"Get to it!" Clinical data sharing: Myth-busting and strategy

Louise Corti (https://www.essex.ac.uk/people/corti68302/louise-corti), Service Director, Collections Development and Producer Relations and Associate Director, and Deb Wiltshire, Senior Access and Support Officer, UK Data Service, introduce a workshop on sharing and archiving data derived from academically-run clinical trials the UK Data Service (http://www.ukdataservice.ac.uk/) ran with Nancy Medley (https://www.researchgate.net/profile/Nancy_Medley) from the University of Liverpool Department of Women's and Children's Health and Cochrane Pregnancy & Childbirth (https://pregnancy.cochrane.org/welcome).

On 5 July Cochrane Pregnancy & Childbirth and the UK Data Service held a half-day workshop aimed at demystifying some of the conceptions and misconceptions around sharing and archiving data derived from academically run clinical trials.

There is no proposal to share any personal data without appropriate security, as outlined below. The purpose of this workshop was to discuss and identify situations where it might be appropriate to share data and how best practice would make this process secure and safe.

On behalf of the UK Data Service (https://www.ukdataservice.ac.uk/), Louise has led discussions with a range of interest groups and policymakers on the practicalities of archiving and sharing data from completed clinical trials. The conversations address calls from funders to improve data sharing of research they have supported and requests from the clinical trials community for practical steps that can be taken to safely archive data.

Why share clinical trials data?

Researchers in both the academic and public sectors are experiencing an increased emphasis on demonstrating research integrity and reproducibility; funders, journals and professional bodies concerned with research conduct expect data usage to be transparent and reproducible. Louise reports:

> "The pressure to conduct open science and share data and thereby enable others to reproduce scientific results is increasing year on year. Threats to scientific integrity, such as fabrication of data and results, have led some journals to request individual-level data, syntax and prior registration of hypotheses as part of a paper's peer-review. In 2016, Penny's article in (https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970) *Nature* (https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970) revealed the worry amongst researchers about a reproducibility crisis; 90% of researchers surveyed agreed there was a crisis. Amongst other such reports and paper retractions, scientific misconduct has been bought to the fore. What followed was an explicit strengthening of integrity initiatives, certainly across European countries."

Making clinical trials data safely available for research enables researchers to analyse data as a secondary source and conduct individual-level meta-analyses across trials, leading to enhanced efficiency, innovation and impact from existing data resources. Lecturers can benefit from having trials data available for teaching purposes, for example, a simplified 'real-life' teaching dataset for training in randomised controlled trial (RCT) design and analysis.

Initiatives aimed at enhancing transparency in clinical trials, such as the AllTrials (http://www.alltrials.net/find-out-more/) campaign championed by Ben Goldacre and colleagues; (http://www.alltrials.net/) led to the requirement by the International Committee of Medical Journal Editors that a data sharing plan be included in each paper, and prespecified in study registration (http://www.icmje.org/about-icmje/faqs/clinical-trials-registration/).

Two of the leading general medical journals, *The BMJ* and *PLOS Medicine*, have implemented a policy of requiring data sharing as a condition for publication of RCT and the Journal of Clinical Epidemiology (https://www.elsevier.com/journals/journal-of-clinical-epidemiology/0895-4356/guide-for-authors#20901) has mandated a data sharing plan and trial registration:

> "The Journal endorses the World Health Organization Standards for registration of all human medical research (http://www.who.int/ictrp). JCE considers original papers reporting results of clinical trials if they have been registered in a clinical trial registry. Authors are asked to provide registration details in the submission letter and the manuscript"

