Written evidence submitted by Trilateral Research (DDA0022)

With contributions from Dr. Joshua Hughes, Vangjel Gjorgjiev, Dr. Mistale Taylor, Nicole Santiago, Panagiota Kourti, Dr. Katrina Petersen, Dr. Filippo Marchetti

Trilateral Research is a British and Irish SME with a long history in providing ethics and data protection advice in research projects, and to commercial clients, as well as now offering data science services. We have significant experience in understanding and promoting privacy in the digital era, including in the health domain.[1] Our response to the Committee's call for evidence draws on experiences from both the research and commercial environments, which gives us a unique platform from which to contribute.

Summary

Trilateral Research recognises the clear benefits of data sharing, particularly in the health context for data-driven responses to diseases (e.g., the COVID-19 pandemic), medical research and generating value. Sharing medical data needs to follow the standards of data ethics and medical ethics, especially the principles of autonomy, beneficence, nonmaleficence, and justice. Where data sharing takes place, it is important that the correct technical and organisational safeguards are implemented to ensure that patient privacy is respected, and no more private information is shared than is necessary for the defined purpose, and that there are appropriate legal agreements in place so that the data processing relationships, and the responsibilities of each party, where data is being shared are clear. Often, practitioners and decision-makers need support to implement the necessary policies and safeguards to protect privacy. With respect to the UK Government's data strategy, Trilateral supports the introduction of a legal basis for research, recommendations to give greater clarity on anonymisation, and inclusion of robust technical and organisational measures to safeguard data-subjects.

In our submission, we focus on the following bullet points from the call for evidence:

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, and the existing barriers to such data sharing;
2. the ethics underpinning the use and sharing of individuals' data in health and care contexts;
3. the extent to which appropriate safeguards and privacy are applied in the usage and sharing of individuals' data; and
4. the extent to which data issues are appropriately addressed by the Government's National Data Strategy, its draft strategy, *Data saves lives: reshaping health and social care with data*, and its consultation *Data: a new direction*.

Each numbered section is used to frame our responses to the above points. With reference to our own experience in the sector, we focus on the benefits of data sharing in the health context.

1. Response to: the potential benefits, including to research, to effectively use and share data between and across Government, other public bodies, research institutions and commercial organisations, and the existing barriers to such data sharing.

[1] More information on the work conducted and services offered by Trilateral Research can be found on our website: www.trilateralresearch.com.

Data sharing can bring great insight to decision-making at all levels from front-line practice to strategy and policy. Enhancing collaboration and data sharing with insight teams across the public and private sectors can enable greater awareness of diseases and our response to them, discovery of previously uncovered insights into medical data, and recognition of value that can be gained from existing datasets. This can help shape and optimise an integrated response to improving outcomes in a variety of medical fields. Data sharing is especially beneficial when it is done in an ethical, secure and legally compliant way.

For example, benefits from sharing health data across governmental agencies, research institutions and commercial organisations can enable faster recognition of medical conditions for faster treatment, more robust research across larger datasets for a more holistic view of health and social issues, and a more effective way of tailoring treatment to an individual to improve their health outcomes.

Data sharing to develop data-driven responses to the COVID-19 pandemic

During the ongoing COVID-19 pandemic, data-sharing has been crucial in developing holistic, data-driven responses to the myriad problems it has presented. For instance, it has allowed new variants to be recognised and tracked within and between different countries. The World Health Organisation has the capability for novel coronavirus infections to be reported and tracked on the international level.[2] Optimising data sharing between medical facilities within, and between, countries could enable swifter collaboration for detecting, analysing and monitoring infectious diseases. As has been evident during the COVID-19 pandemic, the speed at which novel diseases and variants can be understood impacts on when appropriate safeguards (and potentially restrictions) can be implemented. Thus, increasing the speed of data analysis could enable safeguards to be implemented earlier, thereby reducing transmission rates sooner, and contributing to the overall health and wellbeing of society.

Another area where data-driven responses could improve the current situation is the gap between health and social care. The pandemic has shown that data sharing between health and social care institutions is not as smooth or effective as it could be. Optimising data sharing in this context can deepen the understanding of the implications of various approaches to individual and public health. Going further, and potentially beyond the pandemic, this approach could also incorporate data on the impact of medical, social, environmental/pollution, or economic policies and their effects on health.

