

Written Evidence Submitted by
The AI and Digital Healthcare Group of University Hospitals Birmingham NHS
Foundation Trust/University of Birmingham, UK
(GAI0055)

Context and principles of our response

Our response is focused on the governance of AI in a healthcare context, both AI as a medical device (AIaMD) and the wider group of AI health technologies beyond this group.

Our experience is that of a group of experts in the evaluation and regulation of AI technologies comprising a multidisciplinary group of front-line NHS clinicians and researchers who work with regulators, manufacturers and the NHS to:

- (1) identify and understand the distinct risks related to AI technologies,
- (2) identify the ‘gaps’ in the UK’s current regulatory and quality assurance frameworks where they are not fit-for-purpose for AI technologies, and
- (3) provide potential solutions and tools that can support policy-makers, regulators and the NHS in providing a governance framework that works for AI health technologies.

Our view is that governance of AI in a healthcare context should be designed to ensure that **patients can benefit from AI technologies that are effective, safe and equitable**. This can be achieved through:

- (1) supporting patient safety through *effective* regulation and quality assurance systems *and*
- (2) supporting innovation through *efficient* regulation.

Further information on our work in this area is available at www.regulatoryscience.ai

Q. How effective is current governance of AI in the UK?

What are the current strengths and weaknesses of current arrangements, including for research?

The introduction of AI health technologies into the health system can be thought of in three major categories: those that qualify as a Medical Device (AI as a medical device; AIaMD); those that do not qualify as a medical device but still have a relatively direct impact on patient or public health, including direct-to-consumer health apps and wearables; and those that are related to operational efficiencies in health systems and not directly involved in patient care. As discussed below, almost all governance and regulation of AI is focused on the first group, and this largely sits with the Medicines and Healthcare products Regulatory Agency (MHRA) as part of their oversight of all Medical Devices in the UK. The UK model is aligned to the longstanding system still used by the EU, with the MHRA the UK's 'Competent Authority' for the regulation of medical devices, supported by Approved Bodies (known as Notified Bodies in the EU). The latter are responsible for auditing the AIaMD itself, and the manufacturer's quality management system (QMS).^{1,2} Other relevant regulators both for AIaMD and for other AI technologies used within healthcare include NICE, CQC, professional bodies and the HRA, all of whose remits are discussed later.

Strengths

- 1) The UK has well-established health regulators with clear remits that can address areas relevant to the introduction of AI health technologies ranging from devices to services to research.
- 2) The MHRA's internationally-recognised expertise in medical devices provides a strong foundation for the UK to build on so as to create a regulatory framework that is fit-for-purpose for AIaMD. Such a framework should provide high standards for patient safety and make the UK an attractive place for innovation in this area.
- 3) The MHRA Change Programme for SaMD/AIaMD describes a programme of work that would create an exceptional, sector-leading framework for the UK. In doing this the MHRA have recognised that current Medical Device regulation (including regulation specific to Software as a Medical Device, SaMD) does not adequately address AI-specific characteristics of these devices.
- 4) The MHRA Software Team (incorporating Software as a Medical Device, SaMD, and AIaMD) are a small team but well-respected in the field, with a high level of expertise and a track-record of leveraging expertise from outside the Agency (including academic expertise and exploring regulatory science approaches).
- 5) The international regard for the MHRA's expertise in AIaMD is seen in outputs such as the field-leading guidance documents in partnership with international regulators.³

¹ <https://www.gov.uk/government/publications/medical-devices-uk-approved-bodies/uk-approved-bodies-for-medical-devices>

² <https://www.bsigroup.com/meddev/LocalFiles/en-SG/Services/BSI-md-notified-body-guide-brochure-UK-EN.pdf>

³ <https://www.gov.uk/government/publications/good-machine-learning-practice-for-medical-device->

- 6) The UK is a leader in the building of ethical health data infrastructure, which can provide more inclusive datasets for the training and testing of AI health technologies. It is important that the benefit of data sharing for research and innovation continues to be communicated and supported, and investment continues to be put into building efficient, secure health data systems within the UK. Particular attention should be paid to ensuring that these datasets are diverse and representative⁴ of the wider UK population, and measures continue to be taken to address the potential harms of health data poverty.⁵