A first step has been to require the listing of clinical trials (CTs), across the private and public sectors. IT companies are also starting to offer technical platforms for data discovery and access. Clinical Study Data Request (CSDR) and Vivli offer such listing platforms where basic metadata about trails can be searched. A standardised description provides an overview of the trial, its medical condition, intervention, phase and sponsor. Access to the data from these CTs can be requested, typically via an email to the data owner. The storage of trials datasets is at cost to trials producing organisations.
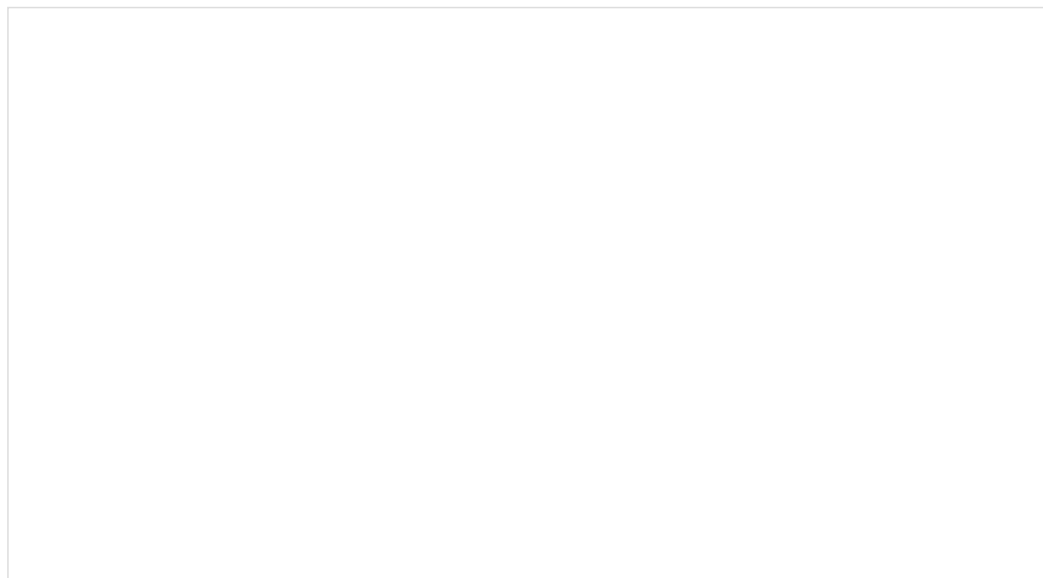
In 2017 funders including the Medical Research Council (https://mrc.ukri.org/), the Wellcome Trust (https://wellcome.ac.uk/) and Cancer Research UK (https://www.cancerresearchuk.org/) financially supported the CSDR system, encouraging all their funded CTs to list themselves on the portal for easy discovery. However, making the raw data available on request is still not yet happening. Indeed, many trails are still unavailable for re-use.

Following the CDSR initiative, these funders met with the UK Data Service to explore the potential for the Service to hold legacy data from completed trials they funded, with the appropriate level of protection put in place for each trial. As a result, these trials were encouraged to offer their data to the UK Data Service, with the MRC actively promoting the call (https://mrc.ukri.org/research/facilities-and-resources-for-researchers/clinical-trials-data-sharing-resources/) to list and share data.

To date the offer has not been taken up. We have further identified opportunities to complement this process by developing additional levels of clarity in areas such as redaction procedures, file formats, data versioning and persistent identifiers.

Pushing forward on clinical trials data sharing

Working with the funders and linking up with some friendly trialists, Louise and colleagues attempted to push forward on some of these more practical issues surrounding sharing. The proactive work aimed to support data owners to look in detail at possible data publishing and access workflows, that would move trials beyond a simple listing in an online registry. Louise notes:

> "I wanted to explore where the 'blockages' were, particularly having witnessed positive sentiments to share trials data, yet a certain inertia in moving forward. I saw real concerns over who is best placed to take responsibility for risk and harm resulting from exposing patients, through IPD sharing. The fear is high, even when the risk might be low, and barriers to sharing are complicated by multiple actors and stakeholders in the 'risk chain'. Pilot studies have been actively sought to work through some of the practical steps and barriers."

In the autumn of 2018 Louise was invited to a meeting of the UK Clinical Trials Network in Leeds, which bought together triallists in a positive space to present both case studies of data sharing legacy trials, legal frameworks and governance strategies, and more practical issues, such as suitable formats and metadata. The tone of the whole meeting was upbeat and positive, with enthusiasm to help motivate data sharing.

