**phg**foundation

making science work for health

# Inquiry on algorithms in decision making: evidence from the PHG Foundation

The PHG Foundation is an independent, not-for-profit health policy think-tank that aims to make science work for health, with twenty years' experience in issues surrounding the responsible and effective use of genomics within health services.

## Summary

• Algorithms are becoming ubiquitous in healthcare – their role in research, diagnosis, prognosis, and monitoring increases. Algorithms now answer complex questions and solve sophisticated problems; as a consequence, these algorithms themselves are increasingly complex and sophisticated.

• Algorithm performance depends on data inputs: quality algorithms require quality data. It is therefore imperative that data flows are integrated and streamlined.

• Algorithms represent unique challenges in regards to transparency. Emphasis on technical and non-technical solutions early in algorithm development is necessary to ensure algorithms remain transparent and their developers sufficiently accountable.

• Public perception and trust are an often-ignored yet important consideration. Public engagement and consideration is imperative to ensure access to health data necessary for the development of healthcare algorithms.

• The Information Commissioner's Office (ICO) has made important steps in providing data controllers with advice as to how to comply with the General Data Protection Regulation (GDPR) transparency requirements. More specific advice is highly desirable to foster good practice, which will require that the ICO should engage with the algorithm development sector.

• The In Vitro Diagnostic Devices Regulation ( IVDR) is another source of potential regulation, which may be a means of ensuring greater transparency. The Medicine and Healthcare products Regulation Agency have responsibility for regulation in this area, and are taking forward IVDR implementation.

phgfoundation

making science work for health

It is clear that algorithms are and will continue to be ubiquitous in healthcare. Further, it is also likely these algorithms will become increasing complex, drawing from a range of rich data sources.

## Extent of current and future uses of algorithms in healthcare

Algorithms are widely used across healthcare, both to guide decision-making, and as an integrated component of many investigative tests, including imaging.

Algorithms have typically supplemented or complemented aspects of clinical care but have not replaced human intervention. Increasingly however, algorithms are being used to replace humans particularly where workforce capacity or expertise is lacking.

The complexity of algorithms used in healthcare has also increased:  once rare and relatively simple, there is now an increase in both number and complexity.

Regarding research itself, text mining algorithms are already at work, assisting researchers in compiling systematic reviews. University College London's EPPI-Centre is one body using these techniques to publish on a variety of topics. Moreover, groups such as Cochrane continue to review machine learning and its increasing usefulness in producing systematic reviews.

Computer-aided diagnosis and prognosis continues to advance. Histopathology is one area likely to benefit from machine learning in the near future. In terms of diagnosis, image analysis algorithms are already set to act as a check to misdiagnosis and reduce pathologist workload.  In terms of prognosis, algorithmic classification is already a part of the process, helping to predict patient outcomes.  Hence, while algorithms may not fully automate processes in the near future, they will and already do complement existing diagnosis and prognosis procedures and processes.

Moreover, more speculative projects such as the SPHERE house envision a future where monitoring technology produces data that, when analysed by algorithms, may assist with early detection and management of chronic diseases or threats to longer term health: sedentary behaviour, depression, anxiety, and the detection of falls.

The examples cited above represent only a small sample of what algorithms do and might do for health. Nevertheless, it is clear that algorithms are and will continue to be ubiquitous in healthcare. Further, it is also likely these algorithms will become increasing complex, drawing from a range of rich data sources.

## Whether good practice can be identified and spread

### Improving algorithm quality through optimising data sharing

The functionality and accuracy of algorithms are heavily dependent upon having access to relevant data, for design, building, testing and verification.

Genomic sequencing is a technology which allows massively parallel analysis of small fragments of DNA sequence, comparison with a reference genome, identification of variants, and reporting of these as being disease causing (pathogenic), or benign in relation to that reference genome.

Accurate, robust disease diagnosis and reporting requires access to relevant data sets. These datasets must contain genotype data, but also relevant phenotype data (i.e. details of a patients clinical condition, signs and symptoms). Both these types of data are needed for accurate interpretation.

Because rare disease causing variants may cluster in families or in particular ethnic groups, laboratory scientists will often need to access data from outside an NHS Trust or even outside the UK. See Data sharing to support UK clinical genetics and genomics services.

