the advice or risk estimates; CDSS effectiveness (impact on patients, decisions and the NHS); whether the CDSS content matches current best evidence (see 'Guidelines: codes and standards frameworks' below); its usability and ease of use in clinical settings; whether its output is worded clearly, and if takes account of patient preferences. These claims should be phrased in professional language, avoiding the extravagant claims about AI often seen in the press.[31 32] Instead of different developers adopting a range of metrics for reporting study results there is a need for a standard CDSS performance reporting 'label' for these assessments, to help clinicians identify, compare and judge the empirical claims being made about competing CDSSs. This is by analogy with European Union (EU) consumer regulations dictating how, for example, tyre manufacturers report on road noise, braking performance and fuel economy for their tyres (figure 1),[33] and EU plans for a health app label.

Ensuring that the truth claims can be verified

First, CDSS developers should be aware of the 'evidence-based medicine' culture,[34] reflected in the top three concerns in table 1. This means that, before clinicians make decisions such as how to treat a patient or which CDSS to use, they look for well designed, carefully conducted empirical studies in typical clinical settings using widely accepted outcomes that answer well-structured questions. This entails a 'critical appraisal' process to identify and reject studies that are badly designed or conducted, or from settings or with patients that do not resemble those where the CDSS will be used.[34] So, it has long been established that empirical evaluation and the evidence it generates are crucial to generating trust.[29] However, a systematic review of empirical research has shown that, when CDSS developers themselves carried out the study, they were three times as likely to generate positive results as when an independent evaluator did so.[35] Therefore, studies that establish these truth claims should be carried out by independent persons or bodies. To counter suspicions of bias or selective reporting, the full study protocol and results should be made openly available, for example, by publication.[36 37] Again, there is