Speakers at the meeting include those who had already sought out practical methods for preparing data and governance frameworks, or had at least estimated the resources needed to do this. Catrin Tudor Smith (https://www.liverpool.ac.uk/translational-medicine/staff/catrin-tudur-smith/publications/) of the North West Hub for Trials Methodology Research (https://www.liverpool.ac.uk/translational-medicine/departmentsandgroups/nwhtmr/) in the Department of Biostatistics at the University of Liverpool presented on how patient or participant data (IPD) from publically funded trials should be shared, based on her 2015 paper. The paper sets out basic requirements for how to share a trials dataset. She further highlighted the results from a 2014 survey of IPD reviewers, who felt that a central repository for storing IPD (https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0097886) would be valuable. Further, with her colleagues, Tudor Smith and Nevitt presented their 2017 work on resource implications (https://www.ncbi.nlm.nih.gov/pubmed/28712359) of preparing data (https://www.ncbi.nlm.nih.gov/pubmed/28712359) from two UK publicly funded clinical trials (the SANAD trial with over 2000 participants with epilepsy recruited from 90 hospital outpatient clinic; and the MENDS trial with 146 children with neurodevelopmental problems recruited from 18 hospitals).

These practical approaches demonstrate that it is, unsurprisingly, very close to the methodology for preparing and publishing cohort study data; in effect clinical trials are a narrowly focused longitudinal study, with clinical measures and limited phenotypic data, but with far less social data, as context.

The Cochrane connection

In the autumn of 2018, Nancy Medley, a researcher affiliated with the Harris Wellbeing Preterm Birth Centre at the University of Liverpool (https://www.harris-wellbeingptbcentre.co.uk/) and with Cochrane Pregnancy & Childbirth (https://pregnancy.cochrane.org/), made contact with Louise to discuss data sharing and how the UK Data Service operated as a data curator, publisher and access point. The Harris Wellbeing/Cochrane research team in Liverpool, led by Professor Zarko Alfirevic, had already secured an NIHR Cochrane Programme grant to

investigate the practicalities of onward sharing of IPD from clinical trials. Their 2 year project aimed to host and facilitate onward sharing of a set of clinical trials of progestogens to prevent preterm birth, from the EPPPIC project (https://doi.org/10.1186/s13643-017-0600-x), led by one of the grant team, Professor Lesley Stewart.

The rationale for Cochrane to endorse a data sharing pilot via an NIHR (https://www.nihr.ac.uk/) Cochrane programme grant was twofold:

- First that patient data would improve Cochrane systematic reviews on priority topics by enabling more in-depth analyses and better answers to an interventions' impact on subgroups, in particular;
- Second, requiring sharing of patient data may raise the bar for future Cochrane systematic reviews and may even help identify fraudulent trials.

In addition to creating a trials data repository at University of Liverpool, with governance systems in place to structure and manage data access, researchers wanted to explore what infrastructure was already available for trials data owners. The UK Data Service was a clear example of a data repository with infrastructure to offer data owners multiple data publishing pathways. The Service also had valuable experience of the legal and governance problems that could delay and frustrate trials' data sharing, and practical solutions to data owners' questions about the technical side of data sharing – formats, metadata, coding frames, future-proofing, and data security. Researchers at the University of Liverpool began exploring a collaboration between Cochrane and the Service to better understand what would be required for Cochrane to enable data sharing.