Research data

In terms of research, data collected by hospitals could be used for valuable research. However, the catchment area of one hospital might be relatively small, and so the data they collect might not be large enough for robust conclusions that can be extrapolated on the national level. Yet, sharing of this data across many different hospitals and trusts would allow that research to take place. Another crucial issue is that some medical conditions, for example, might be more, or less, prevalent in one area than another, and so sharing data on a higher level can balance out some of these traditional data gathering biases. Thus, there are potential scientific benefits and research value that can be gained from sharing data. Higher-quality research can lead to better understanding and, subsequently, improved policy choices.

Data in the technology sector

The vast wealth that has been generated in the tech sector indicates the monetary benefits that data can offer; indeed, data is often referred to as the 'new oil'.[3] The significant amounts of high-quality data generated in the NHS, and from other medical providers in the UK, represent potential value to the healthcare sector and patients. For example, the controversial collaboration between Google's Deepmind and Moorfields Eye Hospital demonstrates the possible financial value that can be

[2] WHO, 'Global Surveillance for human infection with novel coronavirus (2019-nCoV)', 31 January 2020 *available at* https://apps.who.int/iris/bitstream/handle/10665/330857/WHO-2019-nCoV-SurveillanceGuidance-2020.3-eng.pdf.

[3] See, e.g., Bhageshpur, Kiran, 'Data Is The New Oil - And That's A Good Thing', *Forbes*, 15 November 2019.

generated for the technology sector where they have access to health data, and the conceivably beneficial medical tools that could be created using such data.[4] However, the controversy and social concern about this collaboration also demonstrates that it is important for health providers to share data in a trustworthy, ethical, and lawful way that also takes account of their stakeholders and shows they are acting responsibly. Indeed, the extraction of financial value by the private sector from publicly gathered health data should not come at the expense of patient needs, patient privacy or patient care. Below, we offer some points for consideration regarding health data sharing with a focus on the ethical and legal perspectives.

[4] BBC News,' Why Google DeepMind wants your medical records', 19 July 2016.

2. Response to: the ethics underpinning the use and sharing of individuals' data in health and care contexts.

Data sharing in the health sphere should not just be governed by data protection legislation, ethics should also be included as a set of rules for distinguishing between acceptable and unacceptable behaviour. In this context, the use and sharing of data should adhere to both data ethics and medical ethics.

Data ethics

A key principle of data ethics is informed consent from the individual whose data is being collected, processed, shared, and stored. In the realm of healthcare, such individuals can appear in the position of a patient, clinician, caregiver, etc. In order to obtain the individual's informed consent, they need to be provided with all the relevant information pertaining to the use of their data. They also need to understand the potential risks and benefits of data sharing and have the right to object to processing of their data in a particular way, without impacts to their clinical care.

Medical ethics

The basis of medical ethics consists of four crucial principles that need to be followed when processing personal data in the healthcare sphere: autonomy, beneficence, nonmaleficence and justice. Medical research with humans must follow the informed consent procedures and other requirements outlined in the Declaration of Helsinki and the Oviedo Bioethics Convention.[5] Below, we discuss data- sharing in terms of the four principles of medical ethics with considerations from data ethics.

Autonomy

The principle of autonomy refers to the individual's ability to freely make decisions on their own. When it comes to health, this translates into allowing and enabling the patient to make autonomous decisions about their treatments and care; and allowing and enabling the clinician to provide the best healthcare treatment possible. It should be noted that any attempt to persuade a patient into making a choice or undermining a clinician to provide anything but the best healthcare treatment is a violation of the principle of autonomy. Linking this with data ethics, it is important that stakeholders, especially patients, are not pressurised into allowing their data to be shared as a condition of access to other, potentially better, treatments.

Beneficence

Beneficence represents the belief that health treatment needs to be beneficial to the patient. To achieve this, healthcare providers have the obligation to be of benefit to the patient and take positive steps to prevent and remove harm from the patient. The healthcare decisions made for the patient must always be accompanied with the intent of providing the most benefit to the patient. To this extent, any medical data must be well understood by the healthcare provider and have the capacity to adequately process such data. In the context of medical data sharing, providing access to more people in an appropriate way might enable additional insights to be gained that will benefit the health of patients.

