

Background

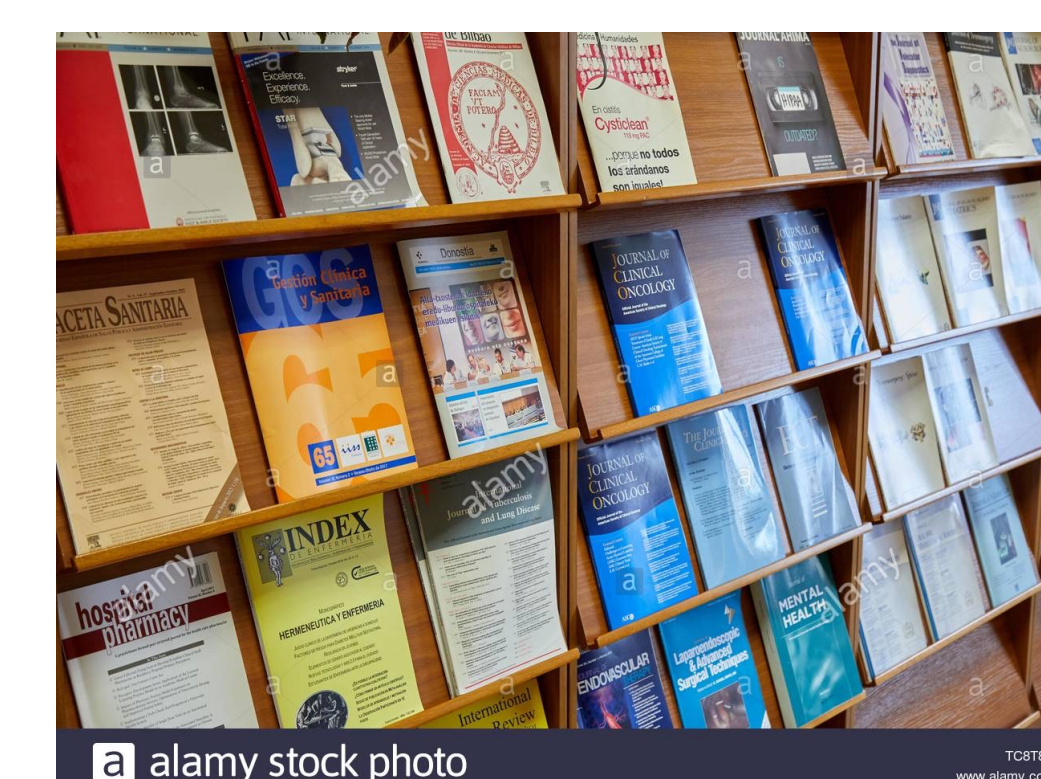
- The NHS is faced with a critical capacity-demand mismatch as the needs of an increasingly elderly population grow faster than its facilities, recruitment or funding (1).
- Technologies incorporating Artificial intelligence (AI) have been identified as key to addressing this unmet need by the Department of Health and Social Care in the NHS Long Term Plan (2).
- They are attractive due to the possibility of consistent performance, improved speed and round-the-clock access compared to conventional care.
- There is justified concern about 'AI Hype', and a need to ensure that technologies are safe, cost-effective and equitable before they are commissioned (3).
- NICE performs Health Technology Assessments (HTAs) on new and existing medicines and treatments, providing commissioners evidence-based guidance on their clinical benefit and cost-effectiveness.
- HTAs are performed using Evidence Standards Frameworks (ESF); an ESF for Digital Health Technologies was published in 2018 (4), providing information to developers, commissioners and other stakeholders on the levels of evidence needed for different innovations based on their intended purpose.
- This ESF however explicitly excluded many types of AI technologies. This includes most advanced forms of AI such as continuously updating algorithms and deep learning, areas where most benefit to patient care is anticipated.
- The evaluation of AI health technologies has unique challenges including failure of generalizability when implemented at new sites, data drifts and shifts causing deterioration in performance over time, differences in performance due to variability in human-AI interaction and ethical concerns regarding perpetuation or amplification of racial or gender bias.
- Ophthalmology is considered a leader in the development and deployment of AI in medicine (5).
- Our group's previous work - SPIRIT-AI and CONSORT-AI - provide foundational principles for the evaluation of sophisticated AI elements; this work is cited by NICE, MHRA, NICE, NHSX, and WHO and has been used as a starting point for evaluation and regulatory frameworks they are now creating (6,7).
- The development of a NICE ESF for AI is a critical step for the NHS in its digital transformation, enabling innovators, trialists, evaluators and commissioners to work from a common standard, enhancing efficiency and encouraging further innovation.
- A balance must be struck however between setting an evidential bar that is too high, depriving the public of interventions that may benefit their health; and one that is too low, possibly exposing them to interventions that are unsafe or unfit for purpose.