Over a six-month period, Liverpool researchers and the Service exchanged calls and procedures and documentation relating to data sharing. The two centres agreed it would be beneficial to design an afternoon workshop for clinical trials' leads and data managers at the Liverpool Women's hospital and the University of Liverpool, as part of the NIHR public engagement component of the Cochrane programme grant. The focus of the workshop would be to move beyond legal agreements, and dig down into the actual and perceived risks in a shared trial dataset, and in outputs arising from reuse of data. To avoid sounding dull to potential audiences, the data sharing workshop was given a playful name, *Get to it! Clinical data sharing myth-busting and strategy.*

The workshop took place on 5 July on the site of the Liverpool Women's Hospital education centre with around 45 in attendance.
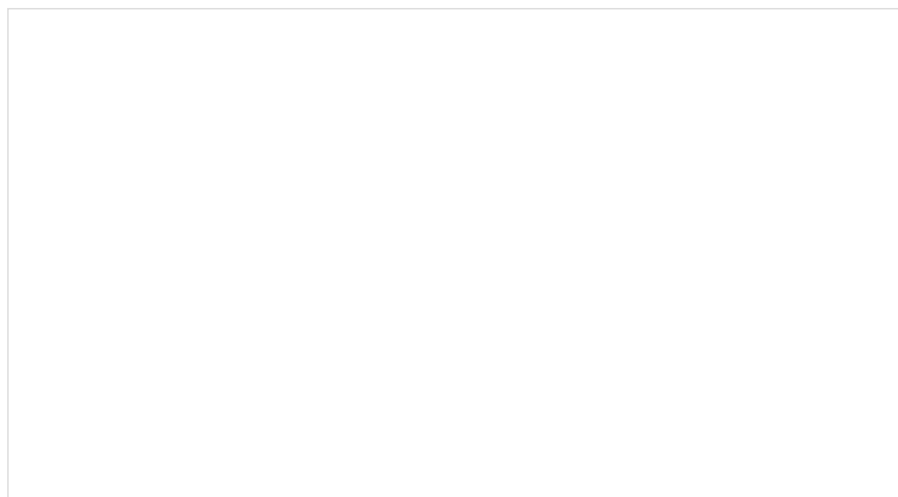
What does a shared dataset look like?

The UK Data Service, supported by the ESRC (https://esrc.ukri.org/), provides long-term data curation and continuous access to key data assets of interest to research and teaching for the broader range of social sciences. It offers managed services for data owners, enabling access to key data assets that are free at the point of use.

We handle thousands of data collections from social and epidemiological surveys, psychology experiments, and fieldwork on a huge rage of topics studied by social scientists. One of our speciality areas is handing large-scale government social surveys and the national long-running UK cohort and longitudinal studies such as National Child Development Study (https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000032), 1970 British Cohort Study (https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=200001), Next Steps (https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000030), Millennium Cohort Study (https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000031) and Understanding Society (https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=2000053). All of these studies include significant biomedical measures.

A robust licensing and governance framework is used to ensure that datasets are made available under terms specified by the data owner. We operate a three-tier **data access policy** where data can be:

- made openly available as **Open** data;
- placed behind gates as **Safeguarded data,** where data are downloadable with authentication and authorisation and, where necessary, owner approval; or
- be available as **Controlled data** through our safe haven, the **UK Data Service Secure Lab**.

The later tier embraces the principles of the **Five Safes** (https://www.ukdataservice.ac.uk/manage-data/legal-ethical/access-control/five-safes)framework (https://www.ukdataservice.ac.uk/manage-data/legal-ethical/access-control/five-safes), so that data classified as personal data can be made available via appropriate legal gateways. Data owners make independent decisions about who can use their data (**Safe Projects)** , mandatory face-to-face training must be undertaken to become a '**Safe Researcher**' and use our ISO 27001-approved remote access **Safe Setting**; and all their research outputs are checked (**Safe Outputs**).

