phg foundation
making science work for health

Regulating algorithms in healthcare

The GDPR and IVDR in practice

UNIVERSITY OF CAMBRIDGE

Authors

Alison Hall and Johan Ordish

Acknowledgements

The PHG Foundation is grateful to the expert contributions provided by all of the Workshop delegates. A list of delegates can be found in the Appendix. Of special mention are guest speakers Brent Mittelstadt and Sandra Wachter of the Oxford Internet Institute and Sara Payne, PHG Foundation Associate. We are also grateful for the assistance of Hannah Murfet of Microsoft Research who reviewed the draft manuscript. The PHG Foundation notes the assistance of members of the Centre for Law, Medicine and Life Sciences, University of Cambridge: Kathleen Liddell, Jeffrey Skopek, David Erdos, and Mateo Aboy.

URLs in this report were correct as of October 2019

This report can be downloaded from:

www.phgfoundation.org

Published by PHG Foundation
2 Worts Causeway
Cambridge
CB1 8RN
UK

+44 (0)1223 761900

May 2019

Correspondence to:
intelligence@phgfoundation.org

How to reference this report:
Regulating algorithms in healthcare: the GDPR and IVDR in practice

PHG Foundation (2019)

ISBN 978-1-907198-33-5

Contents

Executive summary

This report summarises the findings of *Regulating algorithms in healthcare: the GDPR and IVDR in practice*. The workshop brought together those with expertise in data protection and medical device law. These two fields face similar problems and may learn from each other.

The workshop generated the following key findings:

- Different regulatory bodies speak in different and often contradictory terms, using (and sometimes not defining) such terms as 'algorithms', 'software', 'computer programs' and 'automated processing'
- As algorithms increase in complexity and ubiquity, there is growing interest in the ethical challenges associated with their use, and particularly in the novel ethical challenges that may arise
- Determining what constitutes 'meaningful transparency' in a given context is a key priority for the field
- There is disagreement over whether there is in fact a right to explanation for individual decisions under the GDPR and what kind of explanation might satisfy this right if it does exist
- Counterfactual explanations may be one way of satisfying the right to explanation. However, they have limitations and are unlikely to be helpful when making judgments about the fairness of a system
- Many algorithms used in healthcare will be regulated as medical devices and will be subject to the new MDR or IVDR. It is therefore important to understand how these Regulations differ from the previous three Directives
- Arguably, EU and US methods to determine whether a device is regulated as a medical device are converging, both being sensitive to the function and risk a device poses
- Some machine learning devices may pose novel problems for medical device regulation, because they may be human uninterpretable or may retrain and so represent a moving target for regulators

The synergies and findings of this and a subsequent workshop informed another PHG Foundation report (Algorithms as medical devices) and a Wellcome Trust funded PHG Foundation project – Black Box Medicine and Transparency – which take forward some of the topics and challenges noted in this report.

Introduction

Healthcare is undergoing profound change. Like many sectors, the new face of the healthcare industry is exemplified by expansion into digital health, and increased demand and uptake for algorithms. All those within the sector need to respond to a changing industry that brings new challenges and opportunities with it.

Recognising this dynamic environment the PHG Foundation convened a multidisciplinary workshop of academics, developers, researchers, legal practitioners, health professionals, regulators and policymakers to explore the issues that regulating algorithms in healthcare might raise.

Chief amongst our motivations for the workshop was the view that there are two approaches to algorithm regulation that are traditionally siloed: those focusing on algorithms as data consider requirements for data processing under the General Data Protection Regulation (GDPR); those focusing on algorithms as medical devices look to the Regulations on Medical Devices (MDR) and In Vitro Diagnostic Medical Devices (IVDR). However, both approaches face similar regulatory and ethical challenges.

The aim of this workshop was to evaluate the impact of the new EU Regulations on data protection and medical devices. Discussions included: the law, how regulation might impact upon clinical practice and discussion of quality standards.

In particular, this workshop considered:

- Conceptual issues: defining algorithms and software in law
- Ethical issues: the ethics of using algorithms in healthcare
- GDPR and data: the right to explanation and 'counterfactual explanations'
- Medical device law: IVDR, device qualification and machine learning as a medical device

These Regulations will continue to be relevant in the short term irrespective of the outcome of Brexit because of pledges to adopt them in some form into UK law.

Where indicated, sections of the report represent the guest speaker's opinion rather than that of the PHG Foundation.

This workshop identified some synergies between the two approaches and also identified specific areas meriting further research. These included the need for more normative discussion about how much explanation is desirable; the feasibility of complex explanations and the need for worked examples of counterfactual explanations in health.

These findings have led to two further projects – *Algorithms as medical devices* report (also available) and the Wellcome Trust funded *Black box medicine & transparency*.

1 Policy context

Alison Hall, PHG Foundation

Algorithms and software are ubiquitous in healthcare and are fast becoming a core component of healthcare planning and delivery. Recognising this, the PHG Foundation brought together algorithm and software developers, academics, researchers, policy makers and regulators to discuss the impact of three EU Regulations which promise to transform the UK regulatory framework for apps and algorithms. These Regulations: the General Data Protection Regulation (GDPR) (which concerns data processing), the Medical Devices Regulation (MDR) and In Vitro Diagnostic Medical Devices Regulation (IVDR) , consider different aspects of algorithms and software[1-3]. The GDPR regards algorithms and software as data processing tools. Therefore obligations on algorithm developers must mirror the obligations in the GDPR. They are framed in terms of information provision, transparency and explanation. By contrast, the MDR and the IVDR instead regard algorithms and software as medical devices and regulate device performance and validation.

This report describes a workshop held on 20 March 2018, at which, these two different regulatory approaches were evaluated and a number of challenges for software developers were highlighted. The report ends by identifying areas that warrant additional research. The audience for this report spans developers, researchers, health professionals, regulators and policy makers – in short, everybody who is interested in how software and algorithms will be adopted and used for health related purposes in the UK.