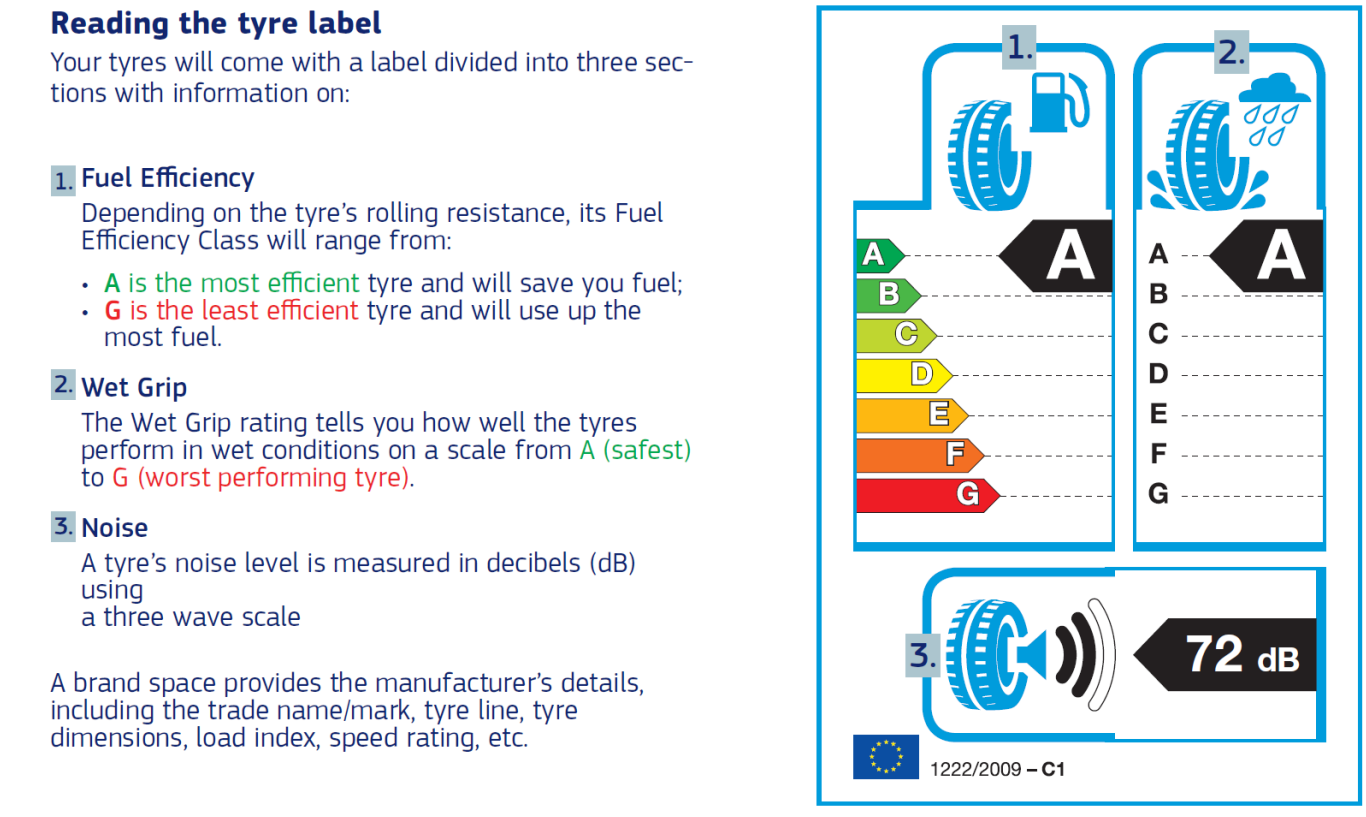

Figure 1 Example of an EU tyre label and how to interpret it.[38]

Table 2 Concerns about professional practice, ethics and liability in Petkus *et al* survey

Concerns about professional practice, ethics and liability	Score
The legal liability of doctors who rely on CDSS advice is unclear	17.5
Some CDSSs act like a 'black box', with no insight possible for the user about how they arrived at their advice or conclusions	15
Doctors may follow incorrect CDSS advice, even if they would make correct decisions without it	13.5
CDSSs can embed unconscious bias, with some patient groups receiving unfair care as a result	12.5

CDSS, clinical decision support systems.

an opportunity to establish standard methods for carrying out performance or impact studies, so that clinicians can trust and compare study results on different CDSSs from different suppliers—as exemplified by EU tyre performance testing standards.[33 38]

Concerns that software developers raise about evaluating CDSS are that these studies are expensive and can take a lot of time,[37] so yield results that can be obsolete by the time they are available. However, choosing the right designs such as MOST (multiphase optimization strategy), SMART (sequential multiple assignment randomized trial), or A/B testing (randomized control experiment to compare two versions, A and B)[39] means that studies can be carried out rapidly and at low cost. Further, if the study not only meets the requirements of the EU Regulation on Medical Devices (see 'UK and EU Regulation on Medical Devices' below), but also provides strong foundations for clinical trust in the CDSS developers, then commissioned independent studies can show a very positive return on investment and could be justified as part of a CDSS's product marketing strategy.

TRUSTING OTHERS' COMMITMENTS

O'Neill asks whether we can trust what others say they will do.[23] Petkus *et al*'s[11] survey also asked clinicians about professional practice, ethics and liability matters, as table 2 shows:

The last two items in table 2 relate to issues of empirical trust (is the advice factually correct?), which can be addressed by following the suggestions in the section on 'Trusting the truth claims made by developers' above. However, the first two concerns (and those found by Liberati *et al*[10]) address not only whether the CDSS provides correct advice, but also whether it does what it claims to do. Clinicians are unable to evaluate concerns about a 'black-box' CDSS because they will likely have no idea about how answers have been arrived at: it demands faith from clinicians that the trust commitments will be met. Rather than a useful support to their practice, such a CDSS may be considered a hindrance to the exercise of clinicians' judgement and critical thinking—as in the trial of PRODIGY (a clinical decision support tool commissioned by the Department of Health to help GPs).[8] There are related concerns about legal liability. What if the clinician relies on the CDSS and this causes harm to a patient? The clinician must trust that a 'black-box' CDSS will do what it is supposed to, and not cause harm for which they may be held legally responsible. Harm could obviously be caused by the CDSS if it is not working as the developers intended (eg, due to software issues). However, even without such issues, if the CDSS utilises a deep learning method such as neural networks, the clinician still has to trust that the mechanism through which conclusions have been derived is sensible, and has only taken into account clinically relevant details, ignoring spurious information such as the patient's name or the presence of a ruler in images of a suspicious skin lesion.[40]

