BMJ Health & Care Informatics

Validation framework for the use of AI in healthcare: overview of the new British standard BS30440

Mark Sujan [ID] ,[1,2] Cassius Smith-Frazer,[3] Christina Malamateniou,[4] Joseph Connor,[5] Allison Gardner,[6] Harriet Unsworth,[7] Haider Husain[3]

[1]Human Factors Everywhere, Woking, UK
[2]Education, Healthcare Safety Investigation Branch, Reading, UK
[3]Healthinnova Ltd, Bath, UK
[4]School of Health and Psychological Sciences, City University London, London, UK
[5]CarefulAI, Cwmbran, UK
[6]National Institute for Health and Care Excellence, London, UK
[7]Wellcome Trust, London, UK

Correspondence to
Dr Mark Sujan;
mark.sujan@gmail.com

INTRODUCTION

The British standard 'BS30440: Validation Framework for the Use of AI in Healthcare' will be published in the second quarter of 2023.[1] It details the evidence required by technology developers to assess and validate products using artificial intelligence (AI) in healthcare settings. Healthcare providers can mandate that their suppliers' products be certified against BS30440 to assure themselves and their service users that the AI product is effective, fair and safe.

For a decade now, there has been growing interest in healthcare AI, especially applications using machine learning approaches, such as deep neural networks.[2] This interest has grown exponentially over the past 5 years, with government bodies and regulatory authorities, non-governmental think tanks, professional associations and academic institutions developing a multitude of relevant guidance to address their local contexts.[3] In the United Kingdom (UK) this includes, for example, the National Institute for Health and Care Excellence Evidence Standards Framework for digital health technologies, NHSX guidance on 'Artificial Intelligence: how to get it right', and guidance on algorithmic impact assessment published by the Ada Lovelace Institute. In addition, there are several international reporting guidelines, including SPIRIT-AI[4] (The Standard Protocol Items: Recommendations for Interventional Trials - Artificial Intelligence) and CONSORT-AI[5] (Consolidated Standards of Reporting Trials - Artificial Intelligence) for clinical trials of healthcare AI technologies.

As a result, the landscape of guidance on how to develop safe and effective AI systems for healthcare is fragmented across hundreds of documents, largely with a focus on products that would be regulated as medical devices. This has led to a lack of formalised guidance for healthcare AI technologies that are out of remit of medical device regulations, such as those with a focus on healthcare resource planning, logistics or general health and well-being support. While regional regulations for AI (such as the European Union AI act) are in development, and national regulators (eg, the UK Medicines and Healthcare Products Regulatory Agency) develop their own regulatory strategies, there is a clear space for well designed and auditable standards to ensure safety, effectiveness and equity. Such standards do not replace legislation but can form the basis for novel regulatory approaches.

Against this backdrop of a multitude of guidance and frameworks, BS30440 is unique in two ways. First, BS30440 has been developed from an extensive review, which synthesises the fragmented healthcare AI landscape into a single, comprehensive framework. It has received additional input from a multidisciplinary panel of experts, two rounds of public consultations, as well as a community and patient engagement panel.

Second, BS30440 represents a fully auditable standard for the assessment of healthcare AI products. Auditing is critical to ensure that healthcare AI products offer demonstrable clinical benefits, that they reach sufficient levels of performance, that they successfully and safely integrate into the health and care environment, and that they deliver inclusive outcomes for all patients, service users and practitioners. Any healthcare AI product that is successfully certified against BS30440, has passed a broad and substantial evaluation across these properties.

This thorough process of synthesis and stakeholder consultation, coupled with the introduction of clear assessment criteria for

auditing, offers significant potential to suppliers who wish to navigate the complex AI guidance landscape by complying with a single framework.

STRUCTURE AND ASSESSMENT CRITERIA

BS30440 is structured around a product life-cycle for healthcare AI, described in five phases: inception, development, validation, deployment and monitoring. For each phase of the product life-cycle, a set of assessment criteria has been defined. The life-cycle within the framework is not intended to be prescriptive or to be thought of in a necessarily linear fashion. However, all the assessment criteria should be addressed during the product life-cycle.