1.1 The regulatory context

This workshop was convened at a time of profound change:

- There has been an exponential rise in use of digital technologies. Global use of mobile phones has enabled a proliferation of mobile apps and algorithms many of which have health related uses
- There is increasing opportunity for citizens to take control over their own health through using digital devices and actively seeking out other interventions. Thus the boundary between health and wellness is becoming increasingly blurred as individuals are urged to take proactive steps to keep themselves well and detect disease through personalised prevention
- Against this backdrop, regulators are struggling to clarify what lies within the scope of their regulatory remit. This is partly due to a lack of clarity in terminology
- Regulatory authorities such as the Food and Drug Administration (US FDA) and the Medicines and Healthcare products Regulatory Agency (UK MHRA) are developing different strategies to cope with their expanding remit. For example, the US FDA retains regulatory oversight over wellness products but has discretion to intervene for those that are deemed to pose a risk to users: in contrast the MHRA excludes wellness products from its regulatory oversight

The challenge for the sector is to maintain a proportionate, yet responsive regulatory system. This would prevent unsafe products reaching the market but allow safe clinically effective products to be used in a timely way.

2 Defining algorithms and software

Johan Ordish, PHG Foundation

How we regulate digital health depends upon how we define the term. However, the terms used in EU and UK to describe digital health's subject matter are inconsistent and changing, with different bodies of regulation speaking in different terms including 'algorithms', 'software', 'computer programs' and 'automated processing' (see Table 1). Even where there is consistency on what term to use, often there is little agreement on what that term means.

This section explores why such inconsistency might exist and aims to:

- Explain how 'software' and 'algorithms' are defined in the MDR, IVDR and GDPR
- Explain why these concepts are difficult to define both conceptually and in law
- Demonstrate instances where ambiguity has led to legal uncertainty or absurdity

Table 1: Illustration of different terminology

Term	Definition	Citation
Software	'Software is defined as a set of instructions that processes input data and creates output data.'	EU MEDDEV 2.1/6 Qualification and Classification of Stand Alone Software Used in Healthcare Within the Regulatory Framework of Medical Devices
Computer program	'Computer program is a syntactic unit that conforms to the rules of a particular programming language and that is composed of declarations and statements or instructions needed to solve a certain function, task, or problem.'	ISO/IEC 2382:2015 Information technology - Vocabulary
Automated processing	'Solely automated decision-making is the ability to make decisions by technological means without human involvement.'	Working Party 29, Guidelines of Automated individual decision-making and Profiling for the purposes of Regulation 2016/679

2.1 Challenges of defining algorithms as a conceptual issue

The conceptual definition of 'algorithm' is contentious among computer scientists. Definitions of algorithm differ widely, from the very broad: 'a set of steps to accomplish a task'[4] to definitions with specific set of criteria, for example, US computer scientist Knuth's definition as follows[5]:

- Definiteness: algorithms require precise description, specifying each step and how this step leads to the desired solution
- Inputs: an algorithm generally takes some value or values as inputs
- Outputs: an algorithm generally produces some output value or values from an initial set of defined inputs and specifies how this output value is derived
- Finiteness: algorithms must have a finite series of steps and terminate upon completion of these steps. If a procedure has all the characteristics of an algorithm but is not finite, then it is a computational process, not an algorithm

One of the foremost disagreements regarding the definition of algorithm is between those that conceptualise algorithms as recursors and those that conceptualise algorithms as abstract state machines. While the exact difference between the two accounts is difficult to articulate, the abstract state machine account will tend to consider the actual coded implementation of any algorithm as 'the algorithm.' In contrast, the algorithms as recursors account will tend to consider any algorithm written in code as merely an implementation of the 'actual algorithm'. Such uncertainty underlies previous legal controversies described below.

2.2 Algorithms in law

The definition of algorithm has consequences in law. Patent law is one example where the terminological ambiguity of algorithm has resulted in legal uncertainty.

For example, consider the mathematical representation of Euclid's algorithm to find the greatest common divisor:

$$< r_{k-2} = q_k r_{k-1} + r_k >.$$

One way to represent this mathematical algorithm in a code is in the python programming language as below:

```
1    a,b = 252,105
2    while b:
3    a,b = b, a%b
4
5    print (a)
```

The question is, are there two algorithms, one mathematical and one in Python, or one algorithm with its mathematical and Python representations?

Such an issue arose in the 1972 US case of *Gottschalk v Benson*. This case regarded a patent application for a method of converting binary numbers into digital format. In that US case, Justice Douglas noted that granting a patent for this subject matter would 'wholly preempt the mathematical formula and in practice the effect would be a patent on the algorithm itself,' tacitly assuming that there was one algorithm with multiple representations[6]. In this way, Justice Douglas equated 'algorithm' with an abstract idea, and tacitly rejected definitions such as the 'abstract state machine' definition above.

The legal position has since moved on with subsequent case law but the point remains, the law often applies implicit definitions of concepts, leading to legal uncertainty[7].

2.3 Software as a conceptual issue

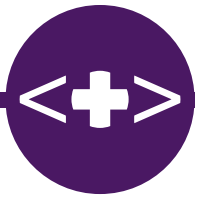
The definition of software is also subject to controversy. Software is typically defined in relation to hardware: something can be either software or hardware but never both. The standard distinction between the two being that software is seen as abstract and hardware is seen as physical. However, this distinction may be untenable and may not get at anything significant. For instance, philosopher James H. Moor notes that many programs that we think of as obvious cases of software may also be physical devices made of lever pulls and pushes.[8] This conceptual uncertainty has consequences for law that relies on a definition of software.

2.4 Software in law

Consumer protection law is an area of law in which the correct definition of software is a source of legal controversy. This controversy concerns whether or not software counts as a product under the UK Consumer Protection Act 1987. Relevant definitions are found below.

Table 2: Definitions of 'product' and 'goods' under the Consumer Protection Act 1987 and Directive 85/374.

Term	Definition	Citation
Product	'any goods or electricity and includes a product which is comprised in another product, whether by virtue of being a component part or raw material or otherwise.'	Section 1(2) Consumer Protection Act 1987
Goods	'"goods" includes substances…'	Section 45(1) Consumer Protection Act 1987
Product	'"product" means all moveables… 'Product' includes electricity.'	Article 2 Directive 85/374/EEC (EU parent Directive to the UK Consumer Protection Act 1987)



Consider the following series of examples that demonstrate that the lines between software versus hardware and product versus non-product are blurred[8].

Defective map: a manufacturer produces a topographic map for air navigation. Carelessly, the manufacturer leaves a mountain off the map, causing a plane to crash.

Fault in rudder: a manufacturer produces rudders for aircraft. Carelessly, the manufacturer produces a batch of rudders with a fault, causing a plane to crash.

