NICE National Institute for Health and Care Excellence

EVALUATING ARTIFICIAL INTELLIGENCE IN HEALTHCARE: EXPANDING NICE'S EVIDENCE STANDARDS FRAMEWORK TO INCLUDE ARTIFICIAL INTELLIGENCE

Harriet Unsworth[1], Viknesh Sounderajah[2], Xiaoxuan Liu[3], Verena Wolfram[1,] Bernice Dillon[4]

[1] Office for Digital Health, National Institute for Health and Care Excellence, UK; [2]Department of Surgery & Cancer, Imperial College London, UK; [3]Institute of Inflammation and Ageing, College of Medical and Dental Sciences, University of Birmingham, UK; [4]Medical technologies evaluation programme, National institute for health and care excellence, UK

Background

The NICE evidence standards framework (ESF) for digital health and care technologies was originally published in 2019. It provides a pragmatic set of evidence standards for digital health technologies (DHTs) in the UK health and care system, including risk-stratified requirements for clinical evidence and graded levels of economic evaluation based on the scale of commissioning. The ESF provides a tool to allow local evaluators in the health and care system, to evaluate DHTs, to spot those technologies that are likely to be clinically and cost effective. In 2021, NHSx commissioned the NICE Office for Digital Health to update the ESF and expand it to include evidence standards for DHTs that include artificial intelligence algorithms (AI-DHTs) that may change over time (adaptive algorithms).

Objectives

The NICE ESF was originally created to aid the local evaluation of a wide range of DHTs. However, AI-DHTs with adaptive algorithms were excluded from the scope of the ESF. Since publication of the ESF in 2019, the use of AI in healthcare has grown significantly, and so it was felt that the addition of standards for AI-DHTs would be beneficial to the health and care system.

The objectives of the 2021/2 update to the NICE ESF were to:

1. Improve alignment of the functional classification to regulatory requirements
2. Improve the functional classification framework to make it easier to use
3. Revise evidence levels to reflect learning and stakeholder feedback
4. Seek further stakeholder and user feedback to improve content and usability

Objectives 1 to 4 are addressed in poster HTA13, poster session 1

5. Specify evidence requirements for AI and data-driven technologies with adaptive algorithms
6. Include AI and data-driven DHTs with adaptive algorithms in functional classification

For objectives 5 and 6, NICE commissioned an academic consortium comprising experts from The University of Birmingham, Imperial College London, and the Alan Turing Institute.

Methods

The evidence standards for AI-DHTs were created through a 5-stage process:

Stage 1: Scoping review (Academic consortium partners)
Summarise current classification systems for AI and understand how these relate to healthcare and HTA

Stage 2: Expert surveys and interviews (Academic consortium partners)
Characterise suitability of AI classification systems for use in the ESF, elicit expert opinion on what standards could be used to evaluate AI-DHTs

Stage 3: Modified Delphi process (Academic consortium partners)
To seek consensus on key factors for inclusion in the AI standards

Stage 4: Combining Delphi outputs into the NICE ESF (NICE)
Incorporate Delphi outcomes into the ESF, de-duplicate where needed and reword to improve accessibility

Stage 5: stakeholder workshops and public consultation (NICE)
Present the updated ESF to key stakeholders and use their feedback to improve the standards. Open public consultation to increase range of feedback

Results and conclusions

The updated ESF comprises a classification system for DHTs based on risk associated with intended use. DHTs are stratified into 3 tiers (A, B, C) of increasing risk. There are 24 evidence standards in total, 19 of the standards apply to tier A DHTs, 22 apply to tier B DHTs, and all 24 apply to tier C DHTs.

Care was taken to avoid AI exceptionalism, by presenting a single set of standards that can be applied to any DHTs, with 7 standards being especially relevant to data-driven DHTs (including those that use AI). The full NICE ESF can be viewed at:
https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies

The 7 standards that are most relevant to data-driven DHTs are summarised below (full text available through the link above)

#	Standard	Applies to tiers:		
3	**Consider health and care inequalities and bias mitigation**	A	B	C
4	**Embed good data practices in the design of the DHT**	A	B	C
5	**Define the level of professional oversight**	A	B	C
15	**Show real-world evidence that the claimed benefits can be realised in practice**	A	B	C
16	**The company and evaluator should agree a plan for measuring changes in the DHT's performance over time**	A	B	C
22	**Ensure transparency about requirements for deployment**	A	B	C
23	**Describe plans for communication, consent and training processes in place to allow the DHT to be understood by end users**	A	B	C

Harriet Unsworth
Technical Advisor
National Institute for Health and Care Excellence

Email:
harriet.unsworth@nice.org.uk

nice.org.uk