The 5 Safes set of principles has gained traction with national statistics around the world, yet it is remarkably absent in the narrative of data access for health research. The health domain tends to focus on 'data sharing agreements', and less on training around trust, security and disclosure. The concept of a **Safe Health Researcher** is missing, yet is appealing. We come to this aspect later.
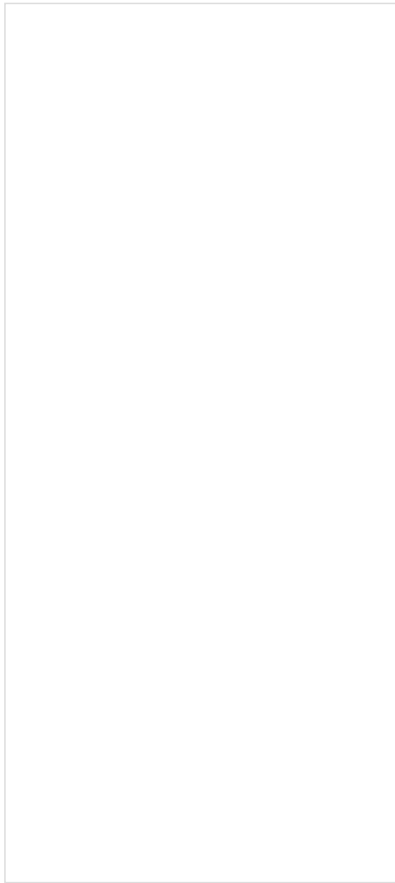
Barriers to sharing

The UK Data Service's workshop on sharing and archiving data derived from academically-run CTs aimed to understand the barriers to sharing these data and included a debate on **To Share or Not to Share** where attendees were invited to take up a pro or contrary position towards data sharing, based on the following research scenario:

> You have undertaken a research and clinical trials programme that investigates the epidemiology and treatment of Chronic Fatigue Syndrome (CFS)/ME in young adults. The work is funded by the Medical Research Council (MRC) and you are expected to share the data. Can you share the data and what are the key issues?

The groups identified some of the following **Opportunities** and **Problems** around sharing these data:

Opportunities

- Improved health outcomes through validated research
- Help progress science, building on a foundation of trusted evidence, which can be connected
- Avoid expensive data collection and duplication
- Help reduce the burden for already over-researched patient groups
- Opportunities for novel research through unanticipated analyses of the data
- Provide greater research transparency/reproducibility/accountability
- Enhance academic impact and credibility, with opportunities for further funding (and risk of not getting research papers published or losing funds if data not shared
- Enable data harmonisation
- Do due diligence for patients who have voluntarily contributed data
- Real life secondary data can bring teaching to life

Problems

- Risk of disclosing personal, confidential/ sensitive information, especially when linked to public sources
- Securing data sharing contracts with local legal entities
- Being 'forced' to share data can lead to risk of misuse of data, with the possibility of damaging outcomes for patients and for data owners
- Lack of detailed knowledge of how to prepare a clean dataset for sharing

- Lack of funds and resource to prepare older datasets
- Limiting consent statements in earlier studies

The problems and opportunities identified by the group reflected the findings from the most recent 2017 Wellcome survey of researchers (Van den Eynden et al, 2017) (https://www.researchgate.net/publication/312464771_Survey_of_Wellcome_researchers_and_their_attitudes_to_open_research). In this case, motivating factors for sharing emphasised the need for following good research practice and offering up opportunities for collaboration and visibility. Key problems highlighted at the workshop included a lack of funding and time and effort to prepare a shareable dataset.

The Wellcome study had noted that established researchers were motivated to share data as an opportunity to secure additional funding, early career researchers viewed sharing as a way of enhancing academic reputation, adding visibility and offering the chance to collaborate and co-author research on projects reusing data. Trials managers agreed that enhanced prospective data management and resourcing would help data sharing from new trials; as would using a third party to manage and access and governance.

Legal agreements

An additional barrier identified for CT data sharing is the practicalities of establishing final legal agreements which meet the needs of all parties. This can be a cause of frustration around creating a mutual understanding of the management of risk and responsibility, particularly for the host institution, such as the University in which trials unit is based. The workshop attendees identified the following potential ways to overcome these issues:

- standard *data licensing and user agreements* via a trusted accredited third party, instead of typical 1-1 Owner-User data sharing contracts
- project approval boards
- transparency around the application and data supply processes, including robust and auditable access control.

University contracts offices might benefit from having a better understanding of the requirements and context of the research and various integrity drivers and mandates. Promoting better connections between contracts staff and institutional research office staff who work on CT grant application and administration may simplify the process.