Navigational software: a manufacturer designs navigational software for aircraft. Carelessly, a software fault that causes the height of mountains to be incorrectly registered is left in, causing a plane to crash.

Navigator: while the navigator has a correct map, he verbally tells the pilot an incorrect height, causing the plane to crash.

In each of the above scenarios (apart from the last), a 'product' could arguably have caused the crash and it is unclear if there is a principled basis on which to distinguish between each of the scenarios. The difficulty seems to hark back to conceptual controversies that surround software. That is, currently there is no good way to distinguish systems of information from software in particular.

Discussion

- What progress can be made with such definitional uncertainty?
- Given that these definitions are inherently contentious should we seek further legal definitions of these concepts?

Delegate response

- Is there virtue in compromising between definitions that are either too specific or too vague? Definitions can fail by being too wide or too narrow.
- Some of the definitions of 'algorithm' refer to a particular field – we have to understand how different sectors use the term
- In regard to pure mathematics and abstract ideas, no one should have a monopoly over these concepts
- Perhaps the problem is less a matter of whether there should or should not be definitional uncertainty but what instrument is best suited to providing such clarification, i.e. ISO standards might be most appropriate?
- Perhaps there should be different definitions in different parts of the law serving different purposes?

3 The ethics of algorithms

As algorithms increase in complexity and ubiquity, there is growing interest in the ethical challenges associated with their use, and particularly with any novel ethical challenges. The workshop focused on these issues in the context of healthcare.

3.1 Mapping the debate

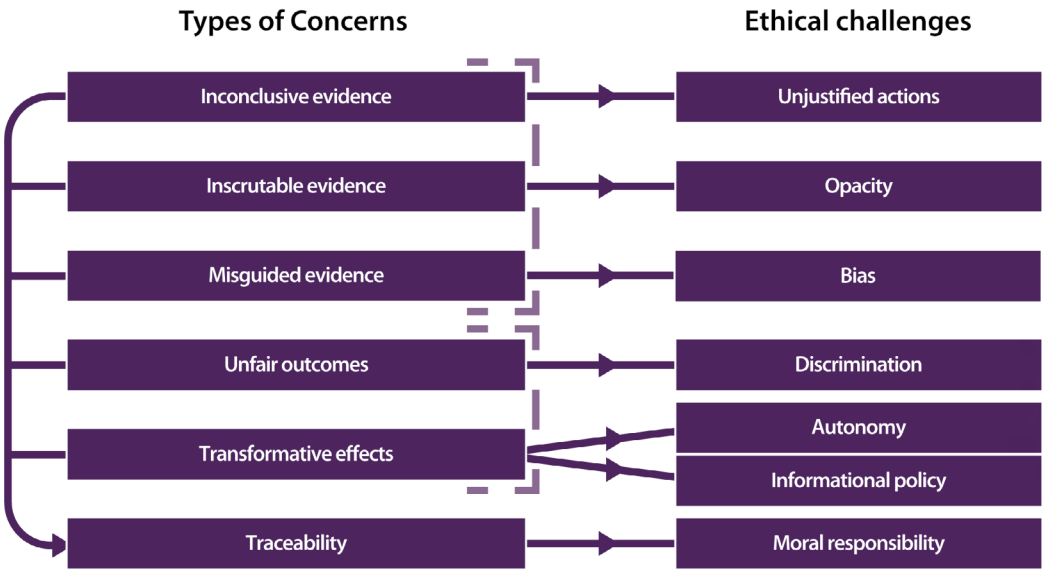
Brent Mittelstadt, Oxford Internet Institute; Alan Turing Institute

Media reports highlighting ethical concerns about the use of algorithms are becoming increasingly common. Standout stories include using software to assign a risk score to defendants to recommend parole terms (e.g. Correctional Offender Management Profiling for Alternative Sanctions (COMPAS)) or to judge teaching performance[10]. Typically such stories concern bias and lack of transparency. These issues arise for a variety of reasons: lack of foresight in design and implementation of algorithms and/or lack of oversight in their use. The use of algorithms in increasingly complex tasks, for example through use of robots in surgical procedures, raises questions about how to distribute responsibility for harms that may result. Harms include physical risk to people as well as risks to privacy.

A review of the academic literature was carried out in Mittelstadt (2016) to assess the ethics of algorithms using the philosopher Robin K. Hill's (2015) definition of an algorithm: 'Mathematical construct with a finite, abstract, effective, compound, control structure, imperatively given, accomplishing a given purpose under given provisions'[11].

This generated a typology of different ethical challenges[12].

Figure 1. Six types of ethical concerns raised by algorithms (adapted from Mittelstadt slide set)



A number of challenges arise from the use of algorithms. There are three dimensions of opacity: interpretability (whether a rational decision making process is meaningful)[13]; accessibility (influenced by trade secrets, willingness and technical barriers) and data provenance (the means of data collection; limitations, biases and gaps). These elements contribute to the feasibility of offering a meaningful explanation, but questions about what constitutes a desirable explanation in terms of scope still need to be resolved.

In situations where an explanation is not possible, other methods can be used to keep systems accountable e.g. ethical certification (ex-post (after the fact)) detection of types, severity and prevalence of the effects of a decision) and auditing.

The extent to which a system is traceable will guide the distribution of moral and legal responsibility across individuals and teams in assigning who should be held responsible for the effects of algorithms. This will help to prevent unjustified actions, as determining when a correlation is sufficiently reliable to be acted on is sometimes challenging. Existing norms of good practice and context may be lost through automation bias (or de-responsibilisation), which may result in people being reduced to their data. It is important that decision making is transparent, traceable and arguably open to public scrutiny and debate. Challenges include the fact that values may be embedded in decision criteria used to guide decision making.

Other ethical challenges include determining who should make ethical decision making rules, what principles should be followed and what data/audit logs are sufficient to apportion blame. Even deciding what constitutes fairness can be problematic. For example, Kleinberg *et al* suggest that three fairness conditions (calibration within groups or balancing for the positive or negative classes) are incompatible with each other, which implies that algorithms cannot be universally fair[14].

