



National Institute for
Health Research

Data and Research



Neil Walker
IT Director

NIHR BioResource for Translational Research

Overview



- A new contract with research participants – the impact of GDPR
- What sorts of data do we collect?
- Is **all** this data available? Yes, but ...
- What does anonymous mean?
- Commercial interests and Ipsos/MORI survey
- Loose ends:
 - Where is this data stored and who can see it
 - New types of data

The impact of GDPR

New EU General Data Protection Regulation (GDPR) comes into force on 25 May 2018

Not asked to re-consent participants, but to:

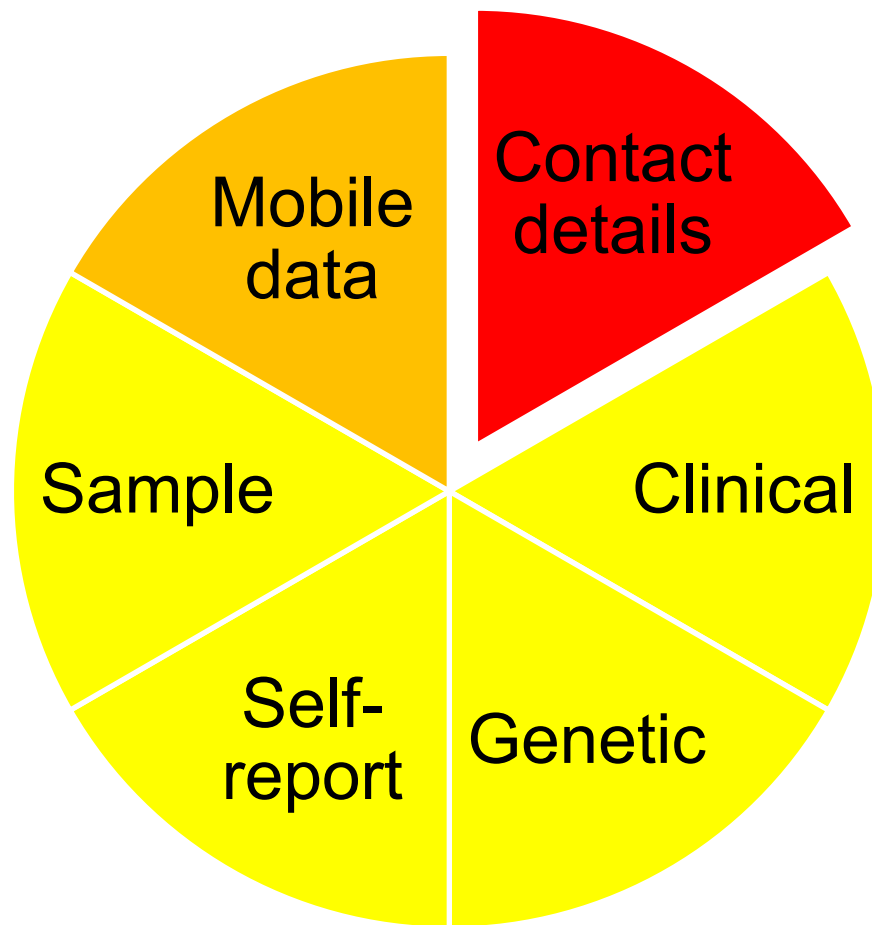
- ✓ Be more transparent
- ✓ Make “data subjects” aware of rights
- ✓ Make it as easy to leave a study as to join



What sorts of data do we collect?

NB: not allowed to collect more data than we will use

Contact details are kept separately- we'll go through this later on in further detail



Is **all** this data available?

A horizontal bar composed of several colored segments: green, dark green, orange, purple, red, dark blue, and light blue.

Yes!

Is **all** this data available?

A horizontal bar composed of several colored segments: green, dark green, orange, purple, red, and blue.

Yes, **but** ...

Depends on how anonymous data is

Depends on who the data is available to

Depends on what the data is used for

What does anonymous mean?

Spectrum of identifiability



Personally identifiable

De-personalised

Anonymous

More identifiable

*anonymised in accordance with the ICO code of anonymisation

Less identifiable

Understanding patient data, please visit goo.gl/wwNyYp

Is it anonymous? It depends who's there...



Google: *white and blue floral shirt*

Does anonymity matter?



