House of Commons Science and Technology Committee Call for Evidence: 'The Right to Privacy: Digital Data'

British Heart Foundation Response - February 2022

About the BHF

The British Heart Foundation (BHF) is the largest independent funder of research into heart and circulatory diseases and the third largest charitable funder of medical research in the UK. Each year, thanks to the generosity of our supporters, we are able to fund around £100 million of new research across the UK. The research we fund has helped halve the number of people dying from heart and circulatory conditions since the 1960s. Despite these breakthroughs, our work is not done, as there are still more than 7 million people living with heart and circulatory diseases in the UK and these diseases cause more than a quarter of all UK deaths.

The BHF supports an extensive portfolio of projects focused on the use of health and care data, and is privileged to work with a world-class community of researchers who utilise the considerable health data asset of the UK on a day-to-day basis:

- In 2019, we announced the launch of a £10 million BHF Data Science Centre (BHF DSC) in partnership with Health Data Research UK (HDR UK), to promote the safe and ethical use of data for research into the causes, prevention and treatment of all diseases of the heart and circulation.[1]
- To date, we have supported 12 BHF/Alan Turing Institute Cardiovascular Data Science Awards, at a value of over £0.6 million. This is a joint funding scheme to support collaborative research between cardiovascular investigators and data scientists seeking to generate data science solutions to key cardiovascular problems.[2]
- We provide significant funding to both HDR UK and the UK Biobank (currently funding £2 million and £3.2 million over five years, respectively).
- In addition to these strategic investments, as of March 2021 we have supported over £30m worth of research projects related to data science, machine learning and artificial intelligence.
- In 2021, we launched a £1 million fund to explore how NHS data can be harnessed to improve delivery of care and patient outcomes. The Cardiovascular Catalyst Awards fund research into the use of advanced analytics to improve cardiovascular care in the NHS, which has suffered devastating impacts due to Covid-19. Funded research projects will utilise NHS data, such as clinical data and healthcare records.[3]
- We have expert patient panels that ensure that our data science work is always supported by a strong patient voice, encouraging, and challenging us on matters of engagement and transparency.

Key messages and recommendations

Patient medical records hold information that can help BHF-funded researchers make vital new life-saving discoveries. However, to unlock the full potential of this data we need a system of data access that protects patient privacy while also enabling cutting edge research. We believe that the public should understand how their data are being used and that they should have the right to object to this use if they wish. Building public trust in data use is key, as is the need for transparency with respect to how individuals' data are used and shared. As such, we welcome the opportunity to share our thoughts on how data can be shared, safely and appropriately, in order to generate maximum benefit for patients and researchers.

Our response focusses on the use of health data for research, including the benefits and barriers to safe and appropriate data sharing, suggestions of how Government can address these barriers, and finally an assessment of the extent to which data issues are appropriately addressed by Government strategies and legislation. We will not be covering issues pertaining to the ethics of data sharing as this topic is outside of our scope and expertise.

1. Connecting data more effectively from different parts of the system holds enormous potential benefits for public health

Health data are an incredibly powerful tool for driving health and research discoveries and innovations, and the UK is in an enviable position in this space, thanks to rich health datasets held by the NHS. Safely unlocking this resource for researchers is vital for tackling every health challenge faced by the UK – from cardiovascular disease to Covid-19.

There is a balance to be struck between allowing access to data for people who can use it to improve our health, such as researchers, and protecting the privacy of individuals.

2. Public trust is key to effective data sharing

Government must prioritise public trust and engagement around data use and sharing, learning lessons from the failures of the GPDPR rollout. In particular, it should address areas where the public has strong concerns, e.g., the sharing of health data with private companies and the use of health data for non-research purposes without clear consent.

Using health data for research purposes will never be completely risk-free; therefore, it is paramount that appropriate safeguards and privacy are put in place to protect it and that these measures and risks are communicated clearly to the public.

BHF recommends that any future reforms to patient data sharing must include a public engagement plan that reaches an audience that accurately reflects the demographics of the UK and clearly communicates the benefits to data sharing, as well as the risks and safeguards so that the public can make an informed decision on sharing their data.

3. Shared data must be representative of the population

Ensuring that shared data are representative of historically marginalised groups, such as women and people from a Black and minority ethnic background, is vital for understanding and tackling health and research inequalities.