Privacy theory and law traditionally protect (identifiable) individuals. Current protection mechanisms include removal, hiding identity in datasets, relying on consent or user agreements, and providing individual oversight of personal data. However, this focus on identifiable individuals does not address new opportunities for privacy violations arising from big data analytics. Groups can suffer discrimination or preferential treatment, driven by analytics without members ever being identified. Insufficient protections exist for ad hoc groups (who are identified externally through having perceived links). Groups can conceivably be ascribed some rights but it is unlikely that members of ad hoc groups have a right to group privacy.

Discussion

- How do we encourage dialogue, and avoid de-professionalisation and de-responsibilisation?
- How do we provide decision makers in healthcare with sufficient opportunities to interrogate different models?
- Determining what constitutes meaningful transparency in a given context is a key priority for the field

Delegate response

- Incremental system changes are incompatible with one off certification. Developing governance that takes account of the dynamic nature of algorithm development is challenging. Consent might be used for very sensitive data but is not suitable for other routine data collection
- Impact assessments may be another method of risk mitigation but fail to take account of risk mitigation methods that do not meet high ethical standards of auditing
- These issues could also be applicable to other areas of medicine such as recruitment for, and efficacy of treatment in clinical trials which show variable rates of effectiveness in different groups

4 Algorithms as data: the GDPR and the right to explanation

The workshop then focused on algorithms as data, concentrating on the regulatory framework created by the GDPR, which came into force on 25 May 2018.

4.1 The context of the right to explanation

The use of machine learning algorithms is becoming increasing common across the NHS, particularly in sectors such as imaging. The GDPR aims to create a robust and harmonised regulatory framework that applies across the European Union. As well as strengthening the rights of data subjects and imposing greater responsibilities on those controlling and processing data (data controllers and data processors), the GDPR contains specific provisions relating to automated decision-making, including profiling.

The strengthened rights for data subjects include general provisions setting out the information that has to be provided to the data subject and other rights of data access, portability and erasure.

This part of the workshop aimed to understand the nature of the 'right to explanation' created by the GDPR, and then assess the potential impact of this right on the algorithms and software sector.

Specific questions for the session included:

- Does the existence of a right of explanation place a burden on developers and users of algorithms and software?
- What are the likely implications for the uptake and dissemination of these technologies in the health sector?

4.2 Does a right to explanation exist?

Sandra Wachter, Oxford Internet Institute; Alan Turing Institute

Automated decision-making and artificial intelligence (AI) are used across multiple sectors. Increasing reliance on AI and black box algorithms create potential challenges of opacity, bias and discrimination. Against that backdrop, the aim of the GDPR is to enforce a harmonised data protection standard in the EU through regulation of the use of personal data and establishment of data processing principles.

Since the GDPR includes provision for substantial fines for non-compliance (up to 4% of total worldwide annual turnover (Article 83)), there is a need for clarity about what the GDPR requires.

There are a number of different types of explanation ranging from system functionality through to a rationale being supplied for an individual decision. Contrary to what some commentators claim, the GDPR does not create a right to explanation about individual decisions. The right to be informed in the GDPR is a limited right to be provided with some general information if automated processes are utilised in data processing[15]. This right to be informed is only triggered if a decision is based solely on automated processing; and if the decision has legal effects or similarly significant effects for the individual (Article 22).

During the legislative process of the Data Protection Bill 2018 through the UK Parliament there were attempts to strengthen the legal obligation to provide an explanation after a decision had been reached but these were unsuccessful.

In the context of healthcare, currently there will usually be some form of human intervention from a healthcare professional in the decision making process. Note that 'legal effect' or 'similarly significant' effect in the health context could include restricting access to health services, or a similar decision which carries significant disadvantage.

Where Article 22 is triggered, the data controller has to implement suitable safeguards to protect the data subject. These include a right of human intervention or for the data subject to express their point of view and contest the decision. Additional provisions in Articles 13 and 14 (rights of information), which are replicated in Article 15 (right of access) also provide that the data subject should be given information on the existence of automated processing, meaningful information about the logic involved as well as the significant and consequences of such processing for the data subject.

Each of these Articles require information about system functionality to be provided before data has been processed. In contrast, Recital 71 (which is not legally binding) provides for the data subject to 'obtain an explanation of the decision reached after such assessment'. This provision was not included in the body of the GDPR, suggesting that the legislators did not want to grant a similar binding legal requirement for data subjects.

The use of robotics and AI increases the possibility that a remedy might be needed if somebody is harmed as a result of a system failure. For example, the European Parliament's Committee on Legal Affairs (JURI) has tabled a report to establish principles of civil liability for damages caused by robots and associated ethical aspects[16]. In this report, provision of information and explanation are regarded as distinct.

Although explanation is important, it does not require an explanation to be given and alternatives to explanation might be needed to provide additional support for data subjects. These include provision of an AI 'watchdog' in the case of complaints; provision of a regulatory body for auditing of algorithms; internal audit and algorithm certification.

Discussion

- Information provision and the right to explanation should be regarded as distinct
- Precise regulation is needed together with better methods to certify, explain and audit inscrutable systems[17]

Delegate response

- Some in the audience felt that Article 22 was not adding much value where the legal basis for processing is consent
- Others felt that it might be helpful to compare the nature and extent of explanation required if data is processed using humans rather than automated processing
- There was discussion about how the rights under Article 13, 14, and 22 might be met where data is processed that generates new personal data. The speaker's view is that the right to be informed arises before data processing rather than after it. Some in the audience disagreed

4.4 Counterfactual explanations and the GDPR

Sandra Wachter, Oxford Internet Institute; Alan Turing Institute

Even if it is possible to generate a satisfactory explanation, this may not be very meaningful to patients or consumers who might have different expectations of what an explanation should deliver. These include:

- Understanding the reasons why a decision has been made
- Obtaining sufficient information to be able to challenge an undesirable decision
- Being able to influence future decisions

For this reason, a counterfactual (i.e. a description of dependency on the external facts that led to a decision) might better serve patients/consumers. The objective of a counterfactual is to convey the smallest change that can be made to obtain a desirable outcome.

Counterfactual example:

You were denied a loan because your annual income was £30,000. If your income had been £45,000 you would have been offered a loan.

Advantages of this approach include that a counterfactual does not require understanding of the internal logic of the algorithm and can be generated even if complex methods are used. For this reason, it is less likely to infringe the rights/freedoms of others or trade secrets and reveals minimal proprietary information[18]. Excessive disclosure of information about the internal logic of a system could infringe the rights of others, either by revealing protected trade secrets or by violating the privacy of individuals whose data is contained in the training dataset.