Appraising risk in data and outputs

Some of the barriers unearthed for the CT community were addressed in the subsequent sessions. Practical solutions that could be used to mitigate against data disclosure risk include:

- appropriate assessment and treatment of actual risk in a dataset, and the likely risk of jigsaw identification (when various independent sources of data are pieced together);
- use of the Five Safes (https://www.ukdataservice.ac.uk/manage-data/legal-ethical/access-control/five-safes) to manage risk, using governance and controls
- for unsafe data e.g. that declared personal under GDPR, the use of legal gateways, and the vetting of projects and of outputs.

The UK Data Service provides regular training (https://www.ukdataservice.ac.uk/news-and-events/events.aspx) on each of these elements.

Disclosure risk assessment

Data owners understand that personal information needs to be treated with due diligence for ethical reasons or legal reasons, such as:

- protecting confidential sensitive or illegal information, or disguising a research location
- avoiding disclosure of personal data under GDPR or reputation damage
- for commercial reasons, such as commercial sensitivity, avoiding revealing trade secrets or patented information.

Individual patient data should never be disclosed, unless a participant has given specific consent to do so or a legally compliant exemption is in place to do so. An example of such an exemption in England and Wales is Section 251 of the NHS Act 2006 (https://www.legislation.gov.uk/ukpga/2006/41/section/251) which provides the statutory power for the NHS to use information which may identify patients who have not given consent, where such data support core NHS activity.

Disclosure occurs when a person or an organisation uses published data in order to find and reveal sensitive or unknown information about a data subject. A person's identity can be disclosed from:

- *direct identifiers* such as names, addresses, postcode, telephone numbers, email addresses, IP addresses, or pictures;
- *indirect identifiers* which, especially in combination with other publicly available information sources, could identify someone, such as information on workplace or residence, occupation or exceptional values of characteristics like salary or age.

Demographic information can be particularly revealing when combined with other data such as a detailed employment description and a very localised geographic area. Text-based and open-ended variables in surveys can contain detailed unique information.

Sensitive information such as health or religious status, or data relating to crime, drugs etc also need to be considered. Where there is no identifying information in a dataset and it cannot be combined, these measures are not in themselves risky. Once assessed, data can be either treated or access controlled.

Data publishers aim to minimise the potential risk of disclosure to an appropriate level while sharing as much data as possible. However, the risk is likely never zero. Data access involves reduction of risk in a manner acceptable to the data owner; risks are mitigated by legal gateways and appropriate safeguards.

Much research data shared for reuse have safeguards placed upon them, including legal agreements signed by users not to disclose any confidential information. Where research data plans are to be widely shared, confidentiality is usually maintained through de-identification or anonymisation. Where there are legal gateways in place, disclosive information can be shared, such as using administrative or register data for research.

Privacy assessment strategies therefore need to devise a plan for striking a balance between removing or replacing disclosive information with retaining as much meaningful information as possible: balancing risk with usefulness. The Information Commissioner's Office (ICO) (https://ico.org.uk/) supplies useful questions and guidance on how a Data Protection Impact Assessment (https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-assessments/) (DPIA) can help identify and minimise data protection risks. A DPIA is required for processing personal data that is likely to result in a high risk to individuals.

There are many different techniques for anonymising data aiming to minimise the risk of patient re-identification within a dataset. Identifying characteristics should be removed or replaced, as recommended by Hrynaszkiewicz and colleagues (https://www.ncbi.nlm.nih.gov/pubmed/20113465), who early on in 2010, identified around 28 potential patient identifiers.

In 2015 the Medical Research Council Hubs for Trials Methodology Research (now part of the MRC-Trials Methodology Research Partnership (https://www.methodologyhubs.mrc.ac.uk/)) published best practice (https://www.methodologyhubs.mrc.ac.uk/files/7114/3682/3831/Datasharingguidance2015.pdf) principles for preparing individual patient data from publicly-funded clinical trials as an 'anonymised dataset' ready for sharing, having determined an appropriate level of anonymisation.