Nonmaleficence

The principle of nonmaleficence is closely tied to the principle of 'do no harm'. It represents the notion of not intentionally creating harm or injury to the patient, either through acts of commission or omission. The mechanics of this principle require the healthcare provider to make a consideration of whether the level of harm is proportionate to the good it might achieve. Therefore, when using and sharing medical data, clinicians must balance the potential harm to the expected benefit. Such actions should always be accompanied by appropriate safeguards and respectful of patient privacy. However,

[5] World Medical Association Declaration of Helsinki, 'Ethical Principles for Medical Research Involving Human Subjects' as amended in 2013; Council of Europe, Convention on Human Rights and Biomedicine, 1997.

if safeguards are not sufficient to mitigate risks of harm, it should be made clear that sharing of medical data is not an appropriate or ethical choice.

Justice

Justice is of vital importance in the medical sphere. This has been highlighted by the COVID-19 pandemic more than ever. The principle finds best application for healthcare in resource allocation and determining priorities. Healthcare providers must rely on a system of fair distribution based on the circumstances at hand, especially where supplies are limited. Therefore, clinicians need to adequately process medical data to deliver fair results that are in line with this principle. Data sharing could facilitate enhanced analysis to provide suggestions on how best to provide an equitable distribution of medical supplies for patients in need.

Below, we provide two case studies from our research project work.

Case example 1: early detection of melanoma (iToBoS project)*

Melanoma (skin cancer) is one of the most common and aggressive cancers. Fortunately, melanoma is curable cancer that can be treated effectively, if detected early. Rapid diagnosis is essential to ensure that treatment is undertaken before local and advanced spreading occurs.

In the last three years, AI systems for the identification of melanomas have become increasingly prevalent. However, most AI systems have yet to differentiate between normal tissue and abnormal skin lesions using images alone. There are also significant bias issues with respect to the skin tones that are used for training such systems. This means that without integrated clinical data, systems continue to underperform compared to the average dermatologist.

Developing AI-based solutions for improving dermatologists' diagnostic accuracy

Responding to this challenge, iToBoS aims to develop a patient-centred AI diagnostic platform for the early detection of melanoma. Specifically, it aims to train an AI system with integrated information from various sources, ranging from dermoscopic images, medical records, family history and genomics data, to increase the precision of clinical decisions in the diagnosis of skin cancer.

Ethical, Privacy and Social Impact Assessment to ensure sustainable innovation

In this project, Trilateral is conducting a comprehensive Privacy Impact Assessment of the iToBoS technologies to ensure the secure handling of patient data complying with existing regulations at both European and national levels.

Trilateral's work also aims to maximise benefits and minimise the risks of AI technology with consideration of ethical and social values. In collaboration with all involved stakeholders, we examine all social and ethical challenges and use these to improve clinician and patient understanding of the role AI can play in healthcare and to assist in the adoption of innovative technologies in the dermatological context.

Trilateral's effort aims at increasing the sustainability of the project's results by producing guidelines for healthcare professionals, patients, care- givers and -receivers and related EU projects.

Ensuring data ethics and algorithmic transparency

Trilateral is developing novel techniques for achieving algorithmic transparency and deal with issues such as algorithmic bias and interpretability. We are conducting an analysis of the AI models' performance, pairing data ethics with AI evaluation assessments of the iToBoS tools. This includes an examination of the algorithms, system integration, and data management systems and their outputs.

*This project has received funding from the EU's Horizon 2020 research and innovation programme under grant agreement No.965221.

Case example 2: Supporting data-driven decision making for policymakers and the healthcare sector (COVINFORM project)*

COVINFORM assesses the social, economic and political impacts of the COVID-19 crisis focusing on government, public health, and citizen responses and the role of information and communication during all stages of the pandemic. The project also examines how European governments define and address vulnerability to COVID-19, if at all. COVINFORM partners are developing solutions, guidelines and recommendations to mitigate health inequalities and improve the wellbeing of the most vulnerable individuals.