The Data Spectrum

Small / Medium / Big data

Personal / Commercial / Government data

Internal access

Employment contract + policies

Sales reports

Named access

Explicitly assigned by contract

Driving licences

Group-based access

Via authentication

Medical research

Public access

Licence that limits use

Twitter feed

Anyone

Open licence

Bus timetable

It would matter if we pretended that data is anonymous when it isn't

Closed

Shared

Open



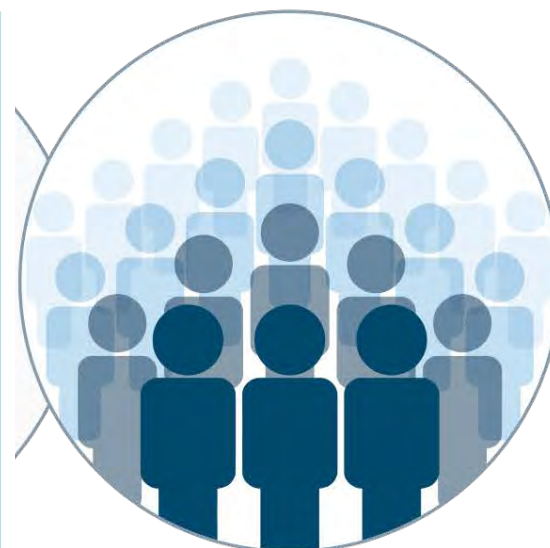
theodi.org/data-spectrum

1. Anonymised data

This is expected research output –
tables and figures in posters and articles

NIHR has an Open Access policy –
available to all without cost

Expect **1000's** of views per publication



- Anonymised*

Anonymous

Understanding patient data, please visit goo.gl/wwNyYp

2. De-personalised data

Researchers apply
for access

**Expect 100's
of applications
from researchers**



De-personalised

**For every study the
process is transparent**

Step 1 Approval process

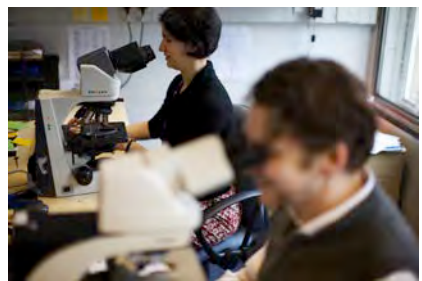
Step 2 Appraisal Criteria

Step 3 List approved
applications

Step 4 Publish Researchers
Findings

Information Commissioner Office's
guide to anonymisation, please visit: goo.gl/WLvWto

Example criteria for reviewing applications for data



- From bona fide researcher
- Negligible risk to participants
- Meets consent criteria
- Overall study ethics are approved
- Participants are not surprised a study was approved
- Results must not upset participants
- Negligible risk of harm to the overall study
- A plain language summary

Please see: <https://www.metadac.ac.uk/>

Digression: Who are all these researchers?



People who collect all the data

Who are all these researchers?



People who pull the data together



Who are all these researchers?

A horizontal bar composed of several colored segments: green, orange, purple, red, and blue.

People who look for the one bit that's different

A background of puzzle pieces. Most are white with black outlines, but one piece in the center is solid green. The pieces are arranged in a grid-like pattern.

Knowing what to ignore is the clever part

Who are all these researchers?

People who are trying to put **lots** of jigsaws together



Who are all these researchers?

Anyone!



**We expect patient groups
to be asking for data**

3. Personally identifiable data



Personally identifiable

Main purpose of **NIHR BioResource** – to allow recall to “experimental medicine” studies:

- Studies must be ethically approved
- Potential participants identified on basis of data, samples held
- Participants are asked if they would like to be contacted
- If so, approached for new study

