DIGITAL TECHNOLOGY

AI in the NHS: a framework for adoption

Authors: Bharadwaj V Chada[A] and Leanne Summers[B]

KEYWORDS: AI, artificial intelligence, NHS

DOI: 10.7861/fhj.2022-0068

Introduction

While we are encouraged by the promise shown by AI in healthcare, and more broadly welcome the use of digital technologies in improving clinical outcomes and health system productivity, we also recognise that caution must be exercised when introducing any new healthcare technology. Working with colleagues across the NHS Transformation Directorate, as well as the wider AI community, we have been developing a framework to evaluate AI-enabled solutions in the health and care policy context. The aim of the framework is severalfold but is, at its core, a tool with which to highlight to healthcare commissioners, end users, patients and members of the public the considerations to be mindful when introducing AI to healthcare settings. By way of a summary, the framework encompasses eight key considerations that policymakers are encouraged to discuss (Table 1).

Table 1. Summary of the AI considerations framework

Consideration	Descriptor questions
Context	What is the reason(s) for using AI in this context? How was the problem addressed previously? What value does AI add here?
Data	What data do we need to access to train the algorithm? Is it representative of the intended population?
Validation	Is the algorithm valid across geographies and over time? Is it generalisable across populations?
Implementation	How will the model work 'in practice'? What are the scope and limitations of using this algorithm? What training is needed for end users?
Surveillance	How is the performance of the algorithm monitored over time? How can users report errors with model performance?
Success metrics	Is the algorithm working as intended? Are we confident the benefits outweigh the costs (eg clinical, health economic and planetary)?
Ethics and governance	What safeguards are in place to protect against algorithmic bias? Is the algorithm transparent and explainable?
Managing change	What else needs to occur alongside introducing a new algorithm to achieve system-wide improvement?

Building on existing work

The past 5 years has seen a proliferation of academic publications and policy initiatives designed to support the deployment of AI in health and care settings, many of which have informed the development of our framework. Lovejoy *et al* outline a number of considerations for the use of AI in healthcare, with a particular focus on context and model design.[1] Similarly, Reddy *et al* present the 'Translational evaluation of healthcare AI' framework, centred around three main components (capability, utility and adoption) and associated subcomponents.[2] Meanwhile, other publications have focused specifically on ethical considerations surrounding the use of AI, such as digital exclusion and worsening clinical outcomes among minority populations.[3,4] These publications have been developed against a backdrop of significant policy activity, notably, these include the Central Digital and Data Office (CDDO) Data Ethics Framework, the Department of Health and Social Care (DHSC) guide to good practice and the National Institute for Health and Care Excellence (NICE) evidence standards framework (ESF) for digital health technologies. The lattermost of these is a particularly welcome addition to the policy landscape, with its focus on the economic impact of using healthcare AI; traditionally an overlooked field of study.

NHS policy initiatives in this context include the Artificial Intelligence Laboratory's (AI Lab's) *A buyer's guide to AI in health and care* (and associated template), which serves to support commissioners in the procurement of AI technologies; the NHS Digital Technology Assessment Criteria for health and social care (DTAC), which is a tool for healthcare organisations to evaluate suppliers through the lenses of user needs and security, as well as regulatory and technical compliance; and the *AI in healthcare: creating an international approach together* report, published jointly by the Global Digital Health Partnership (GDHP) and

Authors: [A]HealthTech innovation fellow, Harvard Medical School, Boston, USA; [B]head of AI strategy, NHS England, London, UK

NHS AI Lab, with the aim of providing AI policy guidance to the international health community.[5,6]

A framework for adoption

The framework presented here builds on, and is in many ways an amalgamation of, much of this work. Notably, it aims to reconcile both the 'ethical' considerations (such as algorithmic bias and transparency) as well as more 'operational' considerations (such as real-world implementation, post-market surveillance and, importantly, change management); as there appear to be few publications that traverse these twin objectives. The framework may temper some of the hype surrounding healthcare AI, and encourages users to be more holistic in their evaluation of AI technologies. The framework was developed following consultation with colleagues within the NHS Transformation Directorate, as well as the broader AI and life sciences community, across a wide range of domain expertise. The breadth of expertise allowed for the identification of several considerations; for example, clinical staff pointed to the importance of a 'lead responsible clinician' overseeing the rollout of a new technology, technical colleagues highlighted the need to monitor the change in an algorithm's performance over time, while ethics experts shed light on issues pertaining to bias and diversity.

While we recognise that the framework is far from exhaustive, we hope that it can, in time, be developed into a more robust assessment tool for healthcare commissioners to oversee the introduction of new technologies. It is also important to note the overlap that exists between each of these eight considerations; for example, 'ethics and governance' should necessarily underpin the entire AI design and deployment life cycle. Likewise, the financial implications of using AI in healthcare (presented here within 'success metrics') will be intertwined with other considerations, such as 'context', 'implementation' and 'managing change'.