Ensuring available health data are representative of the population should be a greater focus of this Government.

4. Introducing compulsory transparency reporting on algorithm use could help to reduce bias and increase public trust

If unregulated, biases in algorithmic decision-making reflect wider inequalities in society which in turn exacerbates existing health and research inequalities.

BHF recommends Government should introduce compulsory transparency reporting on the use of algorithms in decision-making for public authorities, Government departments and Government contractors using public data.

5. Lack of clear standards for Trusted Research Environments (TREs) limits data sharing efficiency and interoperability

TREs are a valuable tool for providing safe access to sensitive data; however, a lack of clear standards for these is a barrier to efficient data sharing and must be improved so that different TREs can work together.

BHF supports NHSX's TRE ambitions to develop a minimum technical specification for TREs, new governance standards and an accreditation framework.

6. Gaps in training for data researchers represent a barrier for safe data sharing

Insight from our cardiovascular research community highlights continued confusion and lack of understanding as to how best navigate data sharing.

BHF recommends that data regulations are communicated clearly with the research community. This should include clarity around applicable regulation, the sharing of practical resources, and strong case studies. We recommend improved guidance from regulators and greater focus on education and upskilling of data custodians, researchers, and host organisations on data protection legislation.

7. Inefficient Government and NHS technology limits data sharing potential

Persistent inefficiencies of 'outdated' Government and health service technology systems are holding back innovation in data sharing.

Government should invest in improvements in the quality, linkage, and interoperability of NHS data systems so research that uses health data is as efficient as possible.

8. The upcoming merger of NHS Digital into NHS England has potential to disrupt data sharing

To transfer management of the National Disease Registration Service (a collection of data on congenital anomalies, rare diseases and all cancers diagnosed in England) from Public Health England to NHS Digital in 2021, data access requests were put on pause for several months. This year, there will be a similar upheaval, as NHS Digital merges into NHS England, and lack of communication around implications of this move is causing uncertainty in the sector.

BHF would welcome more communication on the upcoming merger of NHS Digital into NHS England to protect against unintended consequences and mitigate potential disruptions to data sharing.

9. Changes to data legislation must not jeopardise UK-EU adequacy agreements

As stressed by many in the research sector, it is in the best interests of all patients that UK and EU researchers be allowed to continue exchanging data - an essential part of collaborative research – now that the UK has left the EU.

BHF believes UK-EU data adequacy agreements must be maintained, regardless of changes to UK data legislation.

1. The potential benefits, including to research, to effectively use and share data between and across Government, other public bodies, research institutions and commercial organisations

1.1 Collecting, linking and using patient data is vital for managing an individual patient's care, for example by improving understanding of individual risk or monitoring the safety of a treatment over time. More broadly, if data from many patients are linked up and pooled, researchers and doctors can look for patterns in the data, helping them develop new ways of predicting or diagnosing illness, and identify ways to improve and plan clinical care. The UK has an enviable position in this space, thanks to rich health datasets held by the NHS – a uniquely centralised healthcare system that allows for detailed insights of a large and diverse population.

1.2 In recent history, access to patient data has helped researchers unlock major medical breakthroughs, including:

- Establishing the link between smoking and diseases such as lung cancer and heart disease
- Confirming the health risks of asbestos
- Showing that high blood pressure increases heart disease risk
- Identifying out how the AIDS virus is transmitted
- Demonstrating the benefits of lowering cholesterol using statins, now a standard treatment for patients at high risk of heart disease.[4]

1.3 Enhancing data sharing, while maintaining necessary safeguards, allows for even richer data insights. This provides huge potential for new clinical options and improved care pathways. Below are two case studies which highlight the value of data sharing in BHF-funded projects:

Case Study 1: The BHF Data Science Centre's role in the Covid-19 pandemic

In 2019, BHF announced the launch of a £10 million BHF Data Science Centre (BHF DSC) in partnership with Health Data Research UK (HDR UK). The Centre was established to work in partnership with patients, the public, the NHS, researchers and clinicians to promote the safe and ethical use of data for research into the causes, prevention and treatment of all diseases of the heart and circulation.

