PRIVITAR
Research & Policy

Health Data Sharing Case Studies

Executive Summary

2021

www.privitar.com

Introduction

Sharing health data for research purposes saves lives

The Medical Research Council (MRC) says that using that data responsibly supports research that delivers more effective treatment, helps to identify and track public health risks, and improves service delivery. At the same time, the National Data Strategy cautions that: "used badly, data could harm people or communities, or have its overwhelming benefits overshadowed by public mistrust."
We think that the needle can be threaded by using holistic risk management to enable responsible data innovation.

Organisations considering sharing health data they hold are often unsure about how to evaluate the risks associated with data sharing, the controls they can apply and how to interpret and apply their legal obligations. The law and guidance provide guardrails, but are often expressed in high level principles. This leaves organisations with the challenge of bridging the gap between principles and implementation. Room for discretion can create uncertainty, which slows down the decision-making process. Slower decisions mean researchers struggle to get timely access to the data they need.

Our Health Data Sharing Case Studies report presents two case studies documenting the data sharing processes at Cambridge University Hospital Trust (CUH) and the Centre for Epidemiology Versus Arthritis (CfE) at the University of Manchester. It describes an 'idealised' data sharing process that we hope other organisations can adapt to their needs. We worked with experts in the field, and with a panel of four reviewers including the Office of the National Data Guardian and the MRC Regulatory Support Centre.

This executive summary highlights our recommendations and key findings. It signposts sections of the main report and presents our conclusions without going into the detail underpinning them.

We believe that organisations can overcome uncertainty by developing robust processes for deciding when to share data. A robust health data sharing process should comply with the law, regulation and information governance rules, and allow fast, safe access to data for researchers. The National Data Strategy calls this "data availability." A process that can show it meets these requirements builds trust and confidence in data sharing and data use. We also believe that highlighting data sharing best practice will increase data availability and improve UK health research.

We are grateful to the case study owners, reviewers and others who have given their time and insights to support this project. We welcome your thoughts on this work, and are open to carrying out further case studies as we build a library of best practice. You can contact us on policy@privitar.com.

Key Findings - the challenges and responses

Health data use must meet requirements from legislation and common law, policy and guidance. We found that organisations face broadly similar challenges in responding to requests for access to data.

We group these challenges into the six categories described below. We also found that the case study organisations address these challenges in similar ways, such as by defining a clear process with associated roles and responsibilities.

We identified time to data, in other words the time delay between a researcher submitting a request and receiving data, as a key goal to which many of the challenges relate. The six categories are described in detail at Section 3.1 of the main report. In brief, they are:

1 Managing competing priorities. The case study organisations consider requests for access to data in the round. In other words, they did not assess privacy and data protection issues separately but considered them alongside other factors (including scientific merit or resource constraints). This can lead to different views between stakeholders, making it difficult to find a way forward. The organisation has to balance competing priorities when making decisions which could advance one set of interests but limit others, creating an optimisation challenge.

Organisations can respond by clearly defining their priorities and risk appetite. This could involve the senior leadership team making an explicit statement about the risk that the organisation is willing to accept. It could also emerge through precedent. As an organisation makes decisions on whether or not to share data in response to specific requests, a picture of organisational risk appetite will emerge.

2 Coordinating stakeholders in the decision making process. The data sharing process needs many internal (for example, information governance, legal) and external stakeholders (such as independent approval bodies) to work together. Some stages in the process only happen when other stages are complete. It can be difficult for stakeholders, including the researcher requesting data, to follow the process and to understand their role at each stage. Without proper coordination, this can cause delays.

Organisations can respond by designing processes to avoid unnecessary delays. For example, allowing asynchronous decisions rather than convening a monthly meeting. Organisations can also clearly document their data sharing process to improve coordination and ensure that stakeholders understand their roles. Accelerating reviews by distinguishing between ‘routine’ and ‘non-routine’ requests, frees up resources to focus on the more difficult decisions.