COVINFORM employs data from:

- a cross-analysis of pre-existing studies on national, regional and local levels to map the responses and assess the impact of the pandemic;
- relevant dimensions of health, socioeconomic, political and community vulnerabilities.

Trilateral ensures this data is obtained, processed and used in an ethically-sound manner. The findings will be used to create an online toolkit for governments, the civil service, healthcare sector and researchers.

Social and ethical risk assessment models

As a part of the COVINFORM project, Trilateral is developing a risk assessment model to evaluate the response and impact of the pandemic on national, regional and local levels. The model will encompass desk-based research identifying relevant studies and data on responses to assess pandemic planning and preparedness processes. It will look at different governmental responses and their impact: from lockdown procedures, school closures, advising to wear masks to furlough schemes and economic stimuli. The analysis focuses on various regulatory options to support the ethical and responsible development of pandemic preparedness plans.

*This project has received funding from the EU's Horizon 2020 research and innovation programme under grant agreement No. 101016247.

3. Response to: the extent to which appropriate safeguards and privacy are applied in the usage and sharing of individuals' data.

Moving to the legal impacts of data sharing, it is important to note that incorrect sharing of personal data presents significant data protection and privacy risks. The right to personal data protection is an important corollary to the right to privacy, especially when it comes to personal data processing online, digitised personal data and sensitive personal data, such as health data. An example actor in this sphere is the NHS, as a key player in data sharing in a public sector context. As the NHS is a public authority, it is subject to human rights obligations – such as those incorporated into the UK Human Rights Act - and must also ensure that contractors assist in meeting those standards.

Article 8 of the European Convention on Human Rights, to which the UK is a party by virtue of being in the Council of Europe, is broadly about the right to privacy and reads:

> "Everyone has the right to respect for his private and family life, his home and his correspondence."

Safeguarding personal data: technical and organisational measures

An important way to respect this obligation is for public authorities to abide by data protection legislation and implement technical or organisational safeguarding measures to protect personal data. Biological material itself is not considered personal data, but data relating to it, or otherwise relating to identifiable patients is personal data.[6] It is often suggested that personal data used for research should simply by anonymised. Anonymous data is where the data is treated in such a way that the person to whom they relate can no longer be identified. However, the UK GDPR sets a high standard for anonymisation,[7] and there would be significant data protection risks if personal data were shared under the wrong impression that the data was anonymised. Therefore, pseudonymisation, where direct identifiers of the data-subject are removed, can be a better option. Pseudonymised data is still personal data,[8] but the regulatory burden is lower as pseudonymisation mitigates many data protection risks. Pseudonymising personal data also reduces the risks to a data-subject in case of data breaches or leaks, and so supports the moral obligation to respect patient privacy. Both anonymisation and pseudonymisation are approaches to data minimisation, explored below, which can be a technical safeguard.

Data minimisation

Patients' privacy is incredibly important, and so an appropriate approach to data minimisation is needed to ensure that enough data is revealed for the purpose for which the data is shared to be fulfilled, whilst preventing an unnecessary and disproportionate intrusion into medical records. For example, neither a researcher collecting mortality data for a particular disease, nor an IT contractor developing new ways of storing medical data need access to a patient's full medical record. Most people would consider any unnecessary access to their medical records as a deeply personal intrusion into their privacy. Therefore, the approach needed for appropriate data minimisation should be considered on a case-by-case basis. Yet, where datasets are very large, a more generalised approach that still takes into account the ethical and legal issues might need to be considered.

Data minimisation does not need to be a barrier to fulfilling the purposes of data processing. For example, the researcher might only need minimal information in aggregated statistics, which could be anonymised or strongly pseudonymised depending on what data is needed for their purposes. The IT contractor might not need access to a real medical record for their work, a fictitious example might suffice or privacy-preserving approaches where personal data remains with the data controller and

[6] Art. 4(1), UK GDPR.
[7] Recital 26, UK GDPR.
[8] Art. 4(5), UK GDPR.

results of testing new IT systems are reported back to the contractor could be examined. This also links with the obligation under Article 25 of the UK GDPR to implement data protection by design and default, meaning that business processes should be designed with data protection in mind so that compliance can be assured and privacy respected.