Weaknesses/Areas that need strengthening

- 1) The capacity and resourcing of the MHRA is a major barrier to progress. The demands on the MHRA are greater than ever before with the UK's new regulatory independence from the EU and an ever-increasing number and complexity of drugs and devices. Despite this the MHRA - and particularly the Devices division - has seen its funding decrease in relative terms, and many of its international experts leave the Agency. The excellent proposals within the MHRA Change Programme for SaMD/AlaMD would create an exceptional, patient-centred, innovation-supporting framework for the UK, but this will only be realised if the MHRA is adequately resourced to deliver it, and in a timely fashion. Delay in this area is both a risk to patient safety and a barrier to innovation.
- 2) The capacity of UK Approved Bodies is very limited. There is a very significant certification backlog with most manufacturers reporting greater than 12 month delays. The MHRA are reviewing this, and are making efforts to increase capacity with the addition of new Approved Bodies. This is a slow process however, and despite it capacity will still fall short of that available when the UK was in the EU regulatory environment.⁶
- 3) The requirements for regulatory approval for AlaMD are much more demanding in terms of skills and time compared to most other devices. One component of this is that the speed and ease at which AlaMDs can be changed or updated in response to performance issues means that the frequency of certification/approved body audits may also need to be increased, and new processes may need to be introduced (such as Predetermined Change Control Plans - discussed later) to enable more frequent updating.
- 4) Pre-market evidence fo AlaMD is often tested in a virtual environment, either using pre-existing datasets or new data but without the AlaMD being evaluated in the live pathway.^{7,8} This may exaggerate performance, and fail to

[development-guiding-principles](#)

⁴ www.datadiversity.org

⁵ Ibrahim H et al Health Data Poverty: An Assailable Barrier to Equitable Digital Healthcare. Lancet Digital Health 2021. [https://www.thelancet.com/journals/landig/article/PIIS2589-7500\(20\)30317-4/fulltext](https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30317-4/fulltext)

⁶ <https://www.gov.uk/government/news/mhra-appoints-first-new-uk-approved-body-to-certify-medical-devices-since-brexit>

⁷ <https://www.nature.com/articles/s41591-021-01312-x>

⁸ <https://pubmed.ncbi.nlm.nih.gov/33478929/>

detect problems that may occur when the device is used in the real world. Prospective studies within the health pathway should be encouraged as they provide much better evidence to inform the decisions of NHS procurement teams, and implementation planning (with respect to known risks, potential mitigations and resource requirements). Robust evaluation at this stage has the potential to identify performance issues, inform decisions pertaining to classification of AlaMDs, and enable safe, evidence-based translation to patient care.

- 5) AlaMD require greater emphasis on the post-market phase of device monitoring and our current systems are weak in this regard.
 - a) The yellow card reporting system for medicines and medical devices enables interaction between the regulator (MHRA) and front line staff, but was not envisioned for AlaMD.
 - b) There is little understanding of what counts as an AlaMD related adverse event, and at what stage this should be reported. We advocate for error and adverse event training and awareness for all end-users of AlaMDs including both patients and clinicians.
 - c) The MHRA capacity for post-market aspects of medical devices including reporting and compliance is based on traditional static devices, and not resourced to manage the new requirements of AlaMD

Dealing with the post-market phase of AlaMD will require a *change in approach*, an *investment in new skills* and an *increase in resource* to one or more parts of the system.