During the Covid-19 pandemic, the Centre played a key role in leveraging the UK's exceptional health data to help answer some of the most important heart and circulatory disease research and policy questions. For example, the Centre established the CVD-COVID-UK consortium, bringing together more than 230 interdisciplinary researchers (clinicians, epidemiologists, biostatisticians, computational scientists and others) from across more than 45 institutions in the UK to better understand the relationship between Covid-19 and cardiovascular diseases through analyses of de-identified, linked, nationally collated healthcare datasets across the four nations of the UK.

The breadth and depth of the sixteen projects approved so far under its banner give testament to the importance of such a consortium. Current work includes looking at the risk of blood clots in the arteries (e.g. heart attacks and strokes) and veins (e.g. clots in the veins of the legs and

lungs) following Covid-19 infection; developing new analysis approaches to best handle the huge quantities of data available; and a breadth of other projects, such as predicting future risk of cardiovascular diseases and studying the impact of cardiovascular diseases combined with other conditions (so-called 'multimorbidity'). Results from this work are now coming through, being submitted to journals, and have been used to inform bodies regulating vaccines and other medicines as well as senior government health and science advisers. For example, a recent paper used data from 15.8 million individuals to analyse the adverse impact of the pandemic on cardiovascular disease prevention and management in England, Scotland and Wales, providing evidence for the urgent need to alleviate the backlog of care.[5]

Additional achievements of the Centre include:

- *In April 2020, the BHF and the BHF DSC supported the development of a paper*[6] *that fed into SAGE decision-making, which proposed a way to share data among researchers to rapidly explore the answers to Covid-19 related research questions, across the four nations, using cardiovascular disease as an exemplar.*
- *The Clinical Care for CVD in the COVID-19 Emergency is an initiative, supported by the BHF DSC, which analyses data from nine large hospitals across the UK to assess the impact of Covid-19 on NHS services for cardiovascular diseases, including an online visualisation tool to enable near real-time monitoring of trends. This allowed for rapid analysis of pre- versus post-pandemic trends using existing NHS data.*[7]

Case study 2: Contributions of the BHF/Alan Turing Institute Cardiovascular Data Science Awards to advancements in cardiovascular disease research

To date, we have supported 12 BHF/Alan Turing Institute Cardiovascular Data Science Awards, at a value of over £0.6 million. This is a joint funding scheme which promotes multi-disciplinary research to generate data science solutions to key cardiovascular problems in areas such as data access, privacy and anonymisation, machine learning, image analysis and modelling. The BHF has granted no-cost 6-month extensions to these projects to mitigate against Covid-19 disruption, which has delayed this research.

Examples of funded research projects include:

- *Using machine learning with a large dataset from all patients undergoing surgery in the UK to develop a new 'risk calculator' that identifies patients likely to survive open heart surgery.*[8]
- *Investigating a large number of possible diseases that occur after a heart attack by assessing millions of hospital records.*[9]
- *Improving understanding of genetic, blood related cardiovascular risk factors through algorithmic analysis of images of blood cells from 30,000 healthy people, collected for the University of Cambridge's COMPARE study.*[10]

2. Existing barriers to effective use and sharing of data

Lack of public trust

2.1 The impact of underestimating the importance of public trust was recently demonstrated by the General Practice Data for Planning and Research (GPDPR) programme, designed to

facilitate the extraction of GP patient data for research and planning.[11] Existing public concerns about health data sharing were compounded with confusion between the national data opt-out and the type 1 optout, as well as the opaqueness of communication about the programme. These factors caused more than 1 million people to opt out of the scheme (meaning that they did not want their patient data to be shared for purposes other than their own care), and the programme launch had to be paused pending it meeting specific criteria, including transparent and active engagement with patients and the public. High numbers of people opting out of such schemes, and the negative impact on public trust, represents a significant barrier to data sharing, and it could potentially reduce the richness of the dataset. There is evidence to suggest that some ethnic minority groups may be more likely to opt-out, and there is a risk that this could exacerbate health inequalities. Additionally, there is no centralised data on Type 1 opt-outs, so researchers are unaware of this potential data bias.[12] To avoid a repeat of this, is essential that any changes to the use of patient data are communicated clearly and transparently to the public.