4.5 Will counterfactual explanations satisfy the right to explanation?

Articles 13/14 of the GDPR require that a data subject is provided with meaningful information about the logic involved and the significance and envisaged consequences of such processing for the data subject. Article 12(7) suggests that standardised icons can be used to give an easily visible, intelligible and meaningful overview of the intended processing. However these requirements do not clarify precisely what content is required.

The 'logic involved' includes the main characteristics, source of information and relevance (Article 29 Working Party[19]) and 'categories of data, sources and why data is considered relevant' (UK Information Commissioner's Office). Although existing guidelines are opaque, they are consistent in suggesting that the rights to data access and transparency are not equivalent to a right to explanation.

Although data controllers are required to facilitate the exercise of data subject rights under Articles 15-22, this does not explicitly require data subjects to be informed that they can contest a decision.

Since a counterfactual explanation does not include an explanation of the logic of an algorithm, lack of interpretability becomes a potential problem. Limiting a counterfactual explanation to a specific case also means that it cannot demonstrate whether a system is operating fairly. If models are very dynamic, a counterfactual explanation will not inform future decisions. Large scale statistical analyses would be a better means of uncovering systematic biases.

There may also be wider concerns that the use of algorithms for decision making allows the outsourcing of moral responsibility to AI, suggesting that there may be some applications for which machine learning could be regarded as unethical.

Discussion

- Satisfying a legal requirement is very different from being satisfactory or satisfying for patients. Is a counterfactual explanation useful for patients?
- Does greater reliance on technology in health affect our trust in doctors and other health professionals?
- There may also be wider concerns that the use of algorithms for decision making allows the outsourcing of moral responsibility to AI. Are there some applications for which machine learning should be mandated (if algorithmic performance is better than humans), or others for which machine learning could be regarded as unethical?

Delegate response

- The choice of counterfactual might depend on its utility for the patient, bearing in mind that using multiple counterfactuals might be potentially disclosive. Commercial interests may guide the choice of disclosable characteristics. The choice of which counterfactual to use adds another layer of complexity, including whether selection is made by human intervention or using an algorithm
- Commercial and political interests might influence which characteristics are highlighted in framing a counterfactual example. Research is underway to identify reasonable set of counterfactuals for given applications
- Could a counterfactual explanation constitute a legal representation (i.e. an express or implied statement made by one contractual party to another that could influence the consummation of a deal)?
- Arguably, counterfactual explanations might not satisfy the requirements of the GDPR, since they do not necessarily address the functionality of an algorithm (see above) but rather the outcome of an individual decision[20]

5 Introduction to the IVDR

Sara Payne, PHG Foundation

Algorithms are being used in healthcare to assist in the diagnosis and treatment of patients. Many of these devices will be regulated as medical devices and so be subject to specific requirements and standards to gain and retain CE marking. CE marking confirms that the medical device meets certain requirements under EU legislation and meets the intended purpose. It also shows that the medical device can be freely marketed anywhere in the EEA. Medical device law is the primary means by which authorities establish that medical devices are safe for use in patient populations. This section briefly introduces the EU In Vitro Diagnostic Medical Device Regulation (IVDR) and describes how medical device law regulates software.

5.1 Introduction to the IVDR

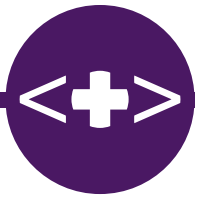
The IVDR is one half of the new set of EU Regulations, the other half being the Medical Device Regulation (MDR). These two replace the previous set of Directives: the Medical Device Directive, In Vitro Diagnostic Medical Device Directive, and Active Implantable Medical Device Directive. The IVDR entered into force on 25 May 2017 and will fully apply in Member States from 22 May 2022.

The aim of introducing the new IVDR and MDR was to tighten up regulation by addressing well-known flaws in the previous set of directives. Specifically, cases such as the widely reported PIP (Poly Implant Prothèse) breast implant scandal drew attention to the inadequacies of the Directives, highlighting that that the Directives often offered low scrutiny to high-risk devices. In drawing up the Regulations, the EU Commission emphasised proportionality - much of the new legislation being sensitive to the potential risks and benefits of devices. Moreover, the new IVDR and MDR also respond to the changing medical device market, with greater numbers of devices now being used by consumers themselves rather than solely by clinicians.

What are the key changes in the IVDR?

The IVDR differs in three key ways from its predecessor Directive:

- The scope of the IVDR is broader, with a wider definition of IVDs that include 'accessories' and components that 'influence' devices[21]. Additionally, there are new definitions that include 'devices for near patient testing', 'self-testing', and 'companion diagnostics'[22]
- The IVDR and MDR introduce a new system to keep track of individual medical devices with unique device identification requirements. This new obligation aims to improve identification and traceability of devices
- The IVDR imposes stricter classification rules, including tighter assessment of 'risk' in the classification rules through Class A (low) through Class B, C, and D (highest)[23] and up classification of riskier devices



What challenges might the IVDR present to algorithms?

To what extent does the IVDR improve upon the provision in the former directives for:

- Smoother functioning of the internal market?
- Better quality and safety of medical and in vitro diagnostic devices?

Discussion

- Will the five year transition period to put IVDR fully in place affect development of algorithms in medical devices?
- Is there a need for a new designated software regulator to properly regulate algorithms in devices?

6 Qualification, validation, and surveillance of software as a MD/IVD

Johan Ordish, PHG Foundation

Software may provide a unique challenge for the dual mission of the Regulations: smoother functioning of the internal market and better safety and quality of medical devices.

This section considers three problems:

- Qualification and intended purpose requirements in both EU and US jurisdictions
- How machine learning might constitute a moving target for regulators
- How the black box nature of some machine learning models might constitute an issue for regulators

6.1 Qualification as a medical device or in vitro diagnostic medical device

As digital health apps proliferate, there is increasing debate about the reach and impact of the revised regulatory framework and whether it is robust enough to protect users from harm. Some commentators (e.g. Quinn 2017[24]) claim that the MDR and IVDR intended purpose test captures too few digital health apps that might pose a significant risk to human health, and argue that the US FDA's strategy is reckoned to be more flexible, capturing devices that pose a risk to human health.