In *approach*, this will require agreement at the national level as to who is responsible for post-market surveillance and how this is delivered. It will require all stakeholders - health institutions, manufacturers and regulators among others - to build collaborative approaches and safety monitoring infrastructure that enables ongoing oversight of performance and safety with reporting at both local and national level. In *skills*, it will require staff within regulatory bodies and within the front-line NHS to build knowledge and experience in the utilisation and evaluation of AI health technologies. The NHS Transformation Directorate/NHSX 'Buyer's Guide to AI in Health and Care' provides an introduction to some of the questions that should be considered prior to procuring and deploying these technologies. In terms of ongoing evaluation after deployment, the *Medical Algorithmic Audit*, a tool created by NHS and academic experts describes a framework for collaborative safety monitoring at a local level.^{9,10} In *resource*, post-market monitoring will require significant staff-time to assemble and analyse the relevant data to ensure ongoing performance, including effectiveness, safety and equity. We have direct experience of undertaking these 'Medical Algorithmic Audits' in the

⁹ <https://transform.england.nhs.uk/ai-lab/explore-all-resources/adopt-ai/a-buyers-guide-to-ai-in-health-and-care/>

¹⁰ Liu et al. The Medical Algorithmic Audit. Lancet Digital Health. 2022. [https://www.thelancet.com/journals/landig/article/PIIS2589-7500\(22\)00003-6/fulltext](https://www.thelancet.com/journals/landig/article/PIIS2589-7500(22)00003-6/fulltext)

NHS and can attest to both the value of the exercise in assuring ongoing performance but also the demands on staff-time, and new skills required. Just as the airline industry commits significant resource to continually monitoring and maintaining its aircraft, so we need to cost in safety monitoring and quality assurance in the NHS for these complex devices. It is not enough to simply build the plane and to provide a 'flight crew'; we also need to build adequate 'ground crew' that can ensure safety and smooth running within a complex health system.

- 6) The current regulatory framework fails to adequately address the issue of 'biased' device performance in terms of differential performance between groups, and how this may cause harm. This is now rightly under the spotlight, including through the Medical Device Equity Review led by Dame Margaret Whitehead.¹¹ We strongly argue that changes to the regulatory system need to provide assurance that AIaMD are not only 'safe on average' but 'safe for all'.¹²

Q. What measures could make the use of AI more transparent and explainable to the public?

General

As in other sectors, reports of the application of AI in health has been hampered by misinformation, both negative (causing unwarranted fear) and positive (causing exaggerated expectations). There is value in continuing to explore effective ways of communicating:

- *what AI is and what it isn't*: emphasis on pattern recognition, and that this is not magic
- *what AI is good at and what it is bad at*: providing realistic expectations, to help people know when an AI solution is likely to be helpful or unhelpful, and to understand the limits of reliability of AI-generated outputs.

Specific

In the context of AI health technologies, communication approaches should include information in lay language through multiple mechanisms such as websites, databases, instruction leaflets and the media. The language used should have been validated as being meaningful to the public. An excellent model for this would be the work of the organisation Understanding Patient Data which has led the UK's national conversation around how we communicate about health data with the wider public, including the language and explanations used. There should also be visible

¹¹ <https://www.gov.uk/government/groups/equity-in-medical-devices-independent-review>

¹² <https://www.regulatoryscience.ai/safe-effective-and-equitable>

accountability mechanisms including public facing database of AI/ML, with safety incidents logs and audit results.

One area which may be challenging to communicate to the public is the extent to which the user can understand and interrogate the process by which an AI model reaches its output (such as a diagnosis). A key distinction is between models which are interpretable (that is, an expert could understand and explain to a patient how a model works, and how it generated its outcome), and so-called 'black box' models (where no such understanding or explanation is possible). There are pros and cons of both approaches - interpretable models generally 'perform' less well than black box models, but the transparency of their mechanism means that the basis on which it makes a decision is visible with greater opportunity for correction in the case of bias. In contrast a 'black box' model does not allow direct visualisation of the mechanism of the decision-making process. However many other interventions in medicine are relatively 'black box' and we would argue that whilst greater explainability for all health interventions is desirable, it should not be an absolute requirement provided that there is adequate evidence of performance including safety and across the target population.

It should be noted that AI technologies (including in health) are potentially disseminated rapidly and at large-scale, and so any hidden bias may affect a large number of people and cause significant harm. Measures need to be put in place to detect this pre-market,¹³ and robust post-deployment surveillance is needed to uncover subgroup performance differences.¹⁴

Q. How should decisions involving AI be reviewed and scrutinised in both public and private sectors?

Are current options for challenging the use of AI adequate and, if not, how can they be improved?