2.2 The Ada Lovelace Institute concluded that the downfall of the GPDPR programme was its lack of public engagement ahead of roll out, and a reliance on a 'decide, announce and defend' decision-making model (i.e., when experts, policymakers and/or politicians announce a solution without considering the public's views, and instead use resources to defend against public backlash).[13] This approach is particularly unsuited to controversial issues like health data sharing, or initiatives where successful implementation is highly dependent on public buy-in. Such an approach can actively undermine the potential for data to be used in ways that benefit patients and the public.

2.3 Analysis from Understanding Patient data (UPD) shows that people are generally comfortable with anonymised data from medical records being used in research, provided there is a public benefit. Crucially, the more informed the public feels, the more likely they are to support the use of their data for health research purposes; however, many are uncomfortable with the idea of companies accessing their health data, and there are particular concerns about information being passed on for marketing or insurance purposes.[14]

2.4 There are many examples of patient data being non-consensually used for purposes outside of research. One of the most damaging examples of this was in 2018, when it came to light that a data sharing system between the Home Office and the NHS was being used to locate patients believed to be breaching immigration rules.[15]. Such breaches of trust can indirectly damage public perception of efforts to use health data to improve health outcomes. A 2017 survey of 1071 UK adults found that 44% of respondents did not trust Government organisations with their personal data. The main reasons for this were that people felt government organisations a) did not have control over their data, b) are not good at keeping data safe and secure and c) do not have individuals' best interests at heart.[16] Low public trust can be partly attributed to concerns about past data breaches in the public sector and data misuses in the private sector.[17] [18] [19]

2.5 To mitigate this, we support the recently added 'no surprises' Caldicott Principle – one of eight principles which are designed to ensure people's information is kept confidential and used appropriately.[20] This principle states that a range of steps should be taken to ensure no surprises for patients and service users, so they can have clear expectations about how and why their confidential information is used, and what choices they have about this. The use of health data must reflect public expectations and should also take into account patient benefit.

2.6 Therefore, we call on Government to ensure that the purpose and process of sharing patient data is communicated transparently, especially around issues we know are of public concern, such as increased access to health data by private companies and Government departments.

2.7 In addition, we recommend that any future reforms or Government strategies that concern the way that patient data are shared include a public engagement plan that reaches an audience that accurately reflects the demographics of the population. This engagement should address the following key questions:

- How will patient data be shared, with whom and for what purpose?
- What are the measures in place to protect against misuse of patient data, and the penalties for breaches of data laws?
- Where can the public access trusted, clear information about the use of their health data and how can they raise any questions they might have?

2.8 Facilitating an open conversation with the public on this issue would help to build trust and public understanding on this topic. It would also allow Government to monitor public attitudes about health data and quickly address concerns that come to light.

2.9 There are several good practice examples of public engagement focusing on the way health data are collected and used. One of the most cited is the OneLondon Citizens' Summit, an event that brought together 100 Londoners to debate and deliberate some of the complex issues around uses of health and care data.[21] The Summit identified Londoners' expectations as to: how their health and care data should be used, who should have access to it, and for what purpose. These expectations were delivered as recommendations in the OneLondon report to a panel of local and national system leaders, politicians and policymakers. This was an example of deliberative public engagement: a unique approach which empowered Londoners to have their say and to inform policy and practice in a way that builds legitimacy and trust.

Case Study 3: Working with patients through the BHF-CRUK Patient Data Panel

The BHF, in collaboration with Cancer Research UK, co-manages a patient data panel consisting of 12 patients. These patients not only share their lived experience with cardiovascular disease and cancer but have also developed expertise in understanding the complicated landscape around health data – both the challenges and the opportunities it provides. The panel allows the two charities to understand the patient perspective when we develop policy initiatives so that no crucial viewpoints are missed. It also speaks to the BHF's wider strategic commitment to work with patients and the public for better health and care.[22]

It was initially formed in response to the NHS opt-out, which allows patients to stop their health data from being used by the NHS for healthcare planning purposes or research – the panel was instrumental in responding to related NHS enquiries. They were also important in guiding early thinking in the establishment of the BHF DSC. The panel has been recognised more widely for their expertise and has been consulted by NHS X, NHS England and Imperial College to ensure early input from patients.

