BMJ Health & Care Informatics

Artificial intelligence projects in healthcare: 10 practical tips for success in a clinical environment

Anthony Wilson ,[1] Haroon Saeed,[2] Catherine Pringle,[3,4] Iliada Eleftheriou,[4] Paul A Bromiley,[4] Andy Brass[4]

To cite: Wilson A, Saeed H, Pringle C, *et al*. Artificial intelligence projects in healthcare: 10 practical tips for success in a clinical environment. *BMJ Health Care Inform* 2021;**28**:e100323. doi:10.1136/bmjhci-2021-100323

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjhci-2021-100323).

Received 07 February 2021
Accepted 23 May 2021

[1]Department of Adult Critical Care, Manchester University NHS Foundation Trust, Manchester, UK
[2]Department of Pediatric Ear Nose and Throat Surgery, Royal Manchester Children's Hospital, Manchester, UK
[3]Children's Brain Tumour Research Network, Royal Manchester Children's Hospital, Manchester, UK
[4]Division of Informatics, Imaging and Data Sciences, The University of Manchester, Manchester, UK

Correspondence to
Anthony Wilson;
anthony.wilson@mft.nhs.uk

ABSTRACT

There is much discussion concerning 'digital transformation' in healthcare and the potential of artificial intelligence (AI) in healthcare systems. Yet it remains rare to find AI solutions deployed in routine healthcare settings. This is in part due to the numerous challenges inherent in delivering an AI project in a clinical environment. In this article, several UK healthcare professionals and academics reflect on the challenges they have faced in building AI solutions using routinely collected healthcare data. These personal reflections are summarised as 10 practical tips. In our experience, these are essential considerations for an AI healthcare project to succeed. They are organised into four phases: conceptualisation, data management, AI application and clinical deployment. There is a focus on conceptualisation, reflecting our view that initial set-up is vital to success. We hope that our personal experiences will provide useful insights to others looking to improve patient care through optimal data use.

INTRODUCTION

There is a growing need for digital transformation in healthcare,[1] utilising the increasing availability of data to improve patient outcomes. Artificial intelligence (AI) is essential to this transformation; it is impossible for healthcare professionals (HCPs) to interpret the wealth of data available to them otherwise. This is reflected in a recent explosion of AI research using routinely collected healthcare data. It is envisaged that every HCP will use AI technology in the future.[2]

Despite this potential, the majority of AI healthcare solutions are yet to make a tangible impact on the frontline.[3 4] This in part reflects the practical barriers that such projects must overcome if they are to be successful. The majority of these barriers are predictable[5] and often similar, regardless of the project. They are a consequence of data privacy and ethical concerns, legislation and a lack of training and of trust among HCPs.

This paper offers 10 tips (table 1) to help AI projects to succeed in a clinical environment. They are based on the experiences of three clinical academics and three health informatics researchers who have experienced the practical barriers first-hand. The tips are categorised under the headings: conceptualisation, data management, application of AI and clinical deployment. This reflects the stages an AI healthcare project must progress through (figure 1) to go from early-stage, proof-of concept to a solution nearing clinical application. The tips range from generic, real-world strategies to those designed to highlight the complexities of using AI in healthcare. Brief biographies for each author are available in online supplemental appendix 1.

PART 1. CONCEPTUALISATION

1. Build a collaborative science team

No individual has all the skills and resources needed to make an AI healthcare project succeed so a collaborative science team (CST) is essential. The team's composition may vary but includes HCPs, data scientists and possibly statisticians, project managers and software engineers. This is the most important tip; the shared knowledge of the CST overcomes many barriers presented to an AI healthcare project. Some barriers are obvious to a data scientist, some to a HCP and vice versa. We propose that such teams are the building blocks of what Cosgriff *et al* have described as the clinical AI departments of the future.[6]

It is also important to identify the project 'gatekeepers' from the outset; senior individuals who may be aware of organisational barriers to success. They could be heads of strategy, department leads or chief informatics officers. Early meetings with gatekeepers are essential; they may recommend additional team members and can access useful resources if a project proposes to change clinical pathways.

Table 1 10 tips to get started with artificial intelligence healthcare projects

Phase 1: conceptualisation	
1	**Build a collaborative science team** Recruit widely, select complementary expertise. Engage gatekeepers early.
2	**Engage frequently with the end user** What is the *clinical* problem to be addressed? End user engagement is a process, not an event. Develop digital transformation, not isolated algorithms.
3	**Build collaboration agreements early** Check for existing agreements. Use contracts and IG teams. Share your experiences with other CSTs.
4	**Ethics: present a balanced view** Present a balanced view of the challenges and benefits of AI projects. Seek out ethics review boards with prior experience. Ethical review may be protracted.
5	**Invest in data science training for healthcare professionals in your team** A common language is needed for the CST to work at its best.
Phase 2: data acquisition and preparation	
6	**Do not underestimate the challenge of data extraction** Understand what data you really need. Identify a data champion. Engage healthcare informatics teams early; it will take longer than you expect. Consider building or contributing to a meta-data catalogue at your institution.
7	**Protect patients' data** Give serious consideration to how you will anonymise and protect patient data from the outset of the project.
8	**Remember that healthcare data varies in quality and reliability** Healthcare data are influenced by a wide variety of factors. Clinicians can play a key role in interpreting and overcoming quality and reliability issues.
Phase 3: AI application	
9	**Design AI that can be trusted with patient care** Use the right algorithms for your data and your end user. Strive for generalisable, well-validated solutions.
Phase 4: translation	
10	**Be mindful of medical device regulation** Be mindful of the requirements for medical device regulation as your healthcare data science project progresses.