Box 1. Scaling AI safely: a hypothetical vignette

Isabel is the clinical director of her local primary care network (PCN), and is also a GP partner at a large surgery in Buckinghamshire. She had been using SkinScanner for the previous 2 years, an AI-enabled tool that can help analyse skin lesions, determine the change in appearance over time and suggest appropriate next steps. Isabel's practice has used SkinScanner to good effect, helping to reduce the number of unnecessary referrals and inappropriate investigations, and Isabel is keen to support the rollout of the software elsewhere in the PCN. However, she recently reviewed a patient, Eric, whom she'd last seen 6 months previously with a mole on his left forearm, which she'd advised him to monitor. Eric hadn't noticed any change in appearance over those 6 months, and Isabel also measured the mole with her dermatoscope, and was reassured to see that it hadn't increased in size since Eric's last appointment. However, she was surprised to find that SkinScanner judged the lesion to have increased in size, and recommended urgent specialist review. She arranged a priority virtual consultation with a consultant dermatologist, Dr Sharma, who agreed with Isabel and Eric's assessment that the lesion was not worrisome and did not warrant further investigation.

Isabel contacted the SkinScanner team to investigate the discrepancy in advice, who promptly paused the software until they could find out more. They conducted a thorough retrospective analysis of all skin lesions assessed using SkinScanner, and found a problem in <1% cases with the measuring ability of their tool. These cases were reviewed by a panel of expert dermatologists and, thankfully, it was found that patient care hadn't been compromised as a result (though the patients were notified of the issue). The developer team at SkinScanner implemented a patch to fix the error and upgraded their software, and continued to enforce vigilant post-market surveillance (PMS).

Reassured that no harm had come to patients as a result, and once again confident in the newly-upgraded software, Isabel was eager to try to scale SkinScanner in a manner that is effective and safe.

GP = general practitioner.

Box 2. Summary of AI considerations

- It is important to be vigilant of algorithmic bias against certain patient populations (ie individuals with darker skin tones); for example, through having diverse research panels, clear ground truths that challenge clinical presuppositions and biases, and public and patient involvement and engagement (PPIE). The underlying datasets must also be carefully curated to ensure generalisability of the model's performance.
- The reason(s) for using AI, the expected indicators of success and any unanticipated consequences of change should be considered at an early stage. This includes conducting a clinical risk assessment, as well as a cost–benefit analysis, considering initial costs as well as costs associated with ongoing maintenance, training and service redesign.
- The algorithm must be continually re-examined for evidence of 'data drift' and changes in environmental conditions, such as the evolving prevalence of skin cancer and other dermatological conditions over time.
- There must be clear guidance and training for staff on the circumstances in which to use the model, and appropriate recourse to revert to model developers with any concerns over the algorithm's performance. Box 1 presents an example of a 'false positive' result, which is perhaps less worrisome than a 'false negative' result, with the latter presenting greater challenges of post-market surveillance and regulation as well as liability and optionality. If the algorithm had failed to identify a patient with skin cancer, where does the burden of accountability lie? This is an area in which there is little consensus, although it is our belief that AI is, for now, an augmentative tool, designed to supplement and not supersede clinician expertise. Clinical staff must maintain oversight and remain accountable for actioning (or not) the recommendations made by an AI model, until such time as there is clearer medico-legal guidance around indemnity and liability. Can we cautiously envisage an AI 'Bolam test' in years to come, wherein an AI algorithm cannot be deemed negligent when other similar algorithms arrive at the same conclusion?
- Policymakers must consider the broader health system dependencies and bottlenecks that must be addressed, alongside the introduction of AI in a given health and care context (eg improved access to treatment and the availability of specialist staff, pathway optimisation/redesign, and greater engagement with marginalised communities).