The consensus view is that the EU and US adopt different strategies and therefore offer substantively different scope to regulate software as a medical device. Arguably however, the two jurisdictions are converging on how they regulate software as a medical device as explained in the following sections.

What is intended purpose under the EU MDR/IVDR?

If a device is to be regulated as a medical device, it must qualify as a medical device.

The broad definition of medical device is found in Article 2(1) MDR:

> "'Medical device' means any instrument, apparatus, appliance, software, implant, reagent, material or other article intended by the manufacturer [author's emphasis] to be used, alone or in combination, for human beings for one or more of the following specific medical purposes…"

'Intended purpose' is clarified in Article 2(12) MDR:

> "'Intended purpose' means the use for which a device is intended according to the data supplied by the manufacturer on the label, in the instructions for use or in promotional or sales materials or statements and as specified by the manufacturer in the clinical evaluation…"

Intended purpose in relation to software is clarified further in Recital 19 MDR:

> '…software in its own right, when specifically intended by the manufacturer to be used for one or more of the medical purposes set out in the definition of a medical device, qualifies as a medical device…'

Recital 19 MDR also distinguishes between software for general purposes and software intended for life-style or well-being purposes, noting that the latter examples of software do not qualify as medical devices.

Under the Directives, the European Court of Justice gave further clarification as to what counts as a 'medical purpose' in 2012 case C-219/11 *Brain Products* noting:

> '…it is not sufficient that it [the device] be used in a medical context, but that it is also necessary that the intended purpose, defined by the manufacturer, is specifically medical'[25].

The current position then is that medical device (and IVD) qualification is resolved by the manufacturer – competent authorities (such as the UK MHRA) may not substitute their own interpretation of the purpose of a device to have it qualify or not qualify[26].

What is the US FDA risk-based test?

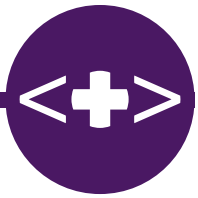
The US method of device qualification differs from the EU strategy. Commentators such as Quinn (2017) have noted that the US strategy emphasises risk. One of the pieces of evidence for this assessment is the FDA's 2015 document Mobile Medical Applications: Guidance for Industry and Food and Drug Administration Staff (Mobile Medical Applications). This document notes that the FDA intends to apply its regulatory oversight to 'only those mobile apps that are medical devices and whose functionality could pose a risk to a patient's safety if the mobile app were to not function as intended'[27].

Released in 2015, this US Mobile Medical Applications guidance is now out of date. Specifically, Section 3060 of the 21st Century Cures Act 2016, relates to software as a device:

> 'Certain software is exempted from requirements for medical devices, including software that provides medical recommendations and the basis for those recommendations to health care professionals. Software remains subject to regulation as a medical device if: (1) the software acquires, processes, analyzes, or interprets medical information; or (2) the FDA identifies use of the software as reasonably likely to have serious adverse health consequences.'

In this way, the US FDA no longer has the ability to regulate low risk devices if these devices do not acquire, process, analyse, or interpret medical information.

Regardless, how do the two jurisdictions compare in the light of one of the most recent ECJ preliminary rulings C-329/16 SNITEM[28]? This issue is the subject of the next section.



Is there a difference between the EU and US FDA approaches?

In considering this question there are two main issues to consider:

- The centrality of 'intended purpose' in each approach
- Does the EU approach have similar scope to capture digital health apps as the US FDA approach?

In regards to the EU strategy, the typical approach is to question the manufacturer's intended purpose and whether this is specifically medical. Broadly, Article 2(12) MDR notes that evidence of this intended purpose is to be found in the material that comes with the device: labelling, promotional material and so on. However, does the recent case of SNITEM signal a shift in this EU approach?

SNITEM recites much of the EU MEDDEV guidance one would expect the Court to cite, i.e. *Brain Products*, and the relevant parts of the directive. However, the approach in SNITEM is rather brisk in comparison to past cases with the key part of the judgment stating:

> 'I[t] follows that software, of which at least one of the functions makes it possible to use patient-specific data for the purposes, inter alia, of detecting contraindications, drug interactions and excessive doses, is, in respect of that function, a medical device'

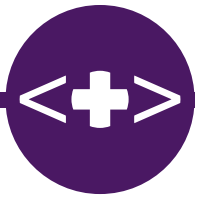
In this way, the Court emphasised the function of the device, noting in strong terms that if a device has a certain function it follows that it qualifies as a medical device. The language used in this passage suggests a shift in the Court's method of determining device qualification. If true, then the EU and US strategies might be closer than commentators such as Quinn previously noted. Arguably, both jurisdictions primarily look to function to determine whether a device qualifies as a medical or in vitro diagnostic device.

6.2 Machine learning as a medical device

The medical device landscape is changing with standalone software becoming more prominent and central to the medical device market. Machine learning is an especially promising technology that has attracted large amounts of venture capital funding to develop its capability as a medical device. In some instances this technology will be regulated as a medical device. Indeed, the US FDA has already given clearance to a select few machine learning medical devices. However, in EU and US regulation machine learning might pose special problems for medical device law and the regulators that have to enforce its standards.

The following sections raise two issues that machine learning might pose:

- Machine learning might lend itself to being more iterative and dynamic, retraining its models often – in this way, it might constitute a moving target
- Machine learning might be more opaque than its explicitly programmed counterparts, being less susceptible to human interpretation



Dynamic machine learning devices

With some simplification, traditional programming combines data with the program to produce the desired output, whereas machine learning combines the data and output to create the program[29]. Machine learning 'trains' using data to create models, these models may be updated frequently, changing to improve performance in light of new data. As a consequence, machine learning, especially those models that utilise incremental learning and streaming data, may constitute an extremely dynamic device to assess. In this way, it may be difficult to ensure the safety and effectiveness of a device that changes so frequently.

Black box validation

Some machine learning models are black boxes, that is, not readily human interpretable. While they may provide highly accurate recommendations, some methods like deep neural decision trees do not lend themselves to explanation. This raises a number of questions:

- Does the opacity of some machine learning models present a barrier to their validation under the MDR, IVDR, and their associated harmonised standards?
- Does the validation and surveillance demanded by medical device law require human interpretability?

Discussion

- Is there much of a difference between the EU and US methods of device qualification?
- How can machine learning algorithms be validated as a medical device?
- What – if any – is the role of explanation in validation?