Sharing personal data: the relationship between data controllers and data processors

When personal data is shared, the data processing relationships between the parties providing and receiving data are clear. Whether the relationships are controller-processor, joint controllers or separate data controllers, different relationships need to be underpinned by appropriate legal agreements. These are, respectively: a data processing agreement, a joint controllership agreement or a data sharing agreement. These agreements provide organisational safeguards by determining accountability for what happens to the data and, in some cases, placing legal limitations on what may be done with the shared data. These agreements can be used to ensure that the sharing and processing of data is done within ethically and socially acceptable limits set by the parties. Development of these agreements can be complex, and so consultation with experts might be needed.

Ongoing protection

All relationships evolve, and the data sharing parties might develop or change their purposes for processing data. The parties should treat data sharing as an ongoing endeavour that needs regular attention and discussion on the terms to ensure that the data processing relationships are still appropriate for the purposes. When needed, agreements should be modified or replaced so that the legal and ethical standards of data sharing are current. Through our research experience we have found examples where decision-makers are hesitant to share data across border, even where the correct safeguards are in place, due to a fear that they will violate data protection legislation. Therefore, engagement with experts is crucial so that concerns can be dealt with to reassure decision-makers.

Case example 3: Smart support platform for pandemic prediction and management (STAMINA project)*

The STAMINA project is developing a decision support tool and policy guidelines to better equip pandemic crises management practitioners at national and regional levels to anticipate and respond to public health crises.

Ethical, legal, social and data protection impact assessments to ensure sustainable innovation

Trilateral manages the research ethics and data protection within the project. Trilateral considers the extent to which the development of the STAMINA system could pose ethical risks, and what measures are needed to mitigate unwanted negative impacts. Trilateral is also responsible for the ethical and data protection aspects of the project's Data Management Plan, to ensure the project complies with GDPR standards.

Cross-border data sharing challenges

STAMINA partners have sometimes shown apprehension around which data protection safeguards are applicable for health data, which has severely limited what data can be shared between partners for research purposes and across borders (organisational, regional and national). This has hampered the ability to explore what kinds of data would be beneficial for joint decision-making, and what tools can support such sharing and decisions. Practitioners and decision-makers need support in implementing data protection policies that balance the benefits of data sharing for public health with the risk to individuals from that sharing.

*This project has received funding from the EU's Horizon 2020 research and innovation programme under grant agreement No. 883441.

4. Response to: the extent to which data issues are appropriately addressed by the Government's National Data Strategy, its draft strategy, *Data saves lives: reshaping health and social care with data*, and its consultation *Data: a new direction*.

Trilateral considers favourably the establishment of clear and separate provisions related to the processing for research purposes, as highlighted in previous consultations in the context of the Government's National Data Strategy. Specifically, we welcome the idea of providing greater clarity regarding the appropriate lawful basis for research, the consolidation and bringing together of research specific provisions and the re-use of personal data for research purposes. However, Trilateral emphasises the need for enhanced transparency through privacy policies and procedures, the significance of valid consent where necessary, the facilitation of data subjects' rights and the implementation of robust technical and organisational measures especially in case of special categories of personal data.

In relation to data minimisation and anonymisation, Trilateral appreciates the recommendations to provide greater clarity on the determination of anonymisation test and the relative approach to be implemented but does not disregard the risks that the relative approach entails for the re-identification of data subjects' personal data.

Innovative data sharing solutions could prove beneficial in the health sphere. The UK Government should capture the need for implementation of robust technical and organisational measures offering enhanced protection to the personal data subject to processing and especially to special categories of personal data.

Conclusion

This contribution demonstrates how effective data sharing between government and industry can be beneficial. Whilst there are, indeed, risks associated with such data sharing, if these are properly considered and mitigated with the use of Ethical, Privacy and Data Protection Impact Assessments; the careful consideration of ethics and human rights standards and the adherence to data protection law, the government can gain the maximum benefit from sharing digital data. Trilateral has a lot of experience with this approach and numerous examples of ethically-sound, privacy-preserving and lawful data sharing across different bodies, which ultimately enables public trust in the process.

January 2022