AI, artificial intelligence; CST, collaborative science team; IG, information governance

2. Engage frequently with the end user

HCPs often have little or no experience with AI. Many will be unfamiliar with the capabilities and limitations of AI[7] which can lead to both unachievable expectations or unfair dismissal from the outset. In addition, HCPs may struggle to specify in advance how they want AI solutions to improve their workflow. Some may have never even mapped out or quantified their daily work. What they *say* they want may change as the project develops. There is a complex interaction of human factors at play which CSTs must not ignore.

Healthcare AI projects must therefore strive to offer digital transformation through 'organisational, service and social innovation' as suggested by Creswell *et al*.[8] The clinical AI departments of the future must be prepared to offer in-reach into the other clinical departments. Defining the problem to be solved may take considerable time and may require CSTs to study clinical workflows carefully, perhaps even shadowing HCPs. In addition to building the solution, consideration should be given to how it will be employed. For example: how it is presented in a human understandable way,[9] its impact on other workflows, on patient–caregiver interaction, the risks of automation complacency[10] and of biasing clinical decision making.[4]

End user engagement is not a single event. We encourage CSTs to adopt an agile approach, where continuous user engagement with rapid modifications creates healthcare tools of real value. There is a tension here between agile development and ensuring patient safety as an AI solution is implemented and refined.

3. Build collaboration agreements early

Although healthcare organisations and universities frequently collaborate, overarching agreements for multisite working and secure data sharing are rare. Therefore, collaboration agreements need to be developed early to avoid major delays later. Conversely, generating new agreements for every project is inefficient. Creating communities of practice can facilitate dialogue between CSTs, supporting shared solutions or lobbying both organisations for high-level agreements.

Data-flow diagrams are an important part of any collaboration agreement (figure 2). They help design a transparent and detailed application for ethical review, and aid

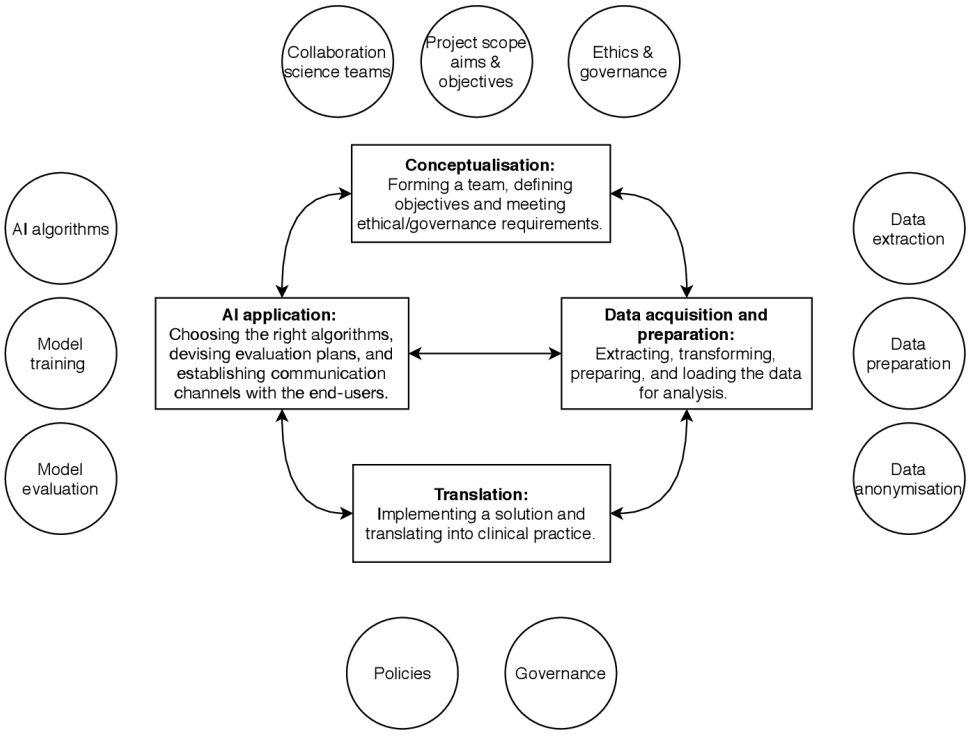

Figure 1 The four typical phases which AI healthcare projects will go through. The circular form highlights the iterative way in which such projects are often conducted. AI, artificial intelligence.

in the formation of a data management plan, which is a prerequisite for collaboration.

Many healthcare organisations and universities have specific teams to help create collaboration agreements. Such teams can also advise on project costs, intellectual property, publication rights, data-sharing agreements and information governance. While drafting agreements, it is important to identify the *data controller*, who designs and supervises the project, and *data processors*, who carry out processing tasks under instruction from the controller.

4. Ethics: present a balanced view

Ethical oversight of AI-driven healthcare research can be problematic. There are numerous tension points that need to be resolved including the role of patient consent in the use of routinely collected healthcare data,[11] the potential for under-representation of certain patient groups in AI projects[12] and concerns about adaptive algorithms whose performance may change with new data. Most approval committees still have limited experience of AI and the review process may be ill-suited to AI-driven research.

The ethical review process can be improved if CSTs follow best practice such as the Standard Protocol Items: Recommendations for Interventional Trials-AI (SPIRIT-AI) protocol guidelines.[13] A moral justification for transitioning to learning health systems supported by data science has been proposed,[14] which may also help to inform the review process. A balanced view of AI healthcare projects should be offered; while AI presents ethical challenges it also offers opportunities to enhance the quality of existing medical evidence in areas where evidence is lacking or subject to bias.[12]

CSTs should be aware that ethical review may be protracted and should consider seeking review by committees with experience of similar projects

5. Invest in data science training for healthcare professionals in your team

There is currently a lack of data science training in healthcare education programmes. Where training programmes do exist, they are often optional. The knowledge demanded of HCPs focuses on interpretation of research evidence, screening tests and the output of randomised controlled trials. Although this is an excellent foundation for traditional medical research, it does not provide a common language for data-driven clinicians of the future. Online supplemental appendix 2 offers a glossary to help build this common language.