This is addressed in other parts of our response, but in general this should be achieved by strengthening the existing regulatory framework including supporting and resourcing the MHRA and other relevant regulators to develop and deliver a total product lifecycle approach, and to ensure that there is increased transparency regarding the pre-market and post-market evidence of performance including safety and equity.

Q. How should the use of AI be regulated, and which body or bodies should provide regulatory oversight?

AI in health should be regulated within the existing health regulatory system

The existing medical regulatory framework is robust and flexible:

¹³ <https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme/software-and-ai-as-a-medical-device-change-programme-roadmap>

¹⁴ Liu et al. The Medical Algorithmic Audit. Lancet Digital Health. 2022.
[https://www.thelancet.com/journals/landig/article/PIIS2589-7500\(22\)00003-6/fulltext](https://www.thelancet.com/journals/landig/article/PIIS2589-7500(22)00003-6/fulltext)

- The **Medicines and Healthcare products Regulatory Agency (MHRA)** regulates medical devices (including AI/MD) as well as drugs and other health products
- Hospitals and care providers deploying AI medical devices are accountable to the public via the **Care Quality Commission**
- The **National Institute for health and Care Excellence (NICE)** evaluates evidence and issues guidelines on the conduct of health and social care, including diagnostic and treatment technologies (such as AI medical devices)
- Individual healthcare professionals using AI medical devices are regulated by their **professional bodies**, including the General Medical Council, General Dental Council, Nursing and Midwifery Council, and the Health and Care Professions Council
- **Health Research Association (HRA)** regulate healthcare research including the development or evaluation of AI health technologies.

It is important to note that there is no one body which is currently accountable for health equity. Although all these bodies may point to equity as a characteristic that they would consider in their evaluations, there has been little emphasis on this traditionally, and there is a risk that equity may 'fall through the gaps'. This poses the risk that systemic inequalities may propagate due to a lack of clear responsibility for their detection and mitigation.

Though AI medical devices are novel, existing methods can be used to assess their efficacy and cost-effectiveness (such as randomised trials and health economic analysis). Creation of new bodies to regulate AI medical devices would cause confusion, redundancy of effort and a risk of borderline technologies falling between existing regulators and a hypothetical AI-specific regulator. Medical device manufacturers (including of AI health technologies) are used to working with the existing regulators and within a sector-specific framework which is designed for the needs of healthcare.

The MHRA and other health regulators need to be supported, resourced and required to demonstrate that they are able to effectively regulate AI health technologies

All health regulators need to rapidly evolve to become fit-for-purpose for AI health technologies. The MHRA have demonstrated strong progress in this area (and indeed their Software Team responsible for AI are recognised internationally).¹⁵ The proposed MHRA roadmap shows a deep understanding and commitment to creating an exceptional regulatory framework for AI-enabled medical devices. The MHRA needs to be adequately resourced to deliver on this. This is key to ensuring patient safety and to accelerating innovation and access to AI health technologies. We have worked with the MHRA in the development of the roadmap, to bring our expertise and NHS perspective in how the NHS can assure safety in the post-market phase;¹⁶ and ensuring that issues of bias and equity are addressed throughout.¹⁷

¹⁵ <https://www.gov.uk/government/publications/good-machine-learning-practice-for-medical-device-development-guiding-principles>

¹⁶ Liu et al. The Medical Algorithmic Audit. Lancet Digital Health. 2022.
[https://www.thelancet.com/journals/landig/article/PIIS2589-7500\(22\)00003-6/fulltext](https://www.thelancet.com/journals/landig/article/PIIS2589-7500(22)00003-6/fulltext)

NICE has recently updated their Evidence Standards Framework for Digital Health Technologies to include AI health technologies, and is actively seeking to increase their focus and capacity in digital technologies.¹⁸ The human factors element (including training of users and operational staff), and this is an area that the CQC and the professional regulators (and educators) need to urgently address.

Q. To what extent is the legal framework for the use of AI, especially in making decisions, fit for purpose?

Is more legislation or better guidance required?