Purpose

- To establish an Evidence Standards Framework for technologies incorporating AI (the *NICE ESF for AI Health Technologies*), facilitating commissioners to make evidence based decisions on their clinical and cost effectiveness whilst maintaining safety, ethicality, inclusivity and a productive research environment.

Methods

- Literature review** - with revamped search terms:
 - Classification systems, regulatory requirements, evaluation frameworks.
- Interview study** – semi-structured interviews with experts from key stakeholder groups to highlight issues not found above and to act as horizon scanning.
- Modified Delphi consensus study over multiple rounds** – to establish consensus on the form and content of a draft ESF.
- Validation and piloting** – a series of consultations with internal and external expert groups along with the assessment and evaluation of specific case studies.
- Dissemination** – a series of white papers, reports and academic papers will be released over the course of the project to communicate our findings to the wider research and regulatory community.



Delegation of tasks

- TM will work with VS and the wider UoB/Imperial/NICE team with involvement at each stage including literature searches, evaluation of sources, design of frameworks, stakeholder interviews and drafting of both the ESF and academic papers.

Results

- Progress so far -**
 - Secured award and set up partnership with NICE** - Jul-Aug 2021
 - Updated literature review** – Aug-Sep 2021-
 - 10, 500 abstracts identified; currently being screened to short-list possible frameworks
- Projected timeline -**
 - Stakeholder interviews** – October 2021
 - Delphi consensus study** – target Jan-Mar 2022
 - Draft ESF** – target Apr-Jun 2022
 - Validation and piloting** – target Jul-Sept 2022
- A final ESF will provide a roadmap for medical technologies incorporating AI to provide the necessary evidential standards in terms of safety, efficacy, inclusivity and equity that assures NICE and enables confident commissioning.
- This will provide an environment in which patients and healthcare systems can safely transition into the AI-enabled "Digital Age".
- This is an important opportunity for the speciality of ophthalmology and the West Midlands area to become leaders in safe, ethical digital transformation within the NHS.



Funding/Disclosure

- This work is funded by awards from NICE and University Hospitals Birmingham NHSFT

References

- NHS pressures in England: Waiting times, demand, and capacity [Internet]. [cited 2021 Sep 29]. Available from: <https://commonslibrary.parliament.uk/nhs-pressures-in-england-waiting-times-demand-and-capacity/>
- NHS Long Term Plan [Internet]. [cited 2021 Jun 25]. Available from: <https://www.longtermplan.nhs.uk/>
- X L, L F, AU K, SK W, DJ F, A B, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. The Lancet Digital health [Internet]. 2019 Oct 1 [cited 2021 Sep 29];1(6):e271–97. Available from: <https://pubmed.ncbi.nlm.nih.gov/33323251/>
- Evidence standards framework for digital health technologies | Our programmes | What we do | About | NICE [Internet]. [cited 2021 Sep 29]. Available from: <https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies>
- J DF, JR L, B R-P, S N, N T, S B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nature medicine [Internet]. 2018 Sep 1 [cited 2021 Sep 29];24(9):1342–50. Available from: <https://pubmed.ncbi.nlm.nih.gov/30104768/>
- X L, S CR, D M, MJ C, AK D. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nature medicine [Internet]. 2020 Sep 1 [cited 2021 Sep 29];26(9):1364–74. Available from: <https://pubmed.ncbi.nlm.nih.gov/32908283/>
- S CR, X L, AW C, AK D, MJ C. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Nature medicine [Internet]. 2020 Sep 1 [cited 2021 Sep 29];26(9):1351–63. Available from: <https://pubmed.ncbi.nlm.nih.gov/32908284/>