Delegate response:

- There is often a gulf between 'intended use' and 'actual use' in practice – a product may be regulated for one use and then used for another
- Intended use is assessed over the lifecycle of the product. Moreover, negative labelling, (i.e. stipulating that a product is not to be used for a particular use) might assist in supplementing the concept of intended purpose
- Does changing the underlying data used for machine learning constitute a change to the algorithm itself? What about retraining?
- Currently it is regarded as too dangerous to introduce new data into a live device for medical purposes as this could cause a significant change in the model and harm patients
- Should the regulatory process speed up to keep pace with the retraining of models etc. Is regulation changing too slowly?
- What algorithmic changes count as 'substantial changes'[30]? How do we assess when a device has changed so much that new performance checks are required?

7 Discussion

There was consensus among the presenters and delegates that the right to explanation is important. Whilst there are a variety of different understandings across different contexts, there is a need to develop explanations that are relevant in terms of content, application and audience. More normative discussion is needed about how much explanation is desirable, in order to understand the scope and impact of the GDPR on the right to explanation. Rather than providing a rational understanding of how an algorithm works, it was agreed that the right to explanation in the GDPR is a means of ensuring accountability.

In a healthcare context, increased reliance on digital health suggests that relying too heavily on complex explanations would not be feasible. This is the case for both healthcare systems and for patients who could quickly become overwhelmed by information and requests for consent. Healthcare professionals will need new competences to fully utilise this high-tech environment.

More work is needed to develop examples of counterfactual explanations in health. Examples of where counterfactual explanations might be relevant are where a patient is denied a healthcare intervention which they request, such as in assisted reproduction, or where medical technicians need to understand black box medical technologies such as in pathology or imaging.

As apps are rolled out more widely other areas of law, such as consumer protection law, could become more relevant. Appropriate ethical standards may also have a role in supporting good governance and in filling gaps. However, there was some concern that there is inadequate ethical oversight in private institutions compared to publicly funded organisations, and there was discussion about how ethical standards can be more effectively mandated across institutions.

Delegates also noted the difference between intent and effect for these apps: there are major risks for the public health system if apps generate information that is then acted on by consumers, creating additional demands on health systems.

The focus of this workshop was on issues arising from forthcoming changes to EU legislation. There are many other challenges arising from the ubiquity of algorithms and software in healthcare. The PHG Foundation held a second workshop in September 2018 to explore the challenges relating to proprietary and non-proprietary rights in algorithms and software, and liability associated with their use. The findings of the first workshop contributed to the PHG Foundation report *Algorithms as medical devices*. It is hoped that these reports will help to inform evolving policy in this area.

Appendices

Workshop agenda

AGENDA

phg foundation
making science work for health

Regulating algorithms in healthcare: the GDPR and IVDR in practice

King's College Cambridge, Tuesday 20 March 2018

09:30	**Arrival - coffee and pastries**	
10:00	Welcome and introductions	**Alison Hall** *Head of Humanities, PHG Foundation*
10:30	Defining algorithms and software	**Johan Ordish** *Senior Policy Analyst (Law and Regulation), PHG Foundation*
11:00	The ethics of algorithms: mapping the debate	**Dr Brent Mittelstadt** *Oxford Internet Institute*
11:45	**Morning tea**	
12:00	GDPR Does a right to explanation exist?	**Dr Sandra Wachter** *Oxford Internet Institute*
12:45	**Working lunch**	
13:30	GDPR continued Counterfactual explanation as a possible solution	**Dr Sandra Wachter** *Oxford Internet Institute*
14:15	IVDR IVDR introduction	**Sara Payne** *PHG Foundation* *MHRA*
14:45	**Afternoon tea**	
15:00	IVDR continued MD / IVD defintion, validation and surveillance	**Johan Ordish** *Senior Policy Analyst (Law and Regulation), PHG Foundation*
15:45	Plenary session: develop action points	**Alison Hall** *Head of Humanities, PHG Foundation*
17:00	**End of workshop**	

@PHGFoundation

CAMBRIDGE UNIVERSITY Health Partners

Workshop delegates

Name	Organisation
Mateo Aboy	Centre for Law, Medicine and Life Sciences, University of Cambridge
Joo Wook Ahn	Department of Clinical Genetics, NHS
Ranveig Berg	Nuffield Council on Bioethics
Tanya Brigden	PHG Foundation
Sarah Cook	PHG Foundation
Cristina Crespo	Centre for Law, Medicine and Life Sciences, University of Cambridge
Sylvie Delacroix	Law School, University of Birmingham
Alison Dennis	Fieldfisher
David Erdos	Centre for Law, Medicine and Life Sciences, University of Cambridge
Guido Fumagalli	MHRA
Jo Gibbs	Department of Infection and Population Health, University College London
Alison Hall	PHG Foundation
Stuart Hogarth	Sociology, University of Cambridge
Hamza Javed	Oxford Computational Health Informatics group
Dimitra Kamarinou	Centre for Commercial Law Studies, Queen Mary University of London
Zisis Kozlakidis	ICONIC, University College London
Stephen Lee	MHRA
Kathleen Liddell	Centre for Law, Medicine and Life Sciences, University of Cambridge
Brent Mittelstadt	Oxford Internet Institute
Johan Ordish	PHG Foundation
Sara Payne	PHG Foundation / MHRA
Paul Quinn	Law, Science, Technology and Society Studies, Vrije Universiteit Brussel
Christopher Russell	Department of Computer Science, University of Surrey
Mark Salmon	NICE
Mark Shaw	School of Computer Science and Informatics, De Montfort University
Jat Singh	Department of Computer Science, University of Cambridge
Jeff Skopek	Centre for Law, Medicine and Life Sciences, University of Cambridge
Allan Tucker	Intelligent Data Analytics Group, Brunel University London
Sandra Wachter	Oxford Internet Institute
Jeremy Wyatt	Wessex Institute of Health, University of Southampton