Limited researcher access and training

2.10 Researchers must be able to access data to conduct research for public benefit in a safe way. Trusted Research Environments (TREs) are secure computing environments that allow researchers to do this, providing approved researchers from trusted organisations with timely and secure access to health and care data. Researchers are given access to their approved data (in accordance with their Data Sharing Agreements), enabling them to collaborate, link data, share code and results within the same research projects, without data leaving the TRE. TREs played a key role during the Covid-19 pandemic, enabling researchers from all over the UK to work together on projects, remotely accessing and analysing the same data, as highlighted in Case Study 1.

2.11 However, a current limitation of TREs, and therefore a barrier to data sharing, is the lack of common governance and technical implementation standards, leading to a divergence between TREs. Providing standards for the technical design and deployment of TREs is essential to ensuring interoperability and providing opportunities for federation across environments. Being clear on where standardisation is needed and where there is room for flexibility will also allow for innovation between different TREs and ensure that collectively, they meet a range of user needs including a reduction in the number of different access requirements to different environments.

2.12 To achieve this, we support the ambitions that NHSX has outlined with regards to TREs:

- A minimum technical specification for TREs which covers core areas such as interoperability, cybersecurity and use of privacy enhancing technologies.
- TRE standards and policy should be developed alongside creating new governance standards and enhancing any relevant existing ones. Policy and best practice guidance for TREs should include when their use will be mandatory and any legitimate exceptions.
- An accreditation framework should be put in place which details the specifications and standards that TREs must adhere to, and how adherence will be assessed and monitored.[23]

2.13 In addition to standardising the way that TREs operate, the high levels of data security that TREs provide should be explained to the public in clear language. This will help to reassure the public that appropriate measures are in place to ensure that their health data are handled in a secure and highly regulated way.

2.14 We share Cancer Research UK's (CRUK) concern around misconceptions that regulations for protecting data and research governance are hindering innovation. Data privacy is not a barrier for research; rather, it is a fundamental control element that encourages appropriate discipline in the application for, and use of, personal data. It can cause difficulties, however, when processes surrounding data handling and collection are not communicated clearly with researchers. Regular insight from our cardiovascular research community highlights continued confusion and lack of understanding as how best to navigate what are perceived to be complex processes. Clearer guidance and accessible learning resources for researchers on how to safely handle sensitive data would help to build confidence and would also improve data security.

2.15 We recommend that data regulations are communicated clearly with the research community. This should include clarity around applicable regulation, the sharing of practical resources, and strong case studies. We recommend improved guidance from

regulators and greater focus on education and upskilling of data custodians, researchers, and host organisations on data protection legislation.

Inefficient Government and NHS technology

2.16 Persistent inefficiencies of 'outdated' Government and health service technology systems are holding back innovation in data sharing. Further to this, there is insufficient integration between different electronic health records and poor interoperability of data systems between NHS trusts, practices and hospitals, and a lack of centralised points of data access for real time clinical data.[24] [25] This has significant impact on patients with heart and circulatory diseases, e.g., heart failure. A BHF report, 'Heart Failure: A Blueprint for Change', found that services across the UK described a lack of reliable, comprehensive, whole pathway data as a significant barrier to improving heart failure services. People working in heart failure services also expressed a lack of confidence that the heart failure data that is routinely available is accurate. This means that data is not informing and driving improvement in ongoing heart failure care after hospital admission. Without improving the collection and sharing of data, it is almost impossible for the system to have a clear picture of heart failure care across the country, and the impact of changes to models of care.[26]

2.17 We recommend that Government should invest in improvements in the quality, linkage, and interoperability of NHS data systems so research that uses health data is as efficient as possible.

2.18 The upcoming merger of NHS Digital into NHS England has potential to disrupt data sharing. To transfer datasets from Public Health England to NHS Digital in 2021, data access requests were put on pause for several months.[27] This year, there will be a similar upheaval, as NHS Digital merges into NHS England. There is not yet any public information on the details of the move, and it has potential to cause delays and confusion; however, the change also offers an opportunity to streamline and improve data access, which we urge the Government and NHS England to take.

2.19 Every effort must be taken to identify and protect against unintended consequences of the merger of NHS Digital into NHS England, and we would urge the Committee to prioritise this issue in its inquiry.