The Topol Review[15] emphasised the need for data science training if NHS staff are to reap the benefits of the digital revolution. There are a range of resources that HCPs can use to learn about the principles of tidy data,[16] how to categorise and address missingness in their data[17] and to gain an overview of AI techniques and their limitations.[7] There are communities of practice[18] who can offer support and growing opportunities for additional self-directed learning.[19]

We recommend that HCPs pay attention to these learning needs and use the expertise of data scientists in their team to guide them. Such preparation can prevent

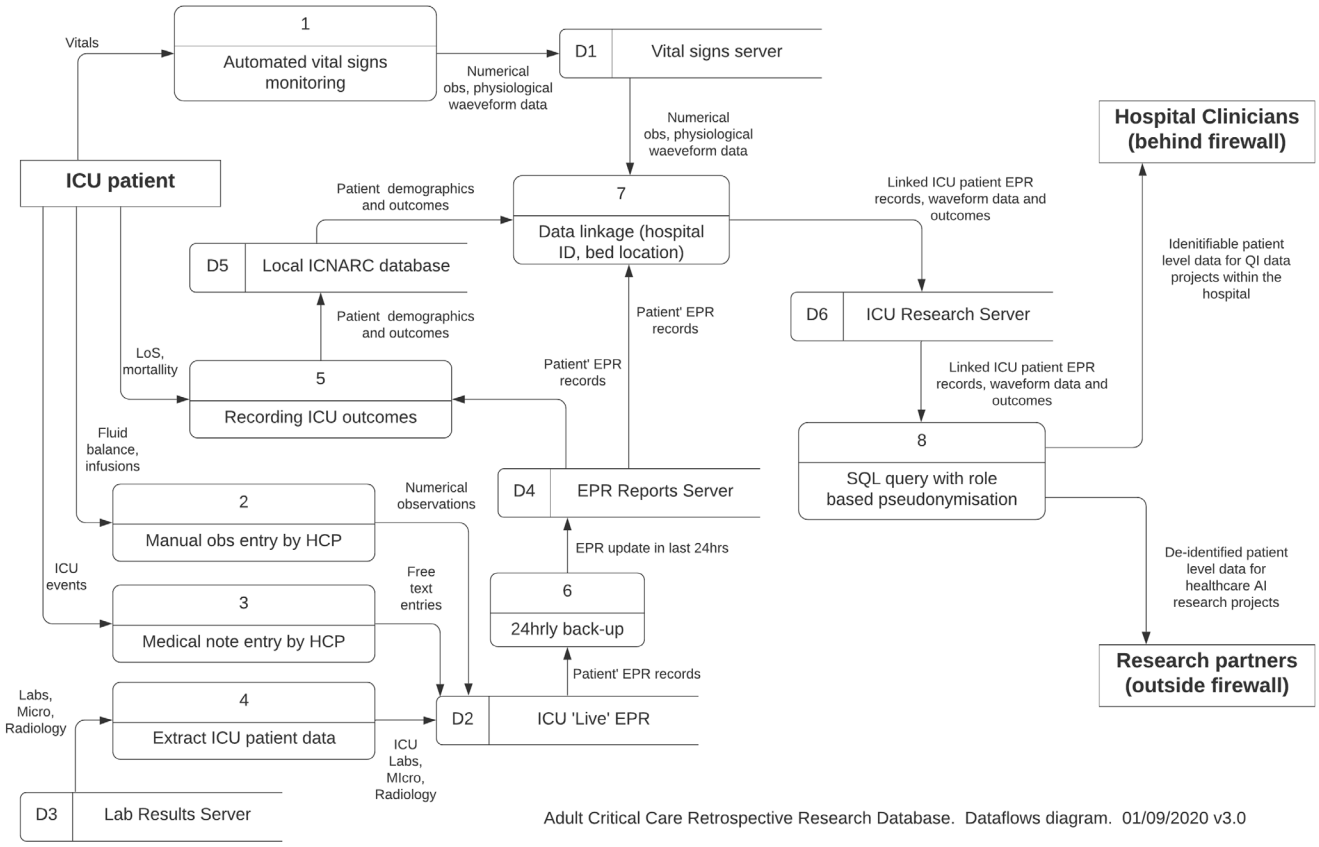

Figure 2 An example of a data flow diagram using the Gane and Sarson nomenclature. In this example, multiple disparate sources of data from individual patients in the ICU are aggregated into a single research database. In-built, role-based access controls allow the data to be accessed by multiple different users while meeting data privacy requirements. EPR, electronic patient record; HCP, healthcare professional; ICNARC, intensive care national audit and research centre; ICU, intensive care unit; SQL, structured query language.

errors of understanding and improve interactions throughout an AI healthcare project.

PART 2. DATA ACQUISITION AND PREPARATION

6. Don't underestimate the challenge of data extraction

Figure 2 illustrates a process of data extraction and linkage in critical care where significant effort was needed to link three disparate data sets before any AI solutions could be considered. Sharma *et al*[20] have highlighted how this lack of data integration is a major barrier to the success of clinical risk prediction tools. HCPs use a huge range of information technology (IT) systems which are not always integrated and are often supported by paper records. These systems were designed for healthcare delivery and manufacturers may offer limited support for research. Furthermore, there is no universal standard for coding and capture of healthcare data despite laudable efforts to promote this.[21 22] Coupled with local variation,[23] this makes it hard to aggregate or extend data sets between institutions. CSTs should anticipate the need for extensive data wrangling to align data sets from different systems and institutions for use in their research.