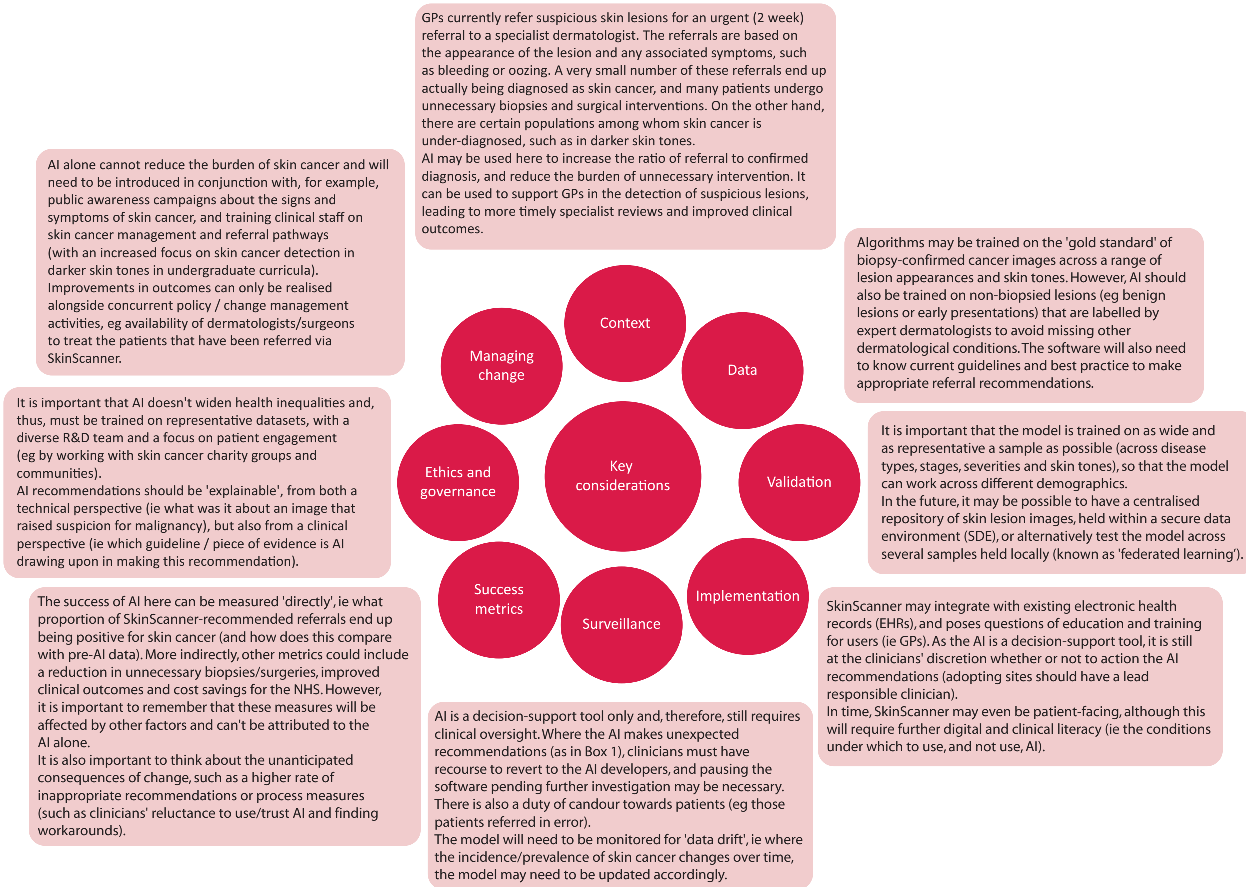

Fig 1. AI considerations framework applied to the vignette in Box 1. GP = general practitioner; R&D = research and development.

Nevertheless, the framework may serve as a useful guide in navigating the adoption of AI in a healthcare system.

We have developed a vignette (Box 1) to showcase a hypothetical use case for AI in the NHS. The framework has then been 'applied' to the vignette (Fig 1) to shed light on issues pertaining to the design and deployment of the algorithm that may otherwise have been overlooked; for example, the importance of algorithmic validation that is reflective of 'data drift' and changing environmental conditions (such as evolving disease prevalence), the scope and limitations of using the model in real-world clinical settings, and the policy measures that must occur in tandem with AI solutions to achieve meaningful system-wide improvement.

Summary and conclusion

This framework may aid policymakers in better understanding the AI landscape in a given health and care context, and highlight the ancillary factors that must addressed if AI is to be used as meaningfully as possible. In the case of the earlier vignette (Box 1), in which AI is being used to detect skin cancer, these factors may be summarised in Box 2.

We commend the excellent work that is currently taking place in using AI to address the highest priority areas of clinical need, and look forward to the more routine and widespread adoption of these tools in healthcare. The technologies must, however, be introduced carefully, using holistic evaluation criteria, multistakeholder engagement and ongoing performance monitoring. ■

Funding

This article and the framework presented here have been developed as part of the Faculty of Medical Leadership and Management (FMLM) National Medical Director's Clinical Fellow Scheme. We welcome feedback and comment in developing this framework further.

References

1 Lovejoy C, Arora A, Buch V, Dayan I. Key considerations for the use of artificial intelligence in healthcare and clinical research. *FHJ* 2022;9:75–8.
2 Reddy S, Rogers W, Makinen V *et al*. Evaluation framework to guide implementation of AI systems into healthcare settings. *BMJ Health Care Inform* 2021;28:e100444.
3 Ibrahim H, Liu X, Zariffa N, Morris A, Denniston A. Health data poverty: an assailable barrier to equitable digital health care. *Lancet Digit Health* 2021;3:e260–5.
4 Adamson A, Smith A. Machine learning and health care disparities in dermatology. *JAMA Dermatology* 2018;154:1247.

5 Joshi I, Cushnan D. *A buyer's guide to AI in health and care*. NHS, 2020.
6 Global Digital Health Partnership. *AI in healthcare: creating an international approach together*. GDHP, 2020.

Address for correspondence: Dr Bharadwaj V Chada, Center for Primary Care, Harvard Medical School, 2nd Floor, 635 Huntington Avenue, Boston, MA 02115, USA.
Email: bharadwaj_chada@hms.harvard.edu
Twitter: @bharadwajchada