In terms of potential legal liability, the situation does indeed appear to be unclear. Searches we carried out in legal databases (Lexis Library, Westlaw, BAILII), and PubMed, for terms around CDSS (adviser, expert system, risk score, algorithm, flowchart, automated tool, etc) turned up blank; nor have other researchers been able to locate published decisions in the UK, Europe or USA.[41] However, it is well established that clinicians are legally responsible for the medical advice and treatment given to their patients, irrespective of the use of a CDSS.[42] They must still reach the standard of the reasonable clinician in the circumstances. This makes it all the more important, if clinical uptake is to be improved, that clinicians have reasons to trust the CDSS developers and in turn their products/services.[43]

How can CDSS developers facilitate this trust?

While developers cannot fix an uncertain legal framework, there are several steps they can take to nurture trust in this area. Most obviously, to ensure that the way the CDSS works and comes to its conclusions are made as clear as possible to users. It may not be realistic to do so completely, particularly as CDSS software becomes more complex via machine learning.[44] However, giving—where possible—some account of the mechanism for how decisions are arrived at; the quality, size and source of any data-sets relied on; and assurance that standard guidelines for training the algorithm were followed (as well as monitoring appropriate learning diagnostics) will probably assuage some clinicians' concerns.[44]

In addition, even if some 'black-box' elements are unavoidable, clinicians' anxieties regarding the dependability or commitment aspects of O'Neill's[23] trust framework may be alleviated by ensuring that frequent updates, fixes and support are all available. This should help clinicians feel more confident that the CDSS is likely to be reliable, and gives them something concrete to point to later to evidence their diligence and reasonableness, for example if they appear in court or at a professional conduct hearing.[9–11]

Box 1 Developer actions that suggest competence and commitment to producing high quality clinical decision support systems (CDSSs)

- Recruit and retain a good development team with the right skills.[58]
- Use the right set of programming tools and safety-critical software engineering processes and methods, for example, HAZOP (Hazard and Operability Analysis) to understand and limit the risks of CDSSs.[17 60]
- Carry out detailed user research for example, user-centred design workshops; establish an online user community and monitor it for useful insights; or form a multidisciplinary steering group of key stakeholders.[13 61]
- Obtain the best quality, unbiased data to train the algorithm; use the right training method and diagnostics to monitor the learning process.[46]
- Implement relevant technical standards, obtain a CE mark (Conformitè Europëenne: the EU's mandatory conformity mark by which manufacturers declare that their products comply with the legal requirements regulating goods sold in the European Economic Area) for their CDSS as a medical device.[45]
- Publish an open interface to their software; carry out interoperability testing.[62 63]
- Build on a prior track record of similar products that appeared safe.[58]
- Follow relevant codes of practice for artificial intelligence and data-based technologies.[46 47]
- Implement continuing quality improvement methods, for example, log and respond to user comments and concerns[60]; deliver updates to the CDSS regularly[61]; seek to become certified as ISO 9000 compliant.

TRUSTING OTHERS' COMPETENCE

O'Neill[23] suggests that we ask whether others' actions meet, or will meet, the relevant standards or norms of competence. Factors that may impact positively on improving clinician trust include, but are not limited to, those listed in box 1.

In this section, we focus on the technical standards,[45] and current codes of practice and development standards frameworks potentially applicable to CDSSs.[46 47] Much more could be said about these approval processes than space permits here. However, the point is not to analyse the merits of the approval processes, but to illustrate how O'Neill's framework helps to highlight their additional importance (beyond being strictly required) as a way to enhance (normative) trust.