3 Managing re-identification risk. An overly risk-averse approach can prevent innovation, but a high tolerance for risk may not be shared by all stakeholders (for example, patients) and could undermine trust in data sharing. Re-identification risk can affect:

a. The balance between the usefulness of the data with the need to protect individuals to avoid losing any key information unnecessarily. This balance depends on the specific research in question.

b. The lawful basis for processing data under the GDPR.[1] The exemptions for processing health data under Article 9 of the GDPR require "suitable and specific measures" to protect individuals.

c. The legal status of the data. This can include determining if a common law duty of confidentiality applies and if the re-identification risk is sufficiently low for the data to be considered anonymous under the GDPR.

Items (b) and (c) show the link between re-identification risk and the legal obligations associated with data.

Organisations can de-identify data (for example, by removing direct identifiers or reducing the level of detail in the data), apply data minimisation and impose conditions on data use, for example contracts stating that recipients will not link data unless authorised. The case study organisations involved requesters in decisions about what controls to apply, to make sure that the controls did not render the data unsuitable for the recipient's intended purpose.

We found that organisations applied two 'layers' of de-identification and data minimisation. A general layer at the data on-boarding stage (for example, deciding not to onboard some values, such as free text notes) and a specific layer to fine tune the controls to match the requester's specific requirements. For example a date of birth could be transformed into age bands, clipped to just month and year of birth or have some random noise added (perturbation) depending on the researcher's requirements.

4 Managing the risk of data misuse. Some organisations share data by providing a data extract. This creates a risk of data misuse, for example using the data for an unauthorised purpose or access by unauthorised individuals.

One of the case study organisations specified that researchers can only use data in specific IT environments managed by the Trust or the University. Both organisations impose terms and conditions on data use. Technical options like trusted research environments are providing alternatives to sharing extracts.

1. The European Union (Withdrawal) Act 2018 transposed the GDPR into UK law as the UK GDPR. The UK GDPR has undergone minor changes to enable it to operate as an independent piece of UK law, but there are currently no substantive differences between the UK GDPR and the GDPR.

5 Providing the right amount of information about the data. Providing information about the data to researchers at an early stage of the process can save time and effort. This information could include how the data is set out (in other words, the schema), what it contains or allowing researchers to perform limited queries on the data to ensure that it is suitable for their intended research.

Initiatives in the health sector, including the **Health Data Research Innovation Gateway,** play an important role in managing this challenge.

6 Continuously improving the data sharing process. Organisations need to monitor the data sharing process to make sure that it works as expected and, if not, intervene to improve it. Organisations also need to demonstrate the benefits of sharing data.

The case study organisations developed proxy indicators, for example the number of academic papers based on data they shared. They also collect feedback from stakeholders and keep their data sharing processes under review.