References

1. Regulation (EU) 2016/679
2. Regulation (EU) 2017/745
3. Regulation (EU) 2017/746
4. Cormen TH. Algorithms Unlocked. MIT Press; 2013
5. Knuth DE. The Art of Computer Programming: Volume 1: Fundamental Algorithms. Addison Wesley; 1997
6. *Gottschalk v. Benson*, 409 U.S. 63 (1972): 409
7. *In re Alappat*, 33 F3d 1526 (Fed Cir. 1994)
   Bilski v. Kappos, 561 U.S. 593 (2010)
   Alice Corp v. CLS Bank International, 573 U.S. 208, 134 S. Ct. 2347 (2014)
8. Moor JH. Three myths of computer science. The British Journal for the Philosophy of Science. 1978; 29(3): 215
9. Stapleton J. Product Liability. Butterworths; 1994
10. Angwin J, Larson J, Mattu S, *et al*. Machine Bias. ProPublica. 2016
11. Hill RK. What an algorithm is. Philosophy & Technology. 2015; 29(1): 35-59
12. Mittelstadt B, Allo P, Taddeo M, et al. The ethics of algorithms: Mapping the debate. Big Data & Society. 2016; 3(2): 1-21
13. Mittelstadt B, Russell C, Wachter S. Explaining Explanation in AI. FAT 2019 Proceedings. 2018: 1-10
14. Kleinberg J, Mullainathan S, Raghavan M. Inherent Trade-Offs in the Fair Determination of Risk Scores. Proceedings of Innovations in Theoretical Computer Science. 2017; 1-23
15. Wachter S, Mittelstadt BD, Floridi L. Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. International Data Privacy Law. 2017; 7(2): 76-99
16. Nevejans N. European Civil Law Rules in Robotics. JURI Committee. 2018
17. Wachter S, Mittelstadt B, Floridi L. Transparent, explainable, and accountable AI for robotics. Science Robotics. 2017; 2(6): 1-2
18. Wachter S, Mittelstadt BD, Russell C. Counterfactual explanations without opening the black box: Automated decisions and the GDPR. Harvard Journal of Law & Technology. 2018; 31(2): 1-44
19. Article 29 Data Protection Working Party. Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. Adopted 3 October 2017
20. Selbst A, Powles J. Meaningful Information and the Right to Explanation. International Data Privacy Law. 2017; 7(4): 14-15
21. Regulation (EU) 2017/746, Article 1(1) and Annex VIII 1.4
22. Regulation (EU) 2017/746, Article 2(5)-(7)
23. Boumans R. Understanding Europe's New In Vitro Diagnostic Medical Devices Regulation: What manufacturers need to know ahead of IVDR implementation. 2016
24. Quinn P. The EU Commission's risky choice for a non-risk based strategy on assessment of medical devices. Computer Law & Security Review. 2017; 33(3): 361-370
25. C-219/11 *Brain Products GmbH v BioSemi VOF, Antonius Pieter Kuiper, Robert Jan Gerard Honsbeek, Alexander Coenraad Metting van Rijn*
26. Vollebregt, E. An unsurprising case of software qualification with an interesting twist [Internet]. Medicaldeviceslegal; 2018
27. FDA. Mobile Medical Applications: Guidance for Industry and Food and Drug Administration Staff. 2013: 4
28. C-329/16 *Syndicat national de l'industrie des technologies médicales (Snitem), Philips France v Premier ministre, Ministre des Affaires sociales et de la Santé*
29. Brownlee J. Basic Concepts in Machine Learning. Machine Learning Mastery; 2015
30. Annex IX, Chapter I, 2.4, Regulation (EU) 2017/745

About the project

Regulating algorithms in healthcare is a PHG Foundation project that clarifies:

- How algorithms in healthcare are regulated
- How algorithms in healthcare should be regulated

Regulating algorithms in healthcare considers how algorithms in healthcare are regulated, starting from the data that is used to train an algorithm through to the question of who is liable if something goes wrong. The project considers the following general spheres of regulation:

- Algorithms as data (the General Data Protection Regulation 2016 (GDPR) and the UK Data Protection Act 2018 (DPA))
- Algorithms as medical devices (the Medical Devices Regulation 2017 (MDR) and In Vitro Diagnostic Medical Devices Regulation 2017 (IVDR))
- Algorithms as intellectual property (including patent, copyright and trade secret protections)
- Algorithms as a source of liability (clinical negligence, product liability and statutory compensation schemes)

Working with the Centre for Law, Medicine and Life Sciences at the University of Cambridge, the project convened two workshops bringing together academics, developers, researchers; legal practitioners, health professionals, regulators, and policy makers.

Workshop 1 – Regulating algorithms in healthcare – the GDPR and IVDR in practice considered the following issues:

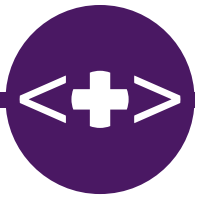
In regards to algorithms as data:

- The particular ethical issues algorithms might pose
- Whether the GDPR contains a right to explanation?
- Whether counterfactual explanations might satisfy such a right?

In regards to algorithms as medical devices:

- The position of software under the IVDR
- How software qualifies as a medical device in EU and US law
- The particular problems machine learning might pose for medical device regulation

Workshop 2 – Regulating algorithms in healthcare – intellectual property and liability considered



the following issues.

In regards to intellectual property:

- The current patentability of 'computer implemented inventions'
- The viability of using open source software in the healthcare sector

In regards to the source of liability:

- What scheme of liability might be most appropriate for artificial intelligence?
- The place of predictive analytics in medical malpractice

This report describes the presentations and the discussions from Workshop 1.

The findings from this workshop prompted a more detailed analysis of software as medical devices and publication of a PHG Foundation report - *Algorithms as medical devices.*

phg foundation

making science work for health

About the PHG Foundation

The PHG Foundation is a pioneering independent think-tank with a special focus on genomics and other emerging health technologies that can provide more accurate and effective personalised medicine. Our mission is to make science work for health. Established in 1997 as the founding UK centre for public health genomics, we are now an acknowledged world leader in the effective and responsible translation and application of genomic technologies for health.In April 2018 we became part of the University of Cambridge.
We create robust policy solutions to problems and barriers relating to implementation of science in health services, and provide knowledge, evidence and ideas to stimulate and direct well-informed discussion and debate on the potential and pitfalls of key biomedical developments, and to inform and educate stakeholders. We also provide expert research, analysis, health services planning and consultancy services for governments, health systems, and other non-profit organisations.

phg foundation
making science work for health

PHG Foundation
2 Worts Causeway
Cambridge
CB1 8RN

+44 (0) 1223 761900

@phgfoundation

www.phgfoundation.org