For AI health technologies that qualify as a medical device (AIaMD)

These technologies fall within the existing SaMD regulations and are overseen by the MHRA. This regulatory framework needs to be updated to address AI-specific matters, but this has been recognised by the MHRA and is being addressed by the MHRA's Software and AI as a Medical Device Change Programme¹⁹. From our experience of working with regulators (UK and international), the NHS-Transformation Directorate, NHS Trusts deploying these technologies, developers, patients and others, we would note the following key points:

- 1) Any regulation of AIaMD should sit within the existing SaMD framework as part of the regulation of Medical Devices and overseen by the MHRA as the Competent Authority of the UK. We support the view of the medical device sector and the health sector that it is essential that this remains as vertical health sector-specific regulation (which is extended to address AI-specific issues), and is not switched to an overarching horizontal AI-regulation.
- 2) The current SaMD regulatory framework should be updated to include specific provision for the following:
 - a) Measures to detect, mitigate and prevent AI bias.
 - b) Measures that address the risk of 'failure to generalise' including some form of local assurance where a new AIaMD is evaluated in a local setting and population prior to full deployment.
 - c) Measures that address risk around 'human factors' in which an AIaMD may be more sensitive to accidental/deliberate misuse than other devices, and with greater risk of harm. There is also a need to delineate responsibility between the manufacturer and the user.
 - d) Measures that improve explainability of the AIaMD to facilitate patients and end-users to understand how it has reached a decision, check its reliability and challenge it. We disagree with the idea that **only** explainable AI models should be used, but rather would encourage manufacturers to develop the model that will best address the health need, whilst also maximising the extent to which that model and the basis of its decisions can be understood.

¹⁷ www.data.diversity.org; <https://www.nature.com/articles/s41591-022-01987-w>

¹⁸ <https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies>

¹⁹ <https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme/software-and-ai-as-a-medical-device-change-programme-roadmap>

- e) Measures that provide ongoing assessment of performance in the post-market phase. This needs to include detection of 'drift' of AIaMD performance related to any change including in local setting (e.g. software upgrades of scanners which provide input data) or local population. Tools such as the Medical Algorithmic Audit which has been cited by the MHRA may be valuable in this context.²⁰
- f) Measures that allow for frequent (and potentially continuous) updating of models, specifically the Pre-determined Change Control Plan (PCCP).
- g) Measures that support the training of the workforce to ensure that the professionals are competent to use the devices safely, and to evaluate their performance post-deployment.

The MHRA Roadmap recognises these factors, therefore ongoing Government support and resourcing of this initiative will help advance patient safety as well as an innovative UK health technology sector.²¹

A detailed analysis of this area is expected shortly in the form of the UK's Regulatory Horizons Council (RHC) Report into Regulation of AIaMD (due to be published on gov.uk).²² This includes a multisector gap analysis of AIaMD regulation produced through consultation with patients, health professionals, health tech companies, digital transformation experts, computer scientists, and regulators. As well as providing contemporary assessment of the UK's position in regulation of AIaMD comparisons will also be drawn with the USA, Europe and other leading digital economies, explicitly highlighting evidence of where UK's regulatory framework is not currently fit-for-purpose, and providing specific recommendations for regulatory reform and other measures to address this.

For AI health technologies for consumer use that do not qualify as a medical device
In contrast to the active efforts by the MHRA and others to create an efficient and effective regulatory framework for AIaMD, there is a regulatory void when it comes to health technologies that do not qualify as a medical device. Such products are sold direct-to-consumer, and include many healthcare apps and wearables (such as smartwatches). There is currently no health-related regulatory framework relating to these, although they will be subject to standard data protection and privacy law and consumer protection measures. Platforms retailing such applications have responsibility for some degree of due diligence however this is often left to their own discretion, and in our view does not provide patients with sufficient protection from harm, particularly as these applications may claim to diagnose or manage potentially life threatening conditions such as cardiac arrhythmias, cancer or serious mental illness. This is a major area of concern and should be addressed.