The 100,000 Genomes Project is currently recruiting NHS patients with rare diseases and cancer for genome sequencing. This project is in the process of developing increasingly automated systems to sequence, analyse and interpret sequence data, and Genomics England is working closely with NHS England with the ultimate aim of translating and integrating this into NHS clinical services.

Other initiatives such as the Deciphering Developmental Disorders project has pioneered the use of the DECIPHER browser and algorithms to improve the rates of diagnosis amongst children with developmental delay (for example in creating composite images of children who share the same inherited syndrome, without compromising the identities of the children concerned).

In order to deliver these technologies more effectively, NHS clinical genetics laboratories in England are in the process of re-procurement, in order to centralise genomic sequencing services and integrate data flows to streamline work processes.

### Additional ways of promoting good practice

Algorithms, especially those utilising machine learning, represent particular challenges when it comes to transparency.

Algorithmic opacity consists of two general, related problems:

- Frequently, it is unclear exactly what factors lead an algorithm to reach a particular decision.

- Even if the decision process is clear, the decision may be the result of complex interactions, and therefore be hard to communicate in a clear, easily-intelligible manner.

Given the complex nature of some algorithms, alternative strategies that promote accountability through transparency are required.

The Information Commissioner's Office (ICO) in their Big data, artificial intelligence, machine learning and data protection propose three general methods that might encourage algorithmic transparency:

- The use of algorithmic auditing. Processes that identify and monitor factors that lead to algorithmic outcomes should be part of algorithm development and present from the algorithm's inception. Moreover, specialised auditing algorithms and companies that specialise in algorithm transparency might further assist in securing transparency.

> Given the complex nature of some algorithms, alternative strategies that promote accountability through transparency are required.

Following ICO's suggestions, it is clear that emphasis on technical and non-technical solutions early in algorithm development is necessary to ensure algorithms are transparent, and their developers are accountable.

- A combination of technical solutions and better use of visualisation and interactivity may assist with communicating otherwise opaque outcomes. That is, automated explanation in the form of Natural Language Generation might constitute a part of how algorithms communicate their decision-making process. Moreover, visualisation techniques like interrelation charts and Venn diagrams are ready solutions to otherwise impenetrable prose. In short, insight into how best to communicate algorithmic decision-making is essential to ensure that data subjects are informed.

- Ethics boards also offer a non-technical method to interrogate algorithmic decision-making. It is submitted that ethics boards are best utilised early in algorithm development and include both computer scientists and lay members. In this way, ethics boards should provoke developers to consider and communicate clearly how their algorithm makes decisions.

Following ICO's suggestions, it is clear that emphasis on technical and non-technical solutions early in algorithm development is necessary to ensure algorithms are transparent, and their developers are accountable.

It is important to foster a regulatory environment that provides appropriate incentives for the development of algorithms by commercial developers as well as the public sector. Maintaining a proportionate regulatory environment which recognises the need for developers to have, and protect proprietary interests over the algorithms they have created is an element that needs to be addressed. The balance between transparency and the appropriate protection of commercial interests may sometimes be difficult to reach.

### Promoting trustworthiness through developing good practice

An important element that is not fully reflected in questions framing the enquiry is the extent to which publics (data subjects and participants) regard governance, systems and processes as being trustworthy.

Although it did not explicitly explore attitudes to algorithms, software and automation, the report The One-Way Mirror: Public attitudes to commercial access to health data explored public attitudes to various data uses.  Many of these findings are pertinent to this enquiry: two of the most salient findings are that:There is a poor level of detailed awareness about how the NHS uses health data and little understanding of safeguarding practices

- A hierarchy of key tests determined public acceptability, each element having to be satisfied before the next is perceived as relevant: why [does the activity's outcome have a provable and sufficient public benefit]; who [can the organisation be trusted to have public interest at heart?]; what [how anonymised or aggregated is the data?] and how? [Does the safeguarding, access and storage protocol provide reassurance that the data will be safe?]

We suggest that these criteria are likely to be more widely applicable to public attitudes towards algorithms and software.

The Nuffield Council on Bioethics' report '[The collection, linking and use of data in biomedical research and healthcare](#)' notes that the opacity and complexity of algorithms makes them difficult to regulate and highlights the tension between the need for public scrutiny and protection of commercial secrecy through trade secrets.