The process is always published and transparent

We expect 10's of applications

Commercial access

Not everyone accepts drug company involvement

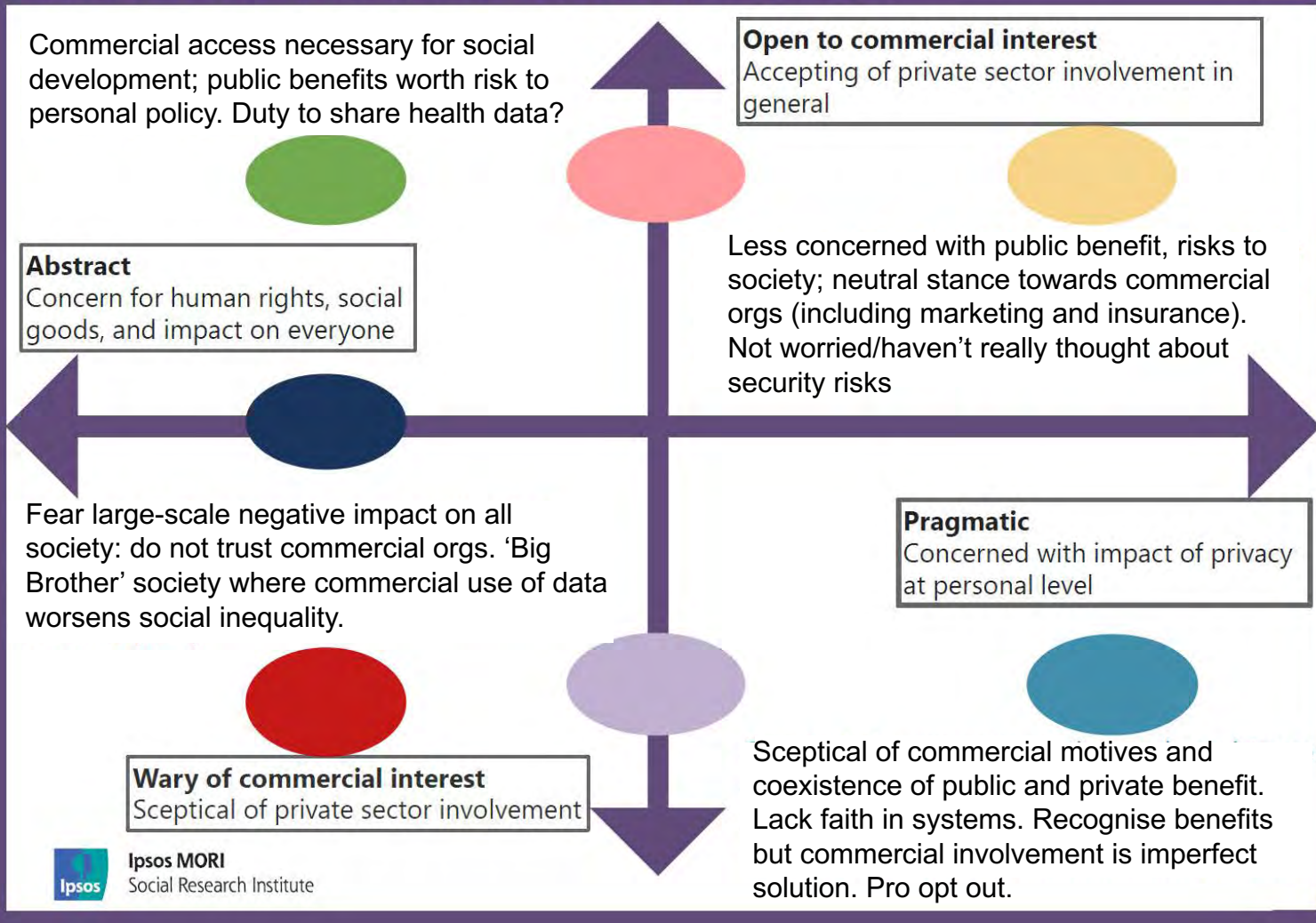
**Wary of
commercial
interest**

**Concerned
about privacy**



Ipsos MORI survey (2016)

Seven mindsets influence views



The full report can be found

<https://goo.gl/SsYT7e>

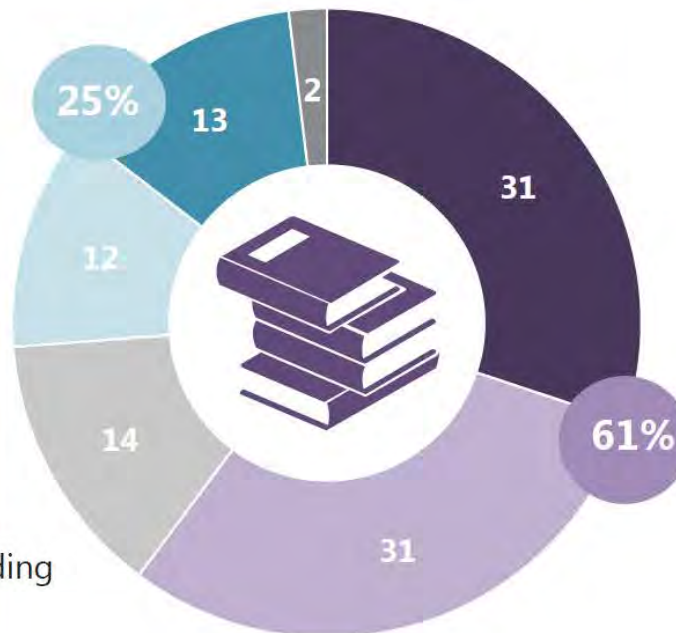
Is there support for commercial access to data?