With regard to the specific example of mental health apps, the Wellcome Trust has recently announced support to MHRA and NICE to address regulation in the context

²⁰ Liu et al. The Medical Algorithmic Audit. Lancet Digital Health. 2022.

[https://www.thelancet.com/journals/landig/article/PIIS2589-7500\(22\)00003-6/fulltext](https://www.thelancet.com/journals/landig/article/PIIS2589-7500(22)00003-6/fulltext)

²¹ <https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme/software-and-ai-as-a-medical-device-change-programme-roadmap>

²² <https://www.gov.uk/government/groups/regulatory-horizons-council-rhc>

of mental health apps,²³ but similar attention needs to go into the regulation of apps outside of mental health. We recognise that not all of these apps use AI, but most of the more sophisticated tools do use some form of AI, and indeed this is often used as a 'selling point' of such tools to infer higher levels of sophistication and implied performance.

For AI health technologies for health system support including operational use

These fall outside of the remit of MHRA, may fall indirectly under the auspices of the CQC in their role assessing the quality of services being delivered by health and care institutions. Many such systems may be valuable for improving operational efficiency but health and care institutions (and regulators) should be alert to such systems indirectly affecting patient care and causing discrimination and harm. For example, an AI system to support booking of outpatient attendances may use rules or learn patterns of behaviour that would discriminate against certain groups who are at risk of lower attendance in order to optimise 'efficiency' of out-patient capacity. When such technologies are being introduced, consideration must be given to unintended consequences regarding equity of access (and indeed whether they can be used to improve access for those who are currently underserved or experience other barriers to access).

Q. What lessons, if any, can the UK learn from other countries on AI governance?

The MHRA is well-regarded internationally, and its Software Group has been an influential contributor to efforts from the IMDRF and leading jurisdictions to address the challenges of AI/ML. The MHRA Software Group are currently leading a programme of regulatory change specific to software as a medical device (SaMD) including AI as a medical device (AI/ML).²⁴ This has included producing the first Good Machine Learning Practice (GMLP) principles, a collaboration with the U.S. Food and Drug Administration (FDA) and Health Canada.²⁵

One area where the UK system should learn from other international leaders (such as the FDA and TGA) is in error and adverse event reporting from AI/ML. Currently in the UK, public and health professionals' access to safety reports on products is limited to online reports of Effective Field Safety Notices (EFSNs) and Field Safety Corrective Actions (FSCAs). Although these are available to the public, there is limited information regarding the medical device of interest, and there is no way of filtering for reports by medical device or manufacturer. In contrast, the U.S. FDA and the Australian Therapeutic Goods Administration (TGA) both have adverse event databases, MAUDE and DAEN respectively.^{26, 27, 28} Both databases can be searched

²³ <https://www.gov.uk/government/news/mental-health-funding-of-18m-welcomed-by-mhra-and-nice-to-explore-regulation-of-digital-mental-health-tools>

²⁴ <https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme/software-and-ai-as-a-medical-device-change-programme-roadmap>

²⁵ <https://www.gov.uk/government/publications/good-machine-learning-practice-for-medical-device-development-guiding-principles>

²⁶ <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>

²⁷ <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/search.cfm>

²⁸ <https://www.tga.gov.au/safety/safety/safety-monitoring-daen-database-adverse-event->

to identify reported adverse events, allowing for both in depth due diligence and post-market safety monitoring. The FDA database goes even further in promoting transparency. Firstly, the FDA has curated a list of AI/Machine Learning (ML) enabled medical devices. This allows for easy identification of AIaMDs, and understanding of what types of medical devices are considered to be “AI/ML enabled”. Secondly, and perhaps more importantly, device summaries are available on the FDA database for all medical devices that have passed regulator approval. The summaries include varying levels of evidence depending on the regulatory pathway that the AIaMD has been approved through. Evidence includes a clinical evaluation summary, risks associated with the AIaMD, a risk benefit analysis summary, and associated warnings.

In the UK, we recommend the introduction of a publicly accessible and searchable adverse event database, curated device lists for all approved SaMD and AIaMD, and summary documents containing key evidence submitted by device manufacturers during the regulatory approval process.

(November 2022)