Noting that algorithm performance declines when applied to more unusual cases, this report highlights the potential for algorithms to systematically compound and reinforce social inequalities or make options of interest invisible. Their recommendations include securing commitments from data controllers regarding data access, transparency and accountability

## Methods for providing regulatory oversight of algorithmic decision making

### General Data Protection Regulation (GDPR)

GDPR Art 12(1), 12(5), 12(7), 13 and 14 all work to provide data subjects with a right to be informed. These articles and their accompanying recitals (58-62) emphasise the need for transparency in meeting this obligation.

However, transparency in algorithmic decision-making (as outlined in paras 25-36) is often no easy task.

Article 22(b) also provides for greater safeguards where decisions about the data subject are based **solely** on automated processing, including profiling 'which produces legal effects' or similarly significant decisions concerning him or her.

Clause 13(2) of the Data Protection Bill clarifies the meaning of 'significant decision.' In its current form, the Bill clarifies that a decision will be 'significant' if the decision either produces legal effects concerning the data subject, or significantly affects the data subject. It ought to be noted that this definition of 'significant' is potentially very wide. Given this, the prohibition on taking significant decisions based solely on automated processing is likely to be broad, and the exception to this prohibition is likely to be narrow.

[Guidance](#) from the Article 29 Working Party on automated individual decision-making and profiling also suggests that the effects of processing '*must be more than trivial*' with '*potential to significantly influence the circumstances, behaviour or choices of the individuals concerned*'.

The [GDPR](#) includes provision for supervisory authorities to publish codes of conduct '*intended to contribute to the proper application*' of the Regulation.

The draft [Data Protection Bill](#) clarifies that the Information Commissioner's Office has powers to prepare appropriate codes of good practice and we suggest that guidance could be developed to clarify what constitutes best practice when processing data, including specific guidance for transparency/accountability.

In this regard, ICO's *Big data, artificial intelligence, machine learning and data protection* report represents a positive step forward. This report offers advice for how algorithms and their developers might secure transparency. Yet, this guidance is given in broad strokes and much of it is speculative in tone.

> The Information Commissioner's Office has powers to prepare appropriate codes of good practice and we suggest that guidance could be developed to clarify what constitutes best practice when processing data, including specific guidance for transparency/accountability.

This constitutes a more onerous requirement for algorithm developers than has previously existed in law. The extent to which these criteria are relevant for post-market surveillance warrants further discussion.

Bodies like ICO must forge partnerships with those bodies offering technical and non-technical solutions to algorithmic transparency.  In short, 'ought' implies 'can'. If the GDPR's right to be informed is to be practicable in the UK, ICO must bolster their current advice with specifics. Guidance distinguishing between different kinds of algorithms and the methods appropriate to secure transparency will provide clarity and a degree of legal certainty.

The National Information Board  is also developing work streams which include aspects of app governance and regulation. In particular, the NIB could incorporate requirements for transparency as part of a number of work streams (e.g.  for accreditation of apps, in work stream 1.2).

### Alternative sources of governance/regulation

Another source of regulatory oversight of algorithms might be through the forthcoming *EU In Vitro Diagnostics Devices Regulation 2017/746* . This Regulation was passed on 5 April 2017 and relevant sections relating to the regulation of software and algorithms will (subject to Brexit discussions) be implemented in May 2022. The Medicines and Healthcare Regulatory Products Agency has responsibility for implementing this Regulation within the UK.

This Regulation includes certain types of software/algorithms within its scope for the first time.

In order to bring a device (including free-standing software/algorithms) to market, requirements for analytical and clinical performance need to be satisfied (Annex I paragraph 9).

Devices (including software) are required to be designed and manufactured in such a way that they are suitable for the purpose of detection, predisposition, treatment or management of diseases requiring elements of performance to be tested in both normal and affected populations.

This may also mean that certain types of risk stratification tool which take account of multiple risk factors to calculate a risk score for an individual (for example of developing certain cancers over a fixed period) would be governed by this Regulation if genotype was included as one of the variables.

Annex II of the Regulation sets out the specifications which need to be provided: these include a description of the data interpretation methodology (Annex II 3(1)(e)); and evidence of the validation of the software including "*summary results of all verification, validation and testing performed in-house and applicable to an actual user environment prior to final release*" (Annex II paragraph 6(4)). This constitutes a more onerous requirement for algorithm developers than has previously existed in law. The extent to which these criteria are relevant for post-market surveillance warrants further discussion.