The assessment criteria were developed through literature reviews and in consultation with a committee of subject matter experts from academia, governmental bodies, healthcare institutions and standards organisations. Patient and public representatives were involved to inform the development and review of the assessment criteria through written contribution and as part of a focus group to ensure diverse and inclusive input.

The standard includes carbon impact criteria because of the anticipated expansion of AI across the sector, which has the potential to result in significant environmental impact if not managed. Feedback from the public consultation on this topic was overwhelmingly positive. The importance of equity and fairness is highlighted as a core criterion for the development of ethical AI products,[6] both in ensuring engagement with the target audience, but also in terms of diversity and inclusiveness of decision-making and development.

Consideration is given to the inclusion of human factors and ergonomics, which runs across all life-cycle phases. The importance of human factors and ergonomics in the healthcare AI product life-cycle is increasingly being recognised,[7 8] and this is reflected in the standard.

In total, BS30440 includes 18 assessment criteria. Each assessment criterion is specified through auditable clauses against which an AI product can be assessed for conformity. An overview is provided in figure 1.

INTENDED AUDIENCE, AUDITING AND COMPLIANCE

The assessment criteria specified in BS30440 are intended to provide assurance of the safety, quality and performance of healthcare AI products. Patients and the public are the main beneficiaries of BS30440 as recipients of healthcare services, but they are not expected to engage directly with the standard.

BS30440 can support healthcare organisations in the procurement and assessment of AI products. Healthcare providers can adopt the standard as a requirement for their suppliers, similar to conformance to other standards such as ISO9001. For a given AI product to be certified against the standard, the product must have been developed and validated following a process aligned to the assessment criteria specified in BS30440. The developer must document evidence for the assessment criteria, which will be evaluated by a competent external auditor. This can provide reassurance to clinicians and staff working with AI products and to patients and their families. It also adheres to core principles of ethical AI in terms of transparency and accountability in providing clarity as to the chain of responsibility and evidence throughout the product life-cycle.

Developers are encouraged to begin with a self-assessment of their current development processes against each of the assessment criteria to establish their current level of conformity, to identify gaps in their development process and documentation, and to decide where they might need to improve their development processes. Developers should create an action plan for how to address any identified gaps to achieve certification and gather evidence as they design and develop their AI product. It is recommended that internal auditors or quality assurance managers work alongside the development team to collate and present the evidence in a systematic and standardised way and to minimise potential rework. Service
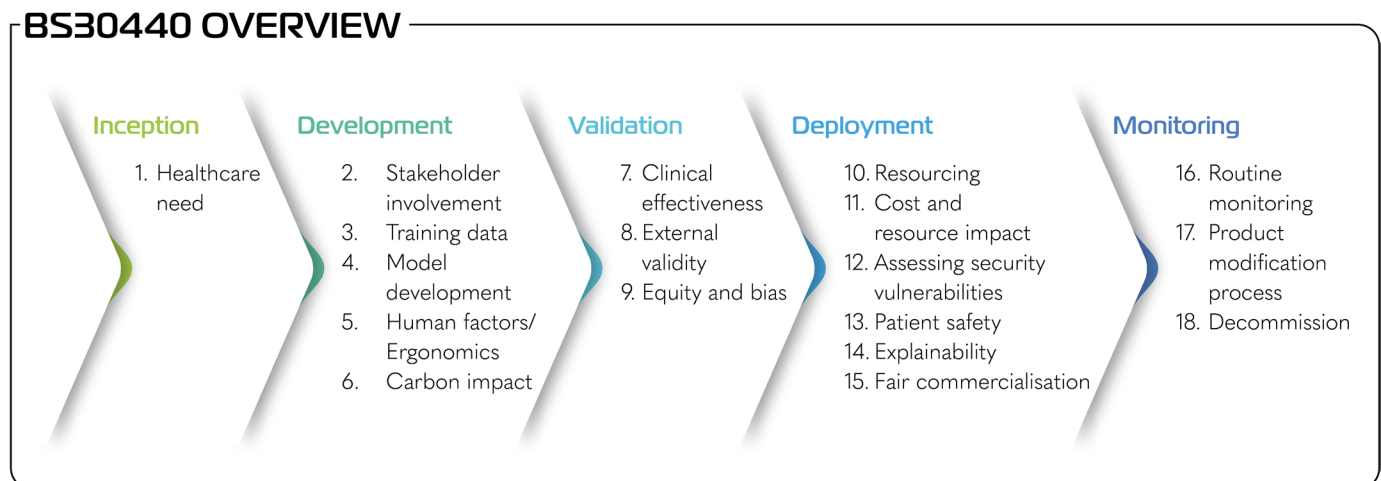