Missing data

2.20 Data sharing is essential for research innovation; however, the utility of data is only as good as its quality, including how representative it is of populations. A 2021 UK survey of 2102 people for the Information Commissioner's Office found that Black and minority ethnic respondents are significantly more likely to have already requested to be forgotten, have inaccurate personal information rectified or completed, and moved their information from one provider to another than white respondents.[28] There are many potential reasons for this, including historic breaches of trust mentioned earlier in this section.

2.21 Ensuring that shared data are representative of minoritised groups e.g., women and people from a Black and minority ethnic background, and that these people do not opt out of data sharing, is essential for tackling health and research inequalities and for levelling up healthcare. Establishing and maintaining public trust is key to achieving this. Datasets that are representative of the whole population helps to develop health interventions that are effective for all members of society. More broadly, technologies and policy interventions developed using

datasets that lack information about certain groups will, at best, not serve whole populations and, at worst, exacerbate inequalities.[29]

2.22 Making available health data representative of the population should be a greater focus in Government's communications and considerations for any legislative change.

3. The extent to which data issues are appropriately addressed by the Government's Strategies and consultations

Health and Social Care Data Strategy

3.1 We welcomed the ambition of the Government's draft strategy on data use in health and care, and in particular the commitment to improve researchers' access to data. Specifically, we were pleased to see ambitions around improving the completeness, quality and interoperability of datasets, ensuring good governance around data collection and management, and upskilling healthcare professionals and data custodians. We were also encouraged to see an emphasis on enabling researchers to access 'data that allows targeting research at those who are most at risk for poorer health, that is to investigate health inequality' – there are currently huge gaps in health data available for research and addressing these should be a priority.

3.2 It was disappointing not to see a more significant section on public engagement in the strategy, but we expect that this will be addressed in the final document. We support the Government's plan to hold regional public engagement events to find out how people would like to use and access their data, as well as working with research charities to establish good practice with regards to transparency and meaningful public involvement in data use. It is important that any public engagement activity is planned so that the audience accurately reflects the demographics of the population.

Data: a new direction

3.3 The BHF submitted a formal response to this Department for Digital, Culture, Media & Sport (DCMS) consultation.[30] This section will summarise our key recommendations that are not covered elsewhere in this response.

3.4 As a major medical research funder and a large charity, we strongly advocate for Government to ensure maintenance of the UK-EU data adequacy agreement, not only in the short-term but when adequacy is being considered for renewal in 2025. To do so will require a firm commitment to making sure that UK data laws do not diverge significantly from the EU's.

3.5 After more than a year of talks, the EU-UK GDPR data adequacy agreement was approved in June 2021 by the European Commission. This was a positive result as data transfers are essential for running UK-EU clinical trials, where researchers need to routinely send patient data and test results across international borders. The BHF also welcomed the news, as around 19% of our collaborators are based in the EU.[31]

3.6 The adequacy agreement is valid for four years, but includes an obligation on the EU Commission to monitor developments with UK data protection laws and revise or revoke the adequacy agreement if the UK system is seen to have diverged from the required standard. If adequacy is not maintained, UK-based researchers involved in UK-EU clinical trials, for example, would have to arrange alternative mechanisms to share personal data with their EU-based partners.

3.7 These alternative data transfer mechanisms would increase the cost of research through added complexity and legal fees, which could deter EU research studies from involving UK-based research and, in turn, reduce patients' access to life-saving innovations. Investigations by the University College London European Institute and British think-tank, the New Economics Foundation, found that it could cost medical-research institutions £50,000 to £100,000 to negotiate a data-sharing agreement with an institution in another country.[32] As stressed by many in the research sector, it is in the best interests of all patients that UK and EU researchers be allowed to continue exchanging data - an essential part of collaborative research.[33]

3.8 We strongly recommend that Government publishes details on the steps being taken to ensure that any legislative changes do not threaten the EU data adequacy agreement.

3.9 Algorithms and artificial intelligence are powerful tools which can help to make great progress in health research and patient outcomes. However, if unregulated, biases in algorithmic decision-making can reflect wider inequalities in society, which in turn exacerbates existing health and research inequalities. For example, a 2019 study by Obermeyer et al[34] found that use of a widely used commercial prediction algorithm for cardiovascular care resulted in significant racial bias in predicting outcomes. Specifically, the algorithm identified White patients to have higher risk scores and were more often selected to receive additional care than Black patients who were equally as sick.