UK and EU Regulation on Medical Devices

The initial question is whether CDSSs are medical devices? Classification as a medical device means that a CDSS will be subject to the EU Regulation on Medical Devices.[45] The European Medicines Agency (the agency responsible for the evaluation and safety monitoring of medicines in the EU) states that 'medical devices are products or equipment intended generally for a medical use'.[48] Article 1 stipulates that 'medical devices', manufactured for use in human beings for the purpose of, inter alia, diagnosis, prevention, monitoring, treatment or alleviation of disease, means: 'any instrument, apparatus, appliance, software, material or other article, whether used alone or in combination, including the software intended by its manufacturer to be used specifically for diagnostic and/or therapeutic purposes and necessary for its proper application'.[45]

In the UK, the Medicines and Healthcare Products Regulatory Agency (MHRA) has indicated that a CDSS is 'usually considered a medical device when it applies automated reasoning such as a simple calculation, an algorithm or a more complex series of calculations. For example, dose calculations, symptom tracking, clinicians (sic) guides to help when making decisions in healthcare'.[49] Hence, although some CDSSs may fall outside this definition (eg, by providing information only), our analysis is directed at those that do fall within the meaning of medical devices.

Accordingly, developers must adhere to the requirements of the EU Medical Devices Regulation[45] and post-Brexit, under domestic legislation, namely the Medicines and Medical Devices Act 2021.[50] These requirements include passing a conformity assessment carried out by an EU-recognised notified body (for medical devices for sale in both Northern Ireland and the EU), or a UK approved body (for products sold in England, Wales and Scotland)[51] to confirm that the CDSS meets the essential requirements (the precise assessment route depends on the classification of the device).[52] The focus of this testing is safety. Following confirmation that the device meets the essential requirements, a declaration of conformity must be made and a CE mark must be visibly applied to the device prior to it being placed on the market[53] (from 1 January 2021 the UKCA (UK Conformity Assessed) mark has been available for use in England, Wales and Scotland,[54] and the UKNI (UK Northern Ireland) mark for use in Northern Ireland).[55] The general obligations of manufacturers are provided in Article 10 of the EU Medical Devices Regulation, including risk management, clinical evaluation, postmarket surveillance and processes for reporting and addressing serious incidents[45]; see also the 'yellow card' scheme operated by the MHRA which allows clinicians or members of the public to report issues with medical devices.[56] Clinical users will rightly mistrust any CDSS developer who is unaware of these regulations, or fails to follow them carefully.

Nevertheless, NHSX (the organisation tasked with setting the overall strategy for digital transformation in the NHS) is seeking to 'streamline' the assurance process of digital health technologies.[57] Similarly, in the USA, the Food and Drug Administration (FDA) is piloting an approach where developers demonstrating 'a culture of quality and organisational excellence based on objective criteria' could be precertified.[58] Such 'trusted' developers could then benefit from less onerous FDA approval processes for their future products due to their demonstrable competence.[25]

Guidelines: codes and standards frameworks

In addition to the generic Technology Code of Practice[59] which should inform developers' practices, there are two sets of guidance specifically focused on the development and use of digital health tools, including data-derived AI tools for patient management – one issued by the Department of Health and Social Care (DHSC) and NHS

England,[46] and the second by the National Institute for Health and Care Excellence (NICE).[47]

The DHSC and NHS England code of conduct aims to complement existing frameworks, including the EU Regulation and CE mark process, to 'help to create a trusted environment',[46] supporting innovation that is safe, evidence based, ethical, legal, transparent and accountable. It refers to the 'Evidence standards framework for digital health technologies' developed by NICE in conjunction with NHS England, NHS Digital, Public Health England, MedCity and others.[47] The aim of this standards framework is to facilitate better understanding by developers (and others) as to what 'good levels of evidence for digital healthcare technologies look like',[47] and is applicable to technologies using AI with fixed algorithms; whereas those using adaptive algorithms are instead governed by the DHSC code (see Principle 7).[46]

Visible and/or certified compliance with these codes and standards would provide developers with normative objective standards to meet, and point clinical users of CDSSs to evidence of their competence.[25] Having confidence in the professionalism of the developers should go some way towards reassuring clinicians as to the safety, accuracy and efficacy of CDSSs, thus potentially fostering greater uptake in practice.

CONCLUSION

O'Neill's[25] approach to trust and trustworthiness, focusing on empirical trust in developers' truth claims and normative trust in their commitment and competence to meet those claims, has proved a useful framework to analyse and identify ways that developers can improve user trust in them, and in turn—it is suggested—the CDSSs they produce. That is, of course, not to suggest that developers are necessarily at fault in any way. It may be that they are unfairly distrusted by (potential) users. We suggest the application of O'Neill's framework has helped to identify ways to facilitate and enhance trust in developers, and by extension, their CDSSs.