Faced with this complexity, help from healthcare IT teams is essential both to access the data, which is often stored on secure servers, and to navigate the data structure, which is complicated in most electronic health records. Unfortunately, many healthcare IT teams are overwhelmed by day-to-day operational work leaving little time to support healthcare AI projects.

CSTs should make allowances for these data encoding and extraction challenges. Accurately defining data required for the project may improve the interaction with hospital IT teams and avoid the need for time-consuming, repeated extraction. Consider promoting and contributing to a meta-data catalogue at your institution,[24] so that other CSTs can benefit from the extraction work you have done.

Identifying a 'data champion'; a gatekeeper (tip 1) who supports the proposed project and can leverage the IT team can be helpful. Even better is to make the healthcare IT team part of the CST. CSTs are often seen as 'service users' by healthcare IT teams. Try to move beyond this relationship by recognising and rewarding the contribution of the healthcare IT team in the research.

7. Protect patients' data

At present, most patients support the use of their data to improve healthcare,[25] but large-scale leaks of confidential information could easily damage trust. Patient data must be anonymised or pseudonymised prior to analysis or storage somewhere other than where it was collected. However, the proliferation of digital data increases the risk of inadvertent release of identifiable information, or identification through deductive disclosure.[26] Moreover, healthcare records often contain unexpected identifiers such as pixelated names in image files or names in narrative entries.

One solution is to bring data scientists and data processing capabilities into the healthcare organisation. An alternative is a trusted research environment in which identifiable patient data are stored with the required analysis tools.[27] Such environments incorporate strict access controls that mitigate many anonymisation concerns. Figure 2 illustrates a hybrid approach in which role-based access control to a research database applies deidentification at source depending on a user's credentials.

Regardless of approach, the aim is to ensure that data are extracted in a manner that allows fluid access for analysis, while adhering to information governance regulations. Protecting patient data is a process, not a one-time event. Teams should develop a strategy to prevent the inadvertent leakage of patient information and consider how they will periodically test and monitor its effectiveness.

8. Remember that healthcare data varies in quality and reliability

Automatically captured observations, such as clinical images or continuous vital signs exemplify the 'cleanest' sources of healthcare data. At the other end of the spectrum are those that rely on manual input (eg, subjective examinations and medical history). As Sujan *et al*[10] note, not all events are captured in the medical record which may make it hard for AI algorithms to function optimally. Furthermore, an individual's healthcare record rarely describes the wider context of their care. Periods of bed pressure and emergency situations are examples of circumstances that may affect an individual's record, for example, by influencing the timing of hospital discharge. This wider context reduces the reliability of some seemingly definitive, time-specific events. It can help to try to identify 'anchors':[28] events in the healthcare record that are independent of these human factors.

Paying attention to data quality and reliability can pay huge dividends. It has been demonstrated that data cleaning and artefact removal can yield greater improvements in model prediction than applying more sophisticated AI algorithms to the original data.[29] CSTs should go to considerable lengths to address missing values and artefacts and to understand the clinical nuances of their variables.

PART 3. AI APPLICATION

9. Design AI that can be trusted with patient care

AI healthcare projects must be trusted to enable safe patient care. Engaging frequently with the end user (tip 2) during the project and educating HCP in data science techniques (tip 5) can go a long way towards this aim. However, the CST should also ensure that they tailor their algorithm to the data available, promote transparency where possible and make it clear how the AI solution will support rather than replace HCPs. In meeting these aims, the UK government's code of conduct for data-driven health technology,[30] research standards such as CONSORT-AI[31] and clinician focussed checklists[9] are valuable resources.

CST's should tailor the algorithms they use towards characteristics of the data set available. For example, deep neural networks offer excellent performance with large data sets but are prone to overfitting when used for smaller, high dimensional data sets, which may be the only option in some clinical settings. A plethora of other techniques including random forests, linear regression, support vector machines and K nearest neighbours[7] can be utilised and ensembled to produce tangible results in these circumstances.

Choosing the correct algorithm also involves addressing the concerns of HCPs who may worry about the use of AI algorithms in clinical practice if they cannot explain them.[7] In this context, so-called black-box solutions[32] may need to be avoided. Alternatively, CSTs can use a variety of techniques such as graphical explanations, model-agnostic explainers, heuristic examples or counterfactual explanations that may offer insight into how a black-box model reached its conclusions.[33]

Degradation of algorithm performance must be mitigated to minimise error and patient safety issues. The CST must ensure that AI algorithms are as generalisable as possible, functioning equitably for all patient groups and across geographic regions. Algorithms should also demonstrate robustness to the introduction of new data (eg, not prone to 'catastrophic forgetting'[34]). In order to achieve this, it is paramount that there is expertise within the CST to internally validate AI algorithms. There should be a clear plan to externally validate the algorithm with unseen data, using an appropriately sized demographic and geographic cohort across an appropriate time frame.

PART 4. TRANSLATION

10. Be mindful of medical device regulation

Many CSTs may be uncertain about the regulatory status of the AI solutions they develop. Existing US and EU guidelines are reasonably clear:[35] AI software is classed as a medical device if its purpose is specifically medical and its function goes beyond data curation and transmission. There are subtleties related to 'intended purpose' and what constitutes a 'medical purpose' that are beyond the scope of this paper; for an excellent summary see Ordish *et al*.[35]

Medical device regulation is unlikely to impede the initial phases of AI development, but has significant implications for marketing or clinical application. Regulatory challenges include the adaptive nature of some algorithms, where outputs change substantially as new data are processed and the aforementioned black-box dilemma. Initiatives such as the Food and Drug Administration (FDA) AI framework[36] and the National Institute for Health and Care Excellence (NICE) evidence standards framework for digital health technologies[37] have shifted the emphasis to the quality of the algorithm development process, along with postimplementation quality assurance measures, rather than the need to meet medical device standards at a specific time.