More recently, the Medical Research Council Regulatory Support Centre (https://mrc.ukri.org/research/facilities-and-resources-for-researchers/regulatory-support-centre/) authored excellent guidance on Identifiability, anonymisation and pseudonymisation (https://mrc.ukri.org/documents/pdf/gdpr-guidance-note-5-identifiability-anonymisation-and-pseudonymisation/) in line with the Information Commissioner's Office guidance. The idea of controlling both 'content' and 'context' are useful; as is the idea of risk of jigsaw identification; when various independent sources of data are pieced together.

Such approaches affect decisions about sharing much older legacy CTs.

Most importantly, any Duty of Confidentiality (DoC) made at the time of the original research must be upheld. Where participants are dead, or alive but cannot be contacted, or where resources required to contact would be too costly, 'reasonable expectations' test can be created (for example, via a research ethics review committee). Such a test would identify whether a new intended use is in keeping with the broader aims of the study, and/or in the public interest. The wording in much older consent forms may be put to a test, as GDPR and other regulations now support a more precise definition of personal data or data with risk of identification, under GDPR.

The workshop worked through examples and exercises demonstrating how a CT dataset could be evaluated for risk and then treated for sharing; namely by removing direct identifiers and by looking at unique combinations of identifying or unusual information. In the case of identifying example attributes of a mother in the trial giving birth were:

- very high BMI,
- high parity (the number of times a female is or has been pregnant (gravidity) and carried the pregnancies to a viable gestational age)
- death as an outcome.

The examples of drug administration, physiological/ blood results would likely not be.

Safe projects, Safe people and Safe outputs

For data that are under controlled access, i.e. accessible from the Services' Secure Lab safe haven, in addition to project approval by the data owner, training is provided on a regular basis around the 5 Safes, especially around becoming a 'Safe Researcher'. Our Liverpool workshop was a great opportunity to pilot materials oriented towards biomedical researchers.

Safe projects

A very useful discussion was held on how to manage project approvals so that the project is viewed as ethical and legal (what we term a **Safe Project**) and the process of approval is fair and transparent.

Examples of how to run approvals panels, how they could be constituted and show transparency was explained, using the model of the ONS Research Accreditation Pane (https://www.statisticsauthority.gov.uk/about-the-authority/better-useofdata-statistics-and-research/betterdataaccess-research/better-use-of-data/)l (RAP) and the NHS Digital Independent Group Advising on the Release of Data (https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/independent-group-advising-on-the-release-of-data) (IGARD); applications details and decisions are recorded in minutes published from the regular committee meetings.

Both these panels are designed to provide independent review of data applications that meet the terms and conditions specified by a data owner, or legal gateway. Importantly, they are usually made up of multi-stakeholder representatives. The UK Data Service has agreed to run a webinar on Project Approval Panels and how to set up an approval form and an initial screening process.

Safe People

The session on **Safe People** outlined a framework of how we might go about demonstrating 'trust' in researchers to use data assets safely and securely.

Training can help consolidate the less well-defined idea of a 'bona fide' researcher. This status, though not fully defined, is typically required by funders such as the UK's Medical Research Council, as a pre condition for accessing available data holdings. The term 'bona fide' makes some assumptions about the credentials of the researchers, yet does not 'test' them, instead relying on 'trust 'underwritten by the individual's university, and maybe a short online security course.

While purposeful breaches are certainly not common place, a researcher accessing personal or sensitive data does benefit from a structured course that covers aspects of: potential or actual disclosure risk in health data and appropriate publishing and access pathways; safeguards that should be put in place when data with risk are shared; researcher behaviour and expectations; and what might constitute a published 'unsafe' output, i.e. with a risk of disclosure.

The workshop outlined the process of becoming a **Safe Researcher** run by the UK Data Service, **Safe Researcher Training**, and some of the key learning objectives.