In summary, developers should:

- Make relevant claims about system content, performance and impact framed in professional language, preferably structured to a standard that allows clinicians to compare claims about competing CDSSs. These claims need to be supported by well-designed empirical studies, conducted by independent evaluators.
- Minimise 'black box' elements, ensure that internal mechanisms are—so far as possible—explained to users, and that CDSS software comes with a comprehensive update and support package. This could help clinicians gain a sense of control over the CDSS, and thus perceive the technology as a valuable working tool that complements their own skills and expertise.
- Comply with all relevant legal and regulatory (codes and standards) frameworks. Having confidence in the professionalism and competence of the developers should go some way towards reassuring clinicians as to the safety, accuracy and efficacy of CDSSs, thus potentially fosteing greater uptake in their use.

The benefit of applying O'Neill's[23] framework is that it requires us to consider issues associated with different facets of both trust and trustworthiness, maximising the possibilities for enhancing trust and trustworthiness once such concerns or objections are overcome. An implicit or one-directional understanding of trust might result in a narrower conclusion, focused on just one element of O'Neill's framework.[25] For example, an understanding solely based on normative competence might focus on the importance of complying with the regulations (not only to avoid sanctions, but to enhance trust); this is important, but O'Neill's framework demands consideration of different, equally useful, elements of trustworthiness.

This analysis is focused on clinician use of decision support tools, but we believe that a similar analysis would generate useful insights had we looked at other users and information systems, such as the public use of risk assessment apps, or professional use of electronic referral or order communication system advisory tools. The principles of examining the empirical truth claims of the software and the evidence on which they are based, then the competence of the supplier to match these claims and their commitment to do so, seems to generate useful insights no matter who the users are or what digital service is being trusted. Thus, we suggest that O'Neill's[25] framework is considered by health and care informaticians—both those developing and evaluating digital services—as a useful tool to help them explore and expand user trust in these products and services.

Contributors All authors designed and cowrote the draft of the paper. CJ took primary responsibility for wider materials on trust and the section on 'Trusting others' competence'. JT carried out the literature review and led on the section on 'Trusting others' commitments'. JCW contributed points about the evidence base for CDSSs, as well as designing the survey which partly stimulated this work, and led on the section on 'Trusting the truth claims made by developers'. All authors critically reviewed and edited the final draft.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests JW is receiving consultancy payments from NHSX AI Lab for advising on the validation and evaluation of AI systems for the NHS, but the views shared in this article are his own. The authors declare no other competing interest that might be relevant to the views expressed in this article.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; internally peer reviewed.

ORCID iDs
Caroline Jones http://orcid.org/0000-0001-7632-9468
James Thornton http://orcid.org/0000-0001-7847-5696
Jeremy C Wyatt http://orcid.org/0000-0001-7008-1473

