Safe sharing of research data: the role of legal agreements when anonymising

Thursday 25th April 2019
IET London, 2 Savoy Place, London

MRC Regulatory Support Centre
Safe sharing of research data: the role of legal agreements when anonymising

Medical Research Council, Regulatory Support Centre

Research – for the public good

We do some strange things with data / information!
- We collect data we know is ‘out of date’
- We do not usually make any decisions about people on the basis of the information we hold
- We are often interested in the unusual
- We look for small changes and subtle, complex relationships – we often need lots of data
- We have ethics approval, peer review etc
- Research is very collaborative (internationally)
  - we share data

Today’s objectives

- Determine how it is possible to anonymise information
- Explore the risks of information sharing to support research
- Identify some specific parameters that should be considered when ‘calculating’ such risk
- Explore how to mitigate risk
- Explore how we can build ‘trustworthiness’ within and between organisations

Intended output – to add to our draft Identifiably Guidance (with buy-in from Regulators)
Legal frameworks

Common law of confidentiality – governs who has access to confidential information (in line with ‘reasonable expectations’)

General Data Protection Regulation (new Data Protection Act) – governs when and how personal data is processed (fairly, lawfully and transparently)

What makes information identifiable?

Year of birth
What makes information identifiable?

Gender

First part of postcode

Place of birth
What makes information identifiable?

Year of birth
Gender
First part of postcode
Place of birth

What makes information identifiable?

Year of birth 1965
Gender Female
First part of postcode EH32
Place of birth Bristol

Identifiability

Content (weak or strong identifiers)
Context (What other information do you have access to? Beware of the rare or unusual)
Identifiability – a ‘grey scale’

Anonymous — Identifiable

Content (weak or strong identifiers)
Context (What other information do you have access to? Beware of the rare or unusual)

Identifiability – law is binary

Anonymous — Identifiable

Content (weak or strong identifiers)
Context (What other information do you have access to? Beware of the rare or unusual)

Identifiability – reality of working with scale

Anonymous — Identifiable

Inherently anonymous — Identifiable
Identifiability – reality of working with scale

Inherently anonymous – be aware of outliers

Identifiable

Anonymous
Identifiability – reality of working with scale

Anonymous

Identifiable

We can limit direct identifiers and control context (identification not ‘reasonably likely’ by any likely means)

Inherently anonymous

More motivated than most

Anonymous

Identifiable

We can limit direct identifiers and control context (identification not ‘reasonably likely’ by any likely means)

Inherently anonymous

More motivated than most

Anonymous

Identifiable

We can limit direct identifiers and control context (identification not ‘reasonably likely’ by any likely means)

Inherently anonymous

More motivated than most
We can limit direct identifiers and control context (identification not ‘reasonably likely’ by any likely means).

Inherently anonymous

More motivated than most

• NHS England is reviewing the words used by Information Governance (IG) professionals when describing the identifiability of data.
  • There are for example nearly 30 different terms used in IG documents which include legal and colloquial terms.
  • The aim is to ensure greater clarity, reduce duplication and improve consistency.
  • The approach will align with GDPR and focus on whether the data described is in scope (personal) or out of scope of GDPR (anonymous).
  • Guidance and a simple flow diagram will help IG professionals select the appropriate term.
Introduction to workshops

We are being asked to manage the risk of:
1. being in breach of common law
2. being non-compliant with GDPR
when sharing research information (between research groups) by ensuring information is anonymous.
Introduction to workshops

We are being asked to manage the risk of:
1. being in breach of common law
2. being non-compliant with GDPR when sharing research information (between research groups) by ensuring information is anonymous*

*Other ways of mitigating risk - Consent and ensuring lawful, fair and transparent

Risk assessment

Risk of breaking the law vs risk of not sharing information

Workshop 1 – risk of not sharing

What is the risk of not supporting the sharing of information to support research?

Identify the hazards...to your organisation

I will ask each table to give us ONE hazard identified per table...
Risk assessment

Risk of breaking the law vs risk of not sharing information

Workshop 2 – Risk of non-compliance

What is the risk of being:
1. In breach of common law, and
2. Non-compliant with GDPR when sharing information?

We have provided some broad hazards: can you identify what might contribute to the size of risk associated with these hazards? This includes an assessment of what makes a breach more likely and what would make the impact of such a breach more severe?

Please use your own, day-to-day experiences...
No verbal feedback – please use paper provided to write down main points

Risk assessments

- Case by case?
- Fall into broad groups: types of dataset, types of transfer etc?
- All instances the same?
Risk assessments

- Case by case?
- Fall into broad groups: types of dataset, types of transfer etc?
- All instances the same?
- Future proofing?

Risk assessment

Risk of breaking the law vs risk of not sharing information
Risk assessment

Risk of breaking the law vs risk of not sharing information

How do we mitigate the risk?

If possible: Obtain consent and ensure fair, lawful and transparent etc
Research* is not an incompatible purpose

But is it a new purpose?

Transparency!!

Controls

If consent and transparency are not possible – must rely on sharing only anonymous information, therefore must

- Limit the content of the information to be shared
- Limit the context in which the information will be viewed

Build mutual trustworthiness
Limiting the content of information to be shared

- Information must still be useful
- Justify what needs to be shared
- Implications of linkages proposed etc conducted by recipients
- Beware of outliers….

Who is best placed to do this?
Who understands the data sufficiently?

Workshop 3 – Assurances from researchers

What assurances should you take from your local researcher(s)?
How do they demonstrate trustworthiness to you?

I will ask for feedback on
1. Assurances for which issues would you wish to seek in most cases? Give number from handout!
2. An ideas about what these assurances might involve
3. How do we avoid increasing bureaucracy? (One idea)
Controls

If consent and transparency are not possible – must rely on sharing only anonymous information, therefore must

- Limit the content of the information to be shared
- Limit the context in which the information will be viewed

Build mutual trustworthiness

Alastair Nicholson
Senior Development Manager
Health Research Authority

Workshop 4 – Control of context

How should we control the context in which the transferred information is viewed? How do we ensure researchers can trust us?

Between organisations or within organisations?
Should it be risk informed?
How do you link the risk assessment conducted in workshop 2 and the mitigations discussed in workshop 3 with these measures?

I will ask for feedback on
1. An issue already met when using agreements to manage anonymisation, and
2. How do we ensure we are risk proportionate in terms of controlling context? (One idea per table)
Ouputs from today

RSC will:

- Analyse all of your input from today and use this to inform further development of ‘Identifiability, anonymisation and pseudonymisation’ guidance note
- Consult further with regulators to finalise a revised version of guidance for publication (on our website)

Any further feedback on the draft guidance? Please email: info@rsc.mrc.ac.uk

For support and guidance with:
- Consent, ethics and governance;
- Confidentiality, data protection and data access;
- Translational research (e.g. medicines, devices, in vitro diagnostics, cell and gene therapies etc.)
- Human tissue; and more.

mrc.ukri.org/regulatorysupportcentre