CSTs should be mindful of such regulations during project planning. The NHS X AI governance framework[38] is an example of best practice principles that may aid this process.

SUMMARY

There is a huge appetite for digital transformation in healthcare and significant potential for AI solutions. Launching an AI healthcare project can be difficult due to a range of organisational and regulatory hurdles coupled with complexities related to the use of AI in healthcare.

CSTs are critical to starting an AI healthcare project. They need a shared, common language to collaborate effectively. They should involve (and educate) the end user from the outset and be prepared for extensive data wrangling. There may be uncertainty concerning authorisations and regulations. Finally, they must build trust in AI techniques by developing solutions that are generalisable, interpretable and user focused.

Twitter Iliada Eleftheriou @IliadaEle

Acknowledgements The authors are grateful to Dr Michal Pruski and Dr Ang Davies for their reviews of the manuscript and their helpful suggestions.

Collaborators Michal Pruski, Ang Davies.

Contributors All authors have contributed equally to the manuscript.

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 None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; internally peer reviewed.

Data availability statement No data are available to share.

ORCID iD
Anthony Wilson http://orcid.org/0000-0001-7482-5337

REFERENCES

1 The future of healthcare: our vision for digital, data and technology in health and care - GOV.UK. Available: https://www.gov.uk/government/publications/the-future-of-healthcare-our-vision-for-digital-data-and-technology-in-health-and-care/the-future-of-healthcare-our-vision-for-digital-data-and-technology-in-health-and-care [Accessed 7 Feb 2020].
2 Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019;25:44–56.
3 Panch T, Mattie H, Celi LA. The "inconvenient truth" about AI in healthcare. *NPJ Digit Med* 2019;2:1–3.
4 Lyell D, Coiera E, Chen J, *et al*. How machine learning is embedded to support clinician decision making: an analysis of FDA-approved medical devices. *BMJ Health Care Inform* 2021;28:100301.
5 Singh RP, Hom GL, Abramoff MD, *et al*. Current challenges and barriers to real-world artificial intelligence adoption for the healthcare system, provider, and the patient. *Transl Vis Sci Technol* 2020;9:45.
6 Cosgriff CV, Stone DJ, Weissman G, *et al*. The clinical artificial intelligence department: a prerequisite for success. *BMJ Health Care Inform* 2020;27:100183:e100183.
7 Rudin C, Carlson D. The secrets of machine learning: ten things you wish you had known earlier to be more effective at data analysis. *Oper Res Manag Sci Age Anal* 2019:44–72.
8 Cresswell K, Williams R, Carlile N, *et al*. Why digitally-enabled health system transformation needs different forms of innovation. *BMJ Health Care Inform* 2020;27:100173.
9 Scott I, Carter S, Coiera E. Clinician checklist for assessing suitability of machine learning applications in healthcare. *BMJ Health Care Inform* 2021;28:e100251.
10 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;26:100081.
11 Heslop PA, Davies K, Sayer A, *et al*. Making consent for electronic health and social care data research fit for purpose in the 21st century. *BMJ Health Care Inform* 2020;27:100128.
12 McCoy LG, Banja JD, Ghassemi M, *et al*. Ensuring machine learning for healthcare works for all. *BMJ Health Care Inform* 2020;27:100237.
13 Cruz Rivera S, Liu X, Chan A-W, *et al*. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat Med* 2020;26:1351–63.
14 Faden RR, Kass NE, Goodman SN, *et al*. An ethics framework for a learning health care system: a departure from traditional research ethics and clinical ethics. *Hastings Cent Rep* 2013;Spec No:S16-27:S16–27.
15 Topol E. Preparing the healthcare workforce to deliver the digital future the Topol review. an independent report on behalf of the Secretary of state for health and social care. *Nhs* 2019;102 https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf
16 Wickham H. Tidy data. *J Stat Softw* 2014;59:1–23.
17 Little RJA, Rubin DB. *Statistical analysis with missing data*. 3rd Ed. John Wiley & Sons, Ltd, 2019.
18 NHS-R Community - Promoting the use of R in the NHS. Available: https://nhsrcommunity.com/ [Accessed 7 Feb 2020].
19 AI for Healthcare and Digital Transformation Course - FutureLearn. Available: https://www.futurelearn.com/courses/artificial-intelligence-in-healthcare [Accessed 14 Feb 2020].
20 Sharma V, Ali I, van der Veer S, *et al*. Adoption of clinical risk prediction tools is limited by a lack of integration with electronic health records. *BMJ Health Care Inform* 2021;28:100253.
21 SNOMED CT - NHS Digital. Available: https://digital.nhs.uk/services/terminology-and-classifications/snomed-ct [Accessed 7 Feb 2020].
22 OMOP common data model – OHDSI. Available: https://www.ohdsi.org/data-standardization/the-common-data-model/ [Accessed 7 Feb 2020].
23 Kirkendall ES, Ni Y, Lingren T, *et al*. Data challenges with real-time safety event detection and clinical decision support. *J Med Internet Res* 2019;21:1–12.
24 Mauro data mapper / Oxford metadata catalogue. Available: https://metadata-catalogue.org/ [Accessed 20 Apr 2021].
25 How do people feel about the use of data? | understanding patient data. Available: https://understandingpatientdata.org.uk/how-do-people-feel-about-use-data [Accessed 7 Feb 2020].