REFERENCES

1 Roshanov PS, Fernandes N, Wilczynski JM, *et al*. Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials. *BMJ* 2013;346:f657.
2 Varghese J, Kleine M, Gessner SI, *et al*. Effects of computerized decision support system implementations on patient outcomes in inpatient care: a systematic review. *J Am Med Inform Assoc* 2018;25:593–602.
3 Kwan JL, Lo L, Ferguson J, *et al*. Computerised clinical decision support systems and absolute improvements in care: meta-analysis of controlled clinical trials. *BMJ* 2020;370:m3216.
4 Moxey A, Robertson J, Newby D, *et al*. Computerized clinical decision support for prescribing: provision does not guarantee uptake. *J Am Med Inform Assoc* 2010;17:25–33.
5 Kortteisto T, Komulainen J, Mäkelä M, *et al*. Clinical decision support must be useful, functional is not enough: a qualitative study of computer-based clinical decision support in primary care. *BMC Health Serv Res* 2012;12:349.
6 Patterson ES, Doebbeling BN, Fung CH, *et al*. Identifying barriers to the effective use of clinical reminders: bootstrapping multiple methods. *J Biomed Inform* 2005;38:189–99.
7 Pope C, Halford S, Turnbull J, *et al*. Using computer decision support systems in NHS emergency and urgent care: ethnographic study using normalisation process theory. *BMC Health Serv Res* 2013;13:111.
8 Eccles M, McColl E, Steen N, *et al*. Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: cluster randomised controlled trial. *BMJ* 2002;325:941.
9 Rousseau N, McColl E, Newton J, *et al*. Practice based, longitudinal, qualitative interview study of computerised evidence based guidelines in primary care. *BMJ* 2003;326:314.
10 Liberati EG, Ruggiero F, Galuppo L, *et al*. What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. *Implement Sci* 2017;12:113.
11 Petkus H, Hoogewerf J, Wyatt JC. AI in the NHS– are physicians ready? A survey of the use of AI & decision support by specialist societies, and their concerns. *Clinical Medicine* 2020;20:324–8.
12 House of Lords Select Committee on AI. *AI in the UK: ready, willing and able?* London: UK Parliament, 2018.
13 The Topol Review: preparing the healthcare workforce to deliver the digital future. NHS 2019.
14 Commission. Artificial Intelligence for Europe. COM (2018) 237 final.
15 Commission. Liability for emerging digital technologies. SWD(2018) 137 final.
16 Commission. Building Trust in Human-Centric Artificial Intelligence. COM (2019) 168 final.
17 Independent High Level Expert Group on AI. Ethics Guidelines for Trustworthy AI. *European Commission* 2019.
18 Independent High Level Expert Group on AI. Policy and Investment Recommendations for Trustworthy AI. *European Commission* 2019.
19 Expert Group on Liability and New Technologies - New Technologies Formation, Liability for Artificial Intelligence and other emerging digital technologies. EU 2019.
20 Commission. 'Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics' COM (2020) 64 final.
21 European Commission. 'White Paper on AI: A European approach to excellence and trust' COM (2020) 65 final.
22 Adams CE, Montgomery AA, Aburrow T, *et al*. Adding evidence of the effects of treatments into relevant Wikipedia Pages: a randomised trial. *BMJ Open* 2020;10:e033655.
23 O'Neill O. *A question of trust*. Cambridge University Press, 2002.
24 O'Neill O. *Autonomy and trust in bioethics*. Cambridge University Press, 2002.
25 O'Neill O. Linking trust to Trustworthiness. *International Journal of Philosophical Studies* 2018;26:293–300.
26 Jones K. Chapter 11, at 186. In: Archard D, ed. *Distrusting the trustworthy. in reading Onora O'Neill*. Taylor & Francis Group, 2013.
27 Baier A. Chapter 10, at 178. In: Archard D, ed. *What is trust?. in reading Onora O'Neill*. Taylor & Francis Group, 2013.
28 Hatherley JJ. Limits of trust in medical AI. *J Med Ethics* 2020;46:478–81.
29 Wyatt J, Spiegelhalter D. Evaluating medical expert systems: what to test and how? *Med Inform* 1990;15:205–17.
30 O'Neill O. Experts, practitioners, and practical judgement. *J Moral Philos* 2007;4:154–66.
31 Sample I. "It's going to create a revolution": how AI is transforming the NHS. The Guardian, 2018. Available: https://www.theguardian.com/technology/2018/jul/04/its-going-create-revolution-how-ai-transforming-nhs
32 Copestake J. Babylon claims its chatbot beats GPs at medical exam. BBC, 2018. Available: https://www.bbc.co.uk/news/technology-44635134
33 Thimbleby H. *Fix IT: Stories from Healthcare IT*. Oxford: Oxford University Press, 2020.
34 Sackett DL, Rosenberg WM, Gray JA, *et al*. Evidence based medicine: what it is and what it isn't. *BMJ* 1996;312:71–2.
35 Garg AX, Adhikari NKJ, McDonald H, *et al*. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA* 2005;293:1223–38.