Recommendations

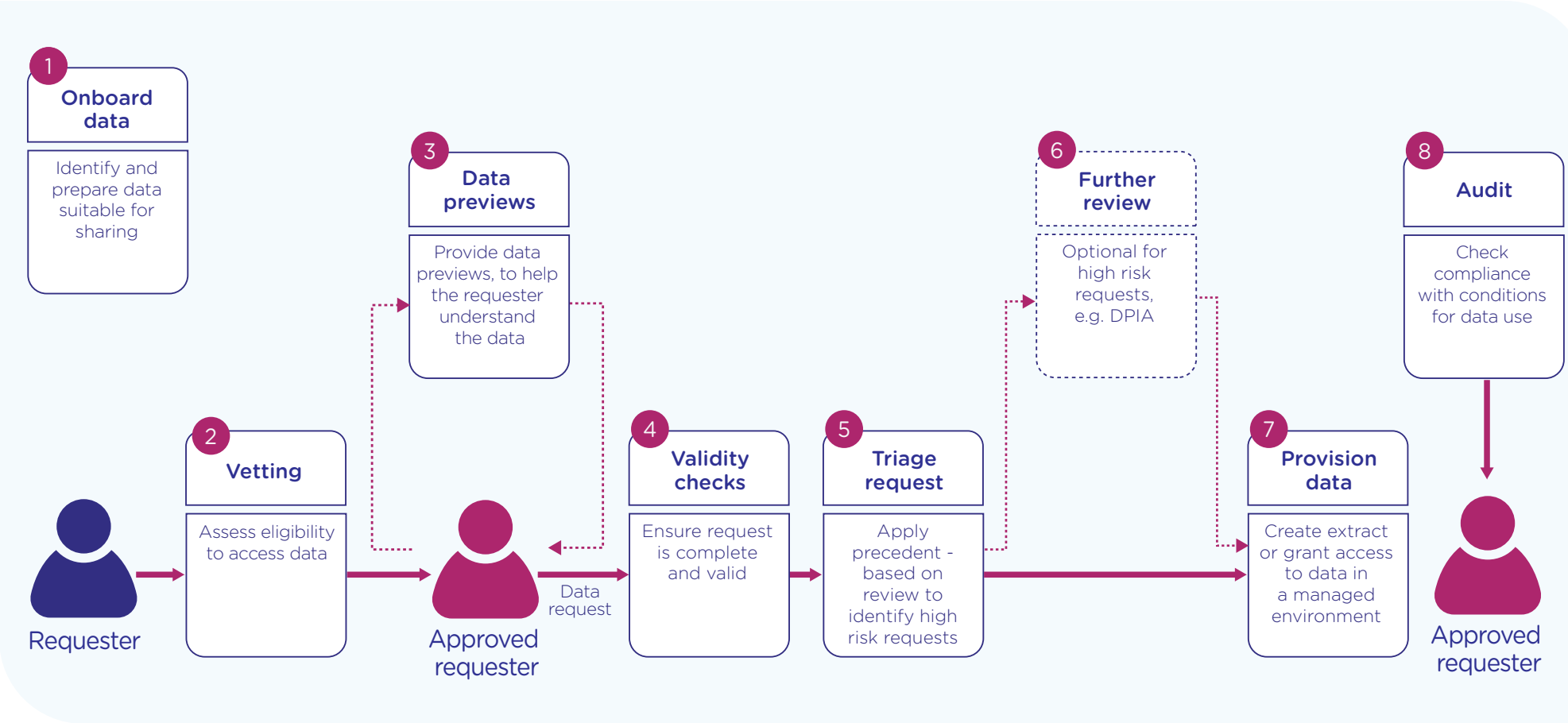
Our recommendations focus on three things an organisation can do when considering data sharing. They can help organisations to respond to the challenges described above. See Sections 3.2, 3.3 and 3.4 of our report for more detail on each recommendation.

1. Set up an information governance framework and a data sharing process. This will generally have four elements:
    a. Developing an information governance framework, a set of internal rules and procedures governing how the organisation will handle data. It helps the organisation to comply with legal and regulatory requirements, defines roles and responsibilities, describes internal processes, and includes an assessment of the risks associated with data sharing.
    b. Preparing to share data. This could include building a dataset suitable for sharing, for example a subset of a 'live' dataset, onboarding data requestors, and providing information about the data.
    c. Manage data sharing. This could include accepting data access requests, applying precedent based triage to evaluate and respond, then select and apply controls to the data and environment as necessary.
    d. Auditing and reporting. This could include internal audits and publishing information (such as what data has been shared with whom and for what purposes) to support transparency.

2. Define the roles and responsibilities needed to support the data sharing process. The table in Section 3.3 of the main report describes these in detail, noting that the names and precise responsibilities will vary between organisations.

3. Define and build a broad range of controls. The case study organisations used a broad, easily explained, auditable range of controls. They tended to use the Five Safes framework developed by the Office for National Statistics (ONS). The Fives Safes encourages organisations to consider risk associated with a project, people, settings, data and outputs.

Example data provisioning process



About Privitar

Privitar is the leader in modern data provisioning. We empower organisations to use data safely, quickly and at scale. Our clients use the Privitar Data Provisioning Platform to share data, unlock insights, keep data safe and support regulatory compliance. Our platform includes state-of-the-art privacy enhancing technologies, and our experts help customers to use them effectively. Only Privitar has the right combination of technology and expertise to create a safe data provisioning ecosystem.

Privitar's Research and Policy teams work with world leading academics, policy makers, regulators, and other experts to investigate how technology can help to preserve privacy while utilising data. Privacy is context specific. Privacy risk to individuals varies depending on factors including what data is being used, by whom and for what purposes. We work to understand the context and the role that technology can play in helping organisations to manage privacy risk.

Contact us:

e: info@privitar.com
t: UK +44 203 282 7136
w: www.privitar.com

PRIVITAR
Research & Policy

www.privitar.com