26 Kaiser K. Protecting respondent confidentiality in qualitative research. *Qual Health Res* 2009;19:1632–41.
27 Ainsworth J, Buchan I. Combining health data uses to Ignite health system learning. *Methods Inf Med* 2015;54:479–87.
28 Halpern Y, Choi Y, Horng S, *et al*. Using anchors to estimate clinical state without labeled data. *AMIA Annu Symp Proc* 2014;2014:606–15.
29 Johnson AE, Kramer AA CG. Data preprocessing and mortality prediction: The Physionet/CinC 2012 challenge revisited - IEEE Conference Publication. *Comput Cardiol* 2014:157–60 https://ieeexplore.ieee.org/document/7043003
30 Code of conduct for data-driven health and care technology - GOV.UK. Available: https://www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology/initial-code-of-conduct-for-data-driven-health-and-care-technology [Accessed 30 Dec 2019].
31 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.
32 Black-box vs. white-box models - Towards Data Science. Available: https://towardsdatascience.com/machine-learning-interpretability-techniques-662c723454f3 [Accessed 7 Feb 2020].
33 Molnar C. Interpretable machine learning. *A Guide for Making Black Box Models Explainable*, 2019. Available: https://christophm.github.io/interpretable-ml-book/
34 McCloskey M, Cohen NJ. *Catastrophic interference in Connectionist networks: the sequential learning problem*, 1989: 109–65.
35 Ordish J, Murfet H, Hall A. Algorithms as medical devices. Cambridge, UK; 2019. https://www.phgfoundation.org/report/algorithms-as-medical-devices [Accessed 07 Feb 2021].
36 US Food and Drug Administration. Artificial intelligence and machine learning in software as a medical device FDA; 2019: 1–20. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device [Accessed 07 Feb 2021].
37 Evidence standards framework for digital health technologies | our programmes | what we do | about | NICE. Available: https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies [Accessed 20 Jan 2020].
38 Joshi I, Morley J. Artificial intelligence: how to get it right. putting policy into practice for safe data-driven innovation in health and care. London, United Kingdom; 2019. https://www.nhsx.nhs.uk/media/documents/NHSX_AI_report.pdf [Accessed 07 Feb 2021].

SUPPLEMENTARY DIGITAL CONTENT

Getting started with artificial intelligence healthcare projects: ten tips from the frontline

APPENDIX 1: Author Biographies

Dr Anthony Wilson (AW)
Anthony is a consultant in anaesthesia and critical care medicine at Manchester University NHS Foundation Trust. He is the research informatics lead for the deployment of the Trust's new, comprehensive electronic patient record. As a PhD scholar at the University of Manchester, he is working to link disparate sources of routinely collected ICU data and to apply AI techniques that utilise this data for better patient care. He is currently exploring the factors that make data-derived recommendations acceptable to clinicians at different points of care.

Mr Haroon Saeed (HS)
Haroon is an ear nose and throat specialist trainee and University of Manchester PhD scholar. His research aims to create a clinical prediction model for congenital progressive hearing loss using machine learning techniques. He has successfully built a collaborative science team who have optimised clinical data collection, navigated the process of research ethics committee approval and created a data processing agreement encompassing multiple NHS organisations. From the outset Haroon has worked to ensure the prediction model has maximum chance of being used on the clinical frontline by engaging with end users throughout development and acting early to ensure methods for model reproducibility and generalisability.

Miss Catherine Pringle (CP)
Catherine is a neurosurgery specialist trainee and University of Manchester PhD scholar. Her PhD focuses on the potential roles of machine learning approaches in outcome prediction and risk stratification of paediatric brain tumours. Paediatric brain tumours present a challenging patient cohort due to the relatively small overall numbers which generate a high dimensional data problem; a deep, multi-faceted, variable rich data set generated from a small number of patients to which traditional survival outcome statistical analysis is often difficult to apply. Through the application of machine learning techniques to radiological, biological and clinical data that accompanies these patients, Catherine is aiming to identify potential biomarkers of tumour diagnosis, risk-stratification and prognosis for children's brain tumours.

Dr Iliada Eleftheriou (IE)
Iliada is a lecturer in healthcare sciences at the University of Manchester. Her expertise lies in mapping complex data landscapes in healthcare settings to identify and address socio-technical challenges stemming from disparate information systems and data formats. She is also an academic consultant at The Christie NHS Foundation Trust where she investigates the feasibility of embedding artificial intelligence models in existing pathways to improve patient outcomes. She is leading a project to automate the validation of chemotherapy prescription regimes in the cancer care setting. The validation process is based on well-defined protocols and pharmacists can use AI-powered decision support tools to accommodate increasing numbers of patients undergoing chemotherapy.

Dr Paul A. Bromiley (PB)
Paul is a Lecturer in health data sciences at the University of Manchester. His research involves the development of machine learning based computer aided diagnostic systems for use in radiology, with a particular focus on musculoskeletal radiology and neuroradiology. He developed the machine learning software used in the ASPIRE™ teleradiology service marketed by Optasia Medical Ltd. (Manchester, UK; www.optasiamedical.com). ASPIRE™ automates the diagnosis of vertebral fragility fractures visualised incidentally in CT images, allowing hospitals to improve diagnostic rates for osteoporosis without increasing radiology department workloads.

Prof Andy Brass (AB)
Andy has over 25 years' experience in building capacity for informatics in healthcare, including developing the world's first Masters in Bioinformatics. His current focus is on working in partnership with Health Education England and the National School of Healthcare Sciences to develop career pathways in bioinformatics and data science for the NHS. As a part of this activity he has worked with NHS trusts across the UK to explore the challenges faced in embedding these new professions and methodologies within the existing workforce and working practices.

APPENDIX 2: GLOSSARY OF TERMS

The following glossary of terms may be useful for health care professionals in collaborative science teams as they seek to develop a shared vocabulary with their data scientist colleagues. The terms are divided into two parts, general healthcare data science terms and AI-specific terms. The former provides a minimum shared vocabulary for all project team members, whilst the latter may be of use to HCPs wishing to gain a deeper understanding of data science techniques.