3.10 To address this serious issue, we support the establishment of mandatory transparency obligation for UK public sector organisations that use algorithms to make decisions affecting people's lives, as advised by the Centre for Data Ethics and Innovation (CDEI).[35]

3.11 Compulsory transparency reporting on the use of algorithms by the public sector would be a positive step towards improving public trust. The current accountability and decision-making processes surrounding algorithms are opaque, which can lead to mistrust from the public as well as the spread of misinformation.

3.12 We strongly advocate for compulsory transparency reporting on the use of algorithms in decision-making for public authorities, government departments and government contractors using public data. Further, this should include standards around reporting in which ensure that it is communicated in clear language. We also recommend that there should be ongoing testing, and reporting, for bias in these algorithms to avoid perpetuating cross-cutting inequalities.

3.13 Transparency will not only help to build public trust in algorithms and the use of their health data, but it could also help to increase public understanding and literacy in this area and evidence a commitment to real engagement between technicians and the public to help develop algorithms that more accurately reflect the real world, without bias.

4. The extent to which appropriate safeguards and privacy are applied in the usage and sharing of individuals' data

4.1 Health data are highly sensitive; therefore, it is paramount that appropriate safeguards and privacy are put in place to protect it and that these are communicated clearly to the public. Many safeguards are already in place, and we join The Association of Medical Research Charities (AMRC) in supporting the continuation of these measures:

- Clear red lines on certain uses of health data e.g., data cannot be used for marketing or insurance purposes.

- Data should be anonymised as much as possible, and researchers must only be given access to the minimum amount of information needed to answer their research question.
- Strict controls about how companies can use personal health data, including having a legal basis to access identifiable data, data access committees reviewing all applications, and having to sign contracts setting out what they can and cannot do with the data.
- Robust penalties if data are misused.
- Audit processes to check who is accessing data.

4.2 Further, all work on patient data must be carried out within an environment that can be shown to be secure and accessible only to permitted researchers. As highlighted in section 2, TREs are a useful tool for improving data security and building public trust. If their safety measures are communicated clearly to the public, it could help to develop trust and understanding around how data are being used.

4.3 Understanding Patient Data (UPD) describes a 'spectrum of identifiability' for data: personally identifiable, de-personalised and anonymous.[36] Different controls are needed at different points along this spectrum depending on the risk of re-identification. The use of de-personalised data presents a communications challenge for maintaining trust in data-sharing. The public tend to be happier with sharing data when it is done so anonymously, but in practice this is hard to guarantee.[37]

4.4 While safeguards can help to minimise privacy risks, using data will never be risk free. Even when direct identifiers are removed, individual data can potentially be re-identified and linked to an individual. The public recognise that data is never completely safe, but more needs to be done to communicate that the risks of re-identification are very small with the appropriate safeguards in place; this will ensure people are making a truly informed choice about data sharing and use.

Case study 4: Safeguards and privacy measures used by the CVD-COVID-UK project

The CVD-COVID-UK consortium is a BHF Data Science Centre project that aims to understand the relationship between Covid-19 and cardiovascular diseases through analyses of de-identified, linked, nationally collated healthcare datasets across the UK.

As a collection of electronic health records and TREs, the consortium published a clear and concise statement that explained how data would be collected, stored, managed, and protected.[38] The statement also outlined the types of patient data that would be accessed and how they would be used.

Examples of the safeguards and privacy controls used within the CVD-COVID-UK consortium include:

- *Ensuring consortium members agree to adhering to the 'Five Safes': a framework for helping make decisions about making effective use of data which is confidential or sensitive.[39] This framework breaks down the decisions surrounding data access and use into five related but separate dimensions, including 'safe projects' (is this use of the data appropriate, lawful, ethical and sensible?) and 'safe people' (can users be trusted and are they sufficiently trained to data it in an appropriate manner?)*

- *Before any researcher in the consortium can access the data to perform research analyses, it is first de-identified. This means that any information that could directly identify an individual (e.g., name or NHS number) is removed.*
- *Individual level patient data never leaves the TRE. Access to the data is restricted to approved, named researchers with certified training in the safe use of health-related data for research. These approved researchers access the data held in each environment remotely through a highly secure authentication system.*
- *Before researchers can share research results outside the TRE, the results of analyses conducted within the environment are first checked to ensure that no individual level or potentially identifiable data are included.*
- *The research is approved by the NHS Research Ethics Committee and is reviewed and approved by an independent panel within each devolved nation.*