All researchers wishing to access the UK Data Service Secure Lab (Safe Setting) must pass a test at the end of the course allowing us to formally test their understanding of the key principles of safe use of personal data, which acts to reassure data owners that researchers can be trusted. While Safe Researcher Training is largely focused on business and economics data, examples of the training can be easily adapted to suit the needs of health researcher, from clinical trials to epidemiology.

Safe Outputs

The final part of the workshop focused on **Safe Outputs** using research examples using health data. This session began by explaining the principles of statistical disclosure control (SDC), the technique of minimising any residual risk of identification and/or association of information in a statistical output.

The group worked through examples of common output types to demonstrate how identification might occur and how researchers could ensure that their outputs are safe. Whilst this risk may be low, SDC is an important precautionary tool which many safe havens employ to further ensure the safe use of personal or sensitive data.

An understanding of SDC is especially beneficial for all researchers working with sensitive or personal data regardless of through which route they access these data. Researchers often have little or no prior knowledge of output SDC, but would benefit from utilising these SDC techniques themselves. The benefits of a **Safe Health Researcher Training** is that it can supply researchers with this concrete knowledge and understanding.

What next?

Nancy was pleased with the training, commenting:

> "I was particularly impressed with the diverse audience; we had senior and junior clinical researchers, statisticians and data managers alongside research nurses and student midwives. Many people in our clinical community want to share their data but don't know quite the best way to do so, and the Data Service provided practical examples, and importantly, contact details for experts like Louise who are willing to share their knowledge."

Following the workshop Nancy and colleagues sent out a link to a brief survey. When asked about how they would share their own data, 7 of the 11 respondents stated that they would use the UK Data Service. Other respondents cited Dryad or Figshare or sharing the data themselves or via safe data transfer platforms.

8 of the 11 respondents thought a webinar on setting up data access committees and managing applications to share data would be valuable. An open-ended final question about unanswered problems brought up two additional issues:

- whether or not data can be uploaded to multiple platforms,
- methods for ensuring data are anonymous.

Full results of the survey (https://www.surveymonkey.com/results/SM-F97N39L97/) are available.

Safe Health Researcher

The material presented at the workshop draws upon and extends elements of the research data management and publishing training undertaken by the UK Data Service (https://uk.sagepub.com/en-gb/eur/managing-and-sharing-research-data/book240297) and the UK Statistics Authority (https://www.statisticsauthority.gov.uk)-approved 'Safe Researcher Training' course, which leads to Accredited Researcher status (https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme#ons-researcher-accreditation).

Over the coming months Louise Corti and colleagues will be liaising with other health data owners and providers to examine the appeal of developing training to become a **Safe Health Researcher**.

Louise Corti is Service Director, Collections Development and Producer Relations for the UK Data Service. Louise leads two teams dedicated to enriching the breadth and quality of the UK Data Service data collection: Collections Development and Producer Support. She is an associate director of the UK Data Archive with special expertise in research integrity, research data management and the

archiving and reuse of qualitative data.

Leave a Reply

Your email address will not be published. Required fields are marked *

Comment

Name *

Email *

Website

Post Comment

Categories

Select Category ▾

Posts by month

Select Month ▾

Search the Data Impact blog

Search	Search

- UK Data Service (https://www.ukdataservice.ac.uk)
- Our impact (https://impact.ukdataservice.ac.uk/)

- Impact Case Studies
  (https://beta.ukdataservice.ac.uk/impact/case-studies)
- Data Impact Fellows
  (http://blog.ukdataservice.ac.uk/category/data-impact-fellows/)
- Impact and Innovation Lab
  (http://blog.ukdataservice.ac.uk/category/impact-and-
  innovation-lab/)
- Accessibility (http://blog.ukdataservice.ac.uk/accessibility/)

UK Data Service

Funded by the Economic and Social Research Council (ESRC), the UK Data Service provides access to international social, economic and population data.

The Data Impact Blog

The Data Impact blog is run by the UK Data Service and is a hub for anyone interested in maximising the impact of social, population and economic data in research and policy.

Latest Post

Using data for effective health service evaluation: why we need data stewards (http://blog.mimas.ac.uk/ukdataservice/?p=5238)