Term	Definition
Healthcare Data Science Terms	
Algorithm	A formula or set of rules for performing a task. In Artificial Intelligence, the algorithm tells the machine how to go about finding answers to a question or solutions to a problem. Link to article by AI Glossary – AI Trends
Artificial Intelligence (AI)	Artificial Intelligence (AI) has numerous definitions which is due to how the interpretation of the concept of intelligence and also how the field has evolved over time. In this paper, we refer to AI as the following: “Refers to a broad field of science encompassing not only computer science but also psychology, philosophy, linguistics and other areas. AI is concerned with getting computers to do tasks that would normally require human intelligence.” (Stefan van Duin and Naser Bakshi, 2017). Link to article “Part 1: Artificial Intelligence Defined” on Deloitte’s website “the science of making machines do things that would require intelligence if done by people” (McCarthy and Minsky,1955) Link to article by John McCarthy on Stanford University
Big Data	Represents large amounts of data that are unmanageable using traditional software or internet-based platforms. It surpasses the traditionally used amount of storage, processing and analytical power. Link to article by Dash, S., Shakyawar, S.K., Sharma, M. et al. Big data in healthcare: management, analysis and future prospects. J Big Data 6, 54 (2019)
Cloud Computing	The practice of using a network of remote servers hosted on the Internet to store, manage, and process data. Link to article by Matt Dryfhout and Scott Hewer on Scout Technology Guides (2019)

Decision Trees	Decision tree learning is commonly used in data mining, aiming to create a model that predicts the value of a target variable based on several input variables. A classification tree models discrete target variables, where leaves represent final class labels and branches represent conjunctions of features that lead to class assignment, such as tumour classification. Regression tree models have continuous value target variables, for example length of stay in hospital. Link to book by Rokach, Lior; Maimon, O. (2008). Data mining with decision trees: theory and applications. World Scientific Pub Co Inc. ISBN 978-9812771711.
Deep Learning	A subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks. Link to article by Jason Brownlee on Machine Learning Mastery (2019)
General AI	Also referred to as strong AI or deep AI, is the concept of a machine with general intelligence that mimics human intelligence and/or behaviours, with the ability to learn and apply its intelligence to solve any problem. General AI can think, understand, and act in a way that is indistinguishable from that of a human in any given situation. AI researchers and scientists have not yet achieved strong AI. To succeed, they would need to find a way to make machines conscious, programming a full set of cognitive abilities. Link to article by Serena Reece on Codebots (2020)
Machine Learning	The study of computer algorithms that improve automatically through experience. It is seen as a subset of artificial intelligence. Machine learning algorithms build a mathematical model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so. Examples of machine learning algorithms are: Random Forest, Nearest Neighbour, Support Vector Machine (SVM), Deep Learning, etc. Link to book by Sammut, C., & Webb, G. I. *Encyclopedia of machine learning and data mining*. Springer Publishing Company, Incorporated. (2017).
Narrow AI	Also known as weak AI, is a specific type of artificial intelligence in which a technology outperforms humans in some very narrowly defined task. Unlike general artificial intelligence, narrow AI focuses on a single subset of cognitive abilities and advances in that spectrum. Link to article by Techopedia (2020).

Natural Language Processing (NLP)	A branch of artificial intelligence that deals with the interaction between computers and humans using natural language. The ultimate objective of NLP is to read, decipher, understand, and make sense of the human languages, in both speech and written form, in a manner that is valuable. Most NLP techniques rely on machine learning to derive meaning from human languages. Link to article by Jason Browniee (2019)
Neural Network	Artificial Neural Networks (ANNs) consist of a layered network of nodes between input and output layers, with weighted relationship connections between each node. The weight of connection increases or decreases the strength of a signal at a connection, and auto-adjusts as the model learns Link to article on Wikipedia
Robotics	Robotics – is an interdisciplinary branch of engineering and science that includes mechanical engineering, electronic engineering, information engineering, computer science, and others. Robotics deals with the design, construction, operation, and use of robots, as well as computer systems for their control, sensory feedback, and information processing. Link to Wikipedia Robotics page
Semi Structured Data	Semi-structured data does not have the same level of organisation and predictability of structured data associated with databases or data tables, but does contain searchable and retrievable tags or labels. These tags, labels or codes separate semantic elements from the data and enforce hierarchies of records and fields within the data, and is also known as a self-describing structure. Items within a semi-structured class may have different attributes despite being grouped together. Link to article by Buneman, Peter. "Semi structured data." Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems. 1997.
Structured Data	Structured data describes the information that can be stored and viewed in a consistent, organised and reproducible manner, typically in a standardized format such as tables with headings in columns and values in rows. This type of data can be easily read by a machine, analysed over time, and validated against expected values or biologically plausible thresholds. Link to book by Cormen, Thomas H.; Leiserson, Charles E.; Rivest, Ronald L.; Stein, Clifford (2009). Introduction to Algorithms, Third Edition (3rd ed.). The MIT Press.

Unstructured data	This type of data doesn't come pre-defined or with a pre-defined data model, therefore it lacks the organisation and precision of structured data. As a consequence, this type of data cannot be analysed easily by a machine. In order to be used as input to an algorithm, the data needs to be transformed into a structured form and often needs to be manually interpreted and analysed. Unstructured data is often text heavy, any may also contain dates, numerical values and factual information, making the application of traditional programs difficult when compared to structured data. Link to Wikipedia page on Unstructured data