For further information, please contact Monica Dahiya via dahiyam@bhf.org.uk

---

[1] Health Data UK, *BHF Data Science Centre*, accessed January 2022
[2] The British Heart Foundation, *BHF-Turing Cardiovascular Data Science Awards*, accessed January 2022
[3] The British Heart Foundation, *Cardiovascular Catalyst Awards*, accessed January 2022
[4] The British Heart Foundation, *Patient data: what is it and why is it important for research?*, 2016
[5] Dale et al, *The adverse impact of COVID-19 pandemic on cardiovascular disease prevention and management in England, Scotland and Wales: A population-scale analysis of trends in medication data*, 2022
[6] UK health research data alliance, *A national health data research capability to support COVID-19 research questions*, 2020
[7] Ball et al, *Monitoring indirect impact of COVID-19 pandemic on services for cardiovascular diseases in the UK*, 2020
[8] The Alan Turing Institute, *Using machine learning to identify patients likely to survive open heart surgery*, accessed January 2022
[9] The Alan Turing Institute, *Disease trajectories following a heart attack*, accessed January 2022
[10] The Alan Turing Institute, *Blood related risk factors for cardiovascular disease*, accessed January 2022
[11] NHS Digital, *General Practice Data for Planning and Research (GPDPR)*, 2021
[12] Association of Medical Research Charites, *GP Data for Planning and Research (GPDPR) – joint statement from medical research charities*, 2021
[13] Ada Lovelace Institute, *Turning distrust in data sharing into 'engage, deliberate, decide'*, 2021
[14] Understanding Patient Data, *How do people feel about the use of data?*, Accessed January 2022
[15] The Guardian, *Home Office scraps scheme that used NHS data to track migrants*, 2018
[16] Deloitte, *The State of the State 2017-18*, 2018
[17] Reform, *Sharing the benefits How to use data effectively in the public sector*, 2018
[18] Nesta, *Dialogues about Data: Building trust and unlocking the value of citizens' health and care data*, 2021
[19] The Scottish Government, *Public Acceptability of Data Sharing Between the Public, Private and Third Sectors for Research Purposes*, 2013
[20] The UK Caldicott Guardian Council, *The Caldicott Principles*, accessed January 2022
[21] OneLondon, *Public deliberation in the use of health and care data*, 2020
[22] The British Heart Foundation, *We work with patients and the public for better health and care*, accessed January 2022
[23] NHSX, *Joining up the dots: driving innovation, research and planning through Trusted Research Environments*, 2021
[24] National Audit Office, *Digital transformation in the NHS*, 2020
[25] The Association of the British Pharmaceutical Industry, *Unlocking the promise of UK health data*, 2020
[26] The British Heart Foundation, *Heart Failure: a blueprint for change*, 2020
[27] NHS Digital, *Data Access Request Service (DARS)*, accessed January 2020

[28] Information Commissioner's Office, *Information Rights Strategic Plan: Trust and Confidence*, 2021
[29] Nature, *A fairer way forward for AI in health care*, 2019
[30] The British Heart Foundation, *BHF response to DCMS 'Data: a new direction' consultation*, 2021
[31] The British Heart Foundation, Research Evaluation Report: Outputs and impact metrics of BHF funded research 2020-21, 2021
[32] New Economics Foundation, *The cost of data inadequacy*, 2020
[33] Nature Index, *Looming deadline on post-Brexit data agreement has researchers worried*, 2021
[34] Obermeyer et al. *Dissecting racial bias in an algorithm used to manage the health of populations*, 2019
[35] Centre for Data Ethics and Innovation, *CDEI publishes review into bias in algorithmic decision-making*, 2020
[36] Understanding Patient Data, *Identifiability Demystified*, accessed January 2022
[37] The King's Fund, *Using data in the NHS: the implications of the opt-out and GDPR*, 2018
[38] British Heart Foundation Data Science Centre, CVD-COVID-UK programme privacy statement, accessed January 2022
[39] The Five Safes, *The Five Safes*, accessed January 2022