BS30440 OVERVIEW

Inception
1. Healthcare need

Development
2. Stakeholder involvement
3. Training data
4. Model development
5. Human factors/ Ergonomics
6. Carbon impact

Validation
7. Clinical effectiveness
8. External validity
9. Equity and bias

Deployment
10. Resourcing
11. Cost and resource impact
12. Assessing security vulnerabilities
13. Patient safety
14. Explainability
15. Fair commercialisation

Monitoring
16. Routine monitoring
17. Product modification process
18. Decommission

Figure 1 BS30440 structure and assessment criteria.

users and key stakeholders should be involved at all stages of the product life-cycle.

CONCLUSION

BS30440 provides within a single resource an actionable, comprehensive and auditable validation framework for healthcare AI. Conformity with the standard can provide assurance to developers, deploying healthcare organisations and patients.

There is a degree of overlap between BS30440 and other relevant forms of assessment and regulation of healthcare information technology, including medical device regulations and the NHS Digital clinical safety standards (DCB 0129 and DCB 0160). However, BS30440 covers specifically AI products, including those which are not included in current medical device regulations. In this way, such healthcare AI technologies can still be subjected to a process of assessment and certification to ensure a minimum standard across relevant assessment criteria. The standard has been shared with and received input from a wide range of organisations and stakeholders, and, as such, the evidence provided for the assessment criteria should facilitate any necessary regulatory approvals.

BS30440 applies across all development and use contexts of healthcare AI. However, the standard might be especially valuable in contexts where developers and deploying organisations have limited prior experience, knowledge, and resources about suitable healthcare AI development processes, including formal software engineering and assurance processes.

BS30440 assumes suppliers will have knowledge of all relevant design information either because they have developed the algorithm and models themselves or because they can access this information from the developers. This can be problematic in future scenarios, where potentially suppliers might make use of generic AI products, such as the increasingly popular large language models applications. In these situations, the supplier will not have designed the model and they might be unable to explain its origin. In that case, suppliers would not be compliant unless they are able to design an assurance wrapper around the generic model. This is not yet addressed in the standard.

BS30440 has been developed as a national initiative. While international committees including International Organization for Standardization (ISO) / International Electrotechnical Commission Subcommittee (IEC SC) 42 and European Committee for Standardization (CEN) / European Committee for Electrotechnical Standardization (CENELECT) Joint Technical Committee 21 (JTC21) have published standards and are in the process of developing their future work programmes, these initiatives are not specific to healthcare AI. The publication and use of BS30440 can serve as a first testbed to inform subsequent international standardisation activities for healthcare AI.

ORCID iD
Mark Sujan http://orcid.org/0000-0001-6895-946X

REFERENCES

1 British Standards Institution. *BS 30440 validation framework for the use of AI within healthcare*. BSI, 2023. Available: https://standardsdevelopment.bsigroup.com/projects/2021-00605#/section
2 Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019;25:44–56.
3 Malamateniou C, McFadden S, McQuinlan Y, *et al*. Artificial intelligence: guidance for clinical imaging and therapeutic radiography professionals, a summary by the society of Radiographers AI working group. *Radiography (Lond)* 2021;27:1192–202.
4 Liu X, Rivera SC, Moher D, *et al*. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *BMJ* 2020;370:m3164.
5 Liu X, Rivera SC, Moher D, *et al*. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *BMJ* 2020;370:m3164.
6 Wawira Gichoya J, McCoy LG, Celi LA, *et al*. Equity in essence: a call for operationalising fairness in machine learning for healthcare. *BMJ Health Care Inform* 2021;28:e100289.
7 Sujan M, Furniss D, Grundy K, *et al*. Human factors challenges for the safe use of artificial intelligence in patient care. *BMJ Health Care Inform* 2019;26:e100081.
8 Vasey B, Nagendran M, Campbell B, *et al*. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *BMJ* 2022;377:e070904.