AI - technical terms	
Error-driven learning	This is a sub-field of machine learning concerned with how an agent ought to take actions in an environment so as to minimize some error feedback. It is a type of reinforcement learning. Link to error-driven learning page on Wikipedia
Generative Adversarial Network	Algorithmic architectures that use two neural networks, pitting one against the other (thus the "adversarial") in order to generate new, synthetic instances of data that can pass for real data. They are used widely in image generation, video generation and voice generation. Link to article by Ian Goodfellow "NIPS 2016 tutorial: Generative adversarial networks." arXiv preprint arXiv:1701.00160 (2016).
Hidden Layer	In neural networks, a hidden layer is located between the input and output of the algorithm, in which the function applies weights to the inputs and directs them through an activation function as the output. In short, the hidden layers perform nonlinear transformations of the inputs entered into the network. Hidden layers vary depending on the function of the neural network, and similarly, the layers may vary depending on their associated weights. Link to article on DeepAI.
Junction tree algorithm	Also known as 'Clique Tree' is a method used in machine learning to extract marginalization in general graphs. In essence, it entails performing belief propagation on a modified graph called a junction tree. The graph is called a tree because it branches into different sections of data; nodes of variables are the branches. Link to article on Wikipedia

KNN (Nearest Neighbour)	Nearest neighbour algorithms classify a test example by finding its closest neighbours in a multidimensional feature space populated by known examples from a reference (training) data set. The class prediction is estimated to be that of the nearest neighbour, or by a weighted average of the classes of the k nearest neighbours. Link to article by Nisbet, R., Elder, J., & Miner, G. Handbook of statistical analysis and data mining applications. Academic Press. (2009)
Logistic Regression	A supervised learning classification algorithm used to predict the probability of a target variable. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes, but there can be two more categories of target variables that can be predicted by it: multinomial and ordinal. Logistic regression can be used for various classification problems such as spam detection, diabetes prediction, cancer detection etc. Link to article by TutorialsPoint.
Multiple Imputation	Imputation is an approach that deals with missing data within a dataset by imputing values based on a statistical method. Single imputation of missing values often fails to account for the uncertainty about the missing values. Multiple imputation, allows for the uncertainty about the missing data by creating several different plausible imputed data sets and appropriately combining results obtained from each of them. Link to article by Sterne, J. A., White, I. R., Carlin, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. Bmj, 338, b2393. (2009).

Overfitting	Overfitting occurs when a supervised learning algorithm becomes over familiar with a small training set of data, and may therefore fail to fit and analyse newly presented data reliably. The contrary position to this is Underfitting. Link to article by Tetko IV, Livingston DJ and Luik AI. Neural Network studies.1. Comparison of overfitting and overtraining. *J. Chem. Inf. Comput. Sci.* 1995, 35, 5, 826–833.
Principal component analysis (PCA)	Principal component analysis (PCA) simplifies the complexity in high-dimensional data while retaining trends and patterns. It does this by transforming the data into fewer dimensions, which act as summaries of features. PCA projects the data onto a few principal component (PC) directions, without losing too much information about the subjects.

	Link to article by Lever, J., Krzywinski, M. & Altman, N. Principal component analysis. *Nat Methods* **14,** 641–642 (2017).
Random Forests	Random forests use multiple decision trees to enhance the classification and regression outputs at time of training, with the mode of the classes (classification) or mean prediction (regression) of the contributing trees generating the final output. Random decision forests correct for decision trees' habit of overfitting to their training data set. Link to book Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome (2008). The Elements of Statistical Learning (2nd ed.). *Springer*. ISBN 0-387-95284-5.
Reinforcement Learning	The process of teaching a machine to make specific decisions using trial and error in order to optimize output. Reinforcement learning is one of three basic machine learning paradigms, alongside supervised learning and unsupervised learning. However, it differs as it does not require labelled input and output data nor correction of poor outcomes. Instead it aims to balance exploration of uncharted data and exploitation (of current knowledge). Link to paper: Kaebling L,; Littman ML.; Moore AW. (1996). “Reinforcement Learning: A Survey”. Journal of Artificial Intelligence Research. :4: 237-285.

Statistical Shape modelling	Statistical shape models (SSMs) are often used in medical imaging analysis, based upon the idea that a statistical shape model represents the normal shape variation of a class of shapes, which is then used as prior knowledge in an algorithm when encountering new data. SSMs describe the shape of an object through the application of PCA to reference points, and work on the assumption that newly encountered shapes are a deformed version of the reference shape. Link to article by Lüthi, Marcel, et al. "Shape modeling using gaussian process morphable models." Statistical Shape and Deformation Analysis. Academic Press, 2017. 165-191.
Supervised Learning	Supervised learning approaches infer a function from labelled training data consisting of a set of training examples. These learnt inferences can then be applied to unseen data and correctly determine class Link to article by Mehryar Mohri, A.R., Ameet Talwalkar The MIT Press ISBN 9780262018258., Foundations of Machine Learning, . The MIT Press 2012.

Support Vector Machines (SVM)	Support Vector Machines analyse and group labelled input data into classes separated by the widest plane (support vector), often in a non-linear relationship where data is not easily distinguishable. The distance of all data points from the separation plane is calculated, and the two closet data points on each side of the plane are used as reference points to classify new data against. It is used in clinical research, for identifying imaging biomarkers, to diagnose cancer or neurological diseases. Link to article by Erickson, B.J., et al., Machine Learning for Medical Imaging. Radiographics, 2017. 37(2): p. 505-515.
Underfitting	Underfitting occurs when a model or algorithm has not learned enough from the training data set, and cannot capture the trend or patterns within a dataset, resulting in low generalization and unreliable predictions. Link to article by Anas Al-Masri on Towards Data Science (2019)
Unsupervised Learning	Unsupervised algorithms process large amounts of data to identify unknown patterns through clustering techniques to identify previously unseen patterns and associations. In contrast to supervised learning approaches, the supplied dataset does not have any labels or known outcomes. Link to article by Hinton, Geoffrey; Sejnowski, Terrence (1999). Unsupervised Learning: Foundations of Neural Computation. MIT Press. ISBN 978-0262581684.