36 Liu X, Cruz Rivera S, Moher D, *et al*. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med* 2020;26:1364–74.
37 Liu JLY, Wyatt JC. The case for randomized controlled trials to assess the impact of clinical information systems. *J Am Med Inform Assoc* 2011;18:173–80.
38 . Available: https://ec.europa.eu/info/energy-climate-change-environment/standards-tools-and-labels/products-labelling-rules-and-requirements/energy-label-and-ecodesign/energy-efficient-products/tyres_en
39 Murray E, Hekler EB, Andersson G, *et al*. Evaluating digital health interventions: key questions and approaches. *Am J Prev Med* 2016;51:843–51.
40 Narla A, Kuprel B, Sarin K, *et al*. Automated classification of skin lesions: from Pixels to practice. *J Invest Dermatol* 2018;138:2108–10.
41 Fox J, Thomson R. Clinical decision support systems: a discussion of quality, safety and legal liability issues. *Proc AMIA Symp* 2002:1–7.
42 Brahams D, Wyatt J. Decision AIDS and the law. *Lancet* 1989;2:632–4.
43 Cohen IG, Cops GH. Docs, and code: a dialogue between big data in health care and predictive policing. *UC Davis Law Review* 2017;51:437–74.
44 Hart A, Wyatt J. Evaluating black-boxes as medical decision AIDS: issues arising from a study of neural networks. *Med Inform* 1990;15:229–36.
45 EU Regulation on Medical Devices 2017/745.
46 UK Government Department of Health and Social Care. Code of conduct of AI and other data driven technologies. London, 2019. Available: https://www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology
47 National Institute for Health and Care Excellence. Evidence standards framework for digital health technologies, 2019. Available: https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies
48 European Medicines Agency. Medical devices, 2019. Available: https://www.ema.europa.eu/en/human-regulatory/overview/medical-devices
49 Medicines and Healthcare Products Regulatory Agency. Guidance: medical device stand-alone software including apps (including IVDMDs), 2018. Available: https://www.gov.uk/government/publications/medical-devices-software-applications-apps
50 . Available: https://www.legislation.gov.uk/ukpga/2021/3/contents/enacted/data.htm
51 Medicines and Healthcare Products Regulatory Agency. Medical devices UK Approved bodies, 2021. Available: https://www.gov.uk/government/publications/medical-devices-uk-approved-bodies/
52 European Commission. Guidance document - Classification of Medical Devices - MEDDEV 2.4/1 rev.9, 2015. Available: http://ec.europa.eu/DocsRoom/documents/10337/attachments/1/translations
53 UK Government Department for Business, Energy & Industrial Strategy. Guidance: CE marking, 2012. Available:https://www.gov.uk/guidance/ce-marking
54 MHRA. Medical devices: conformity assessment and the UKCA mark. Available: https://www.gov.uk/guidance/medical-devices-conformity-assessment-and-the-ukca-mark
55 Guidance using the UKNI marking, 2021Department for Business, Energy and Industrial Strategy. Available: https://www.gov.uk/guidance/using-the-ukni-marking
56 Yellow Card. Medicines and Healthcare Products Regulatory Agency, 2020. Available: https://yellowcard.mhra.gov.uk/
57 Joshi I, Joyce R. NHSX is streamlining the assurance of digital health technologies, 2019. Available: https://healthtech.blog.gov.uk/2019/11/01/nhsx-is-streamlining-the-assurance-of-digital-health-technologies/
58 FDA. Digital health software Precertification (Pre-Cert) program. from, 2020. Available: https://www.fda.gov/medical-devices/digital-health-

center-excellence/digital-health-software-precertification-pre-cert-program
59 UK Government Digital Service. Technology code of practice, 2019. Available: https://www.gov.uk/government/publications/technology-code-of-practice
60 NHS Digital Clinical Safety team. DCB0129: clinical risk management: its application in the manufacture of health it systems and DCB0160: clinical risk management: its application in the deployment and use of health it systems. from, 2018. Available: https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/dcb0160-clinical-risk-management-its-application-in-the-deployment-and-use-of-health-it-systems
61 Mahadevaiah G, RV P, Bermejo I, *et al*. Artificial intelligence-based clinical decision support in modern medical physics: selection, acceptance, commissioning, and quality assurance. *Med Phys* 2020;47:8.
62 NHS Digital. Interoperability toolkit. Available: https://digital.nhs.uk/services/interoperability-toolkit
63 Walsh K, Wroe C. Mobilising computable biomedical knowledge: challenges for clinical decision support from a medical knowledge provider. *BMJ Health Care Inform* 2020;27:e100121.