Support for commercial access if research at risk

Which of the following statements comes closest to your view of health data being shared with commercial organisations?*

A. I would not want commercial organisations to have access to anonymised health data, even if this means the research does not take place

B. The research should be conducted by commercial organisations if there is a possibility of new treatments for diseases being developed



- Agree much more with B than with A
- Agree a little more with B than with A
- Agree equally with both / don't agree with either
- Agree a little more with A than with B

Base: 974 GB adults, aged 16+



*See report for full question wording

source: Ipsos MORI Social Research Institute

Is there an alternative to sharing data?

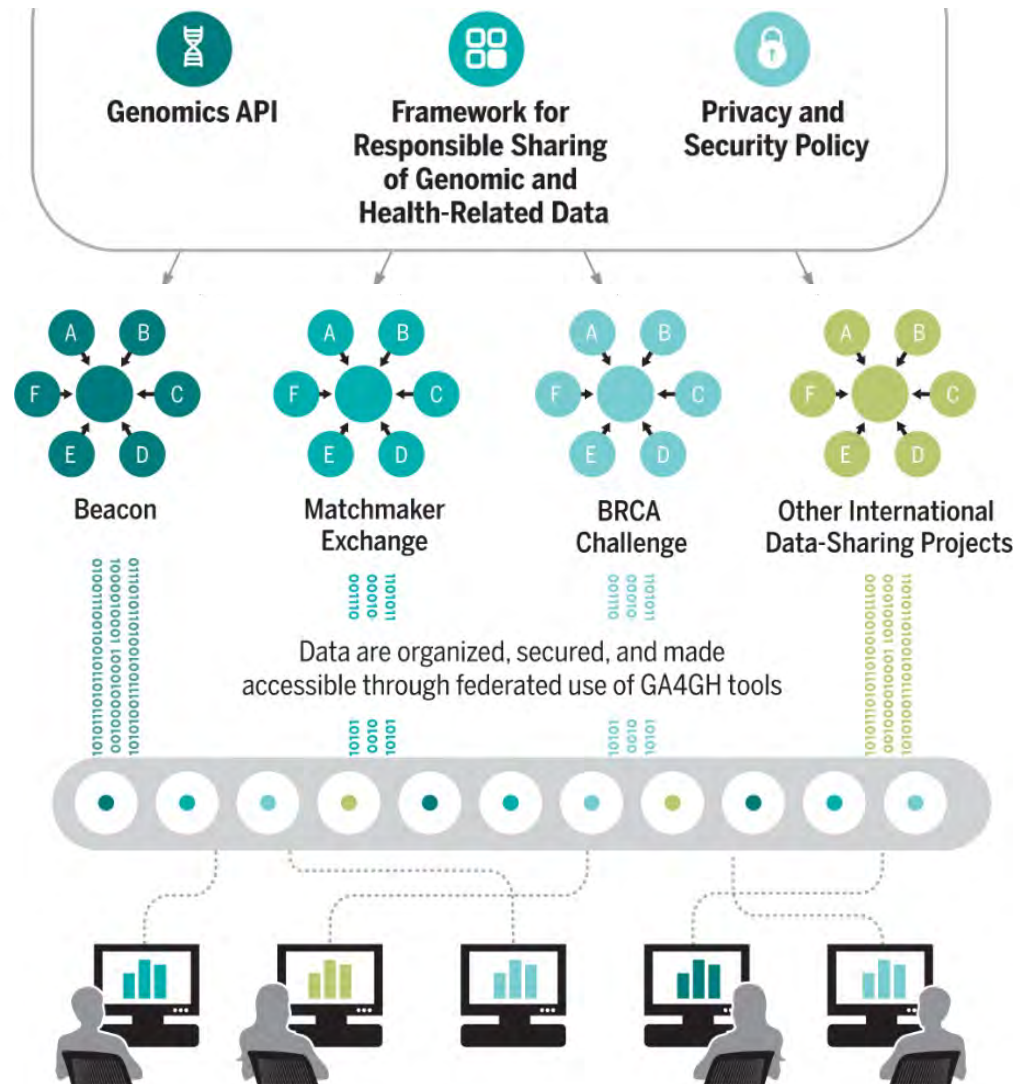


Global Alliance
for Genomics & Health
Collaborate. Innovate. Accelerate.

Technical fix: researchers ask each other whether they have data, without having to see it

Article on a federated data ecosystem for sharing genomic, clinical data

<https://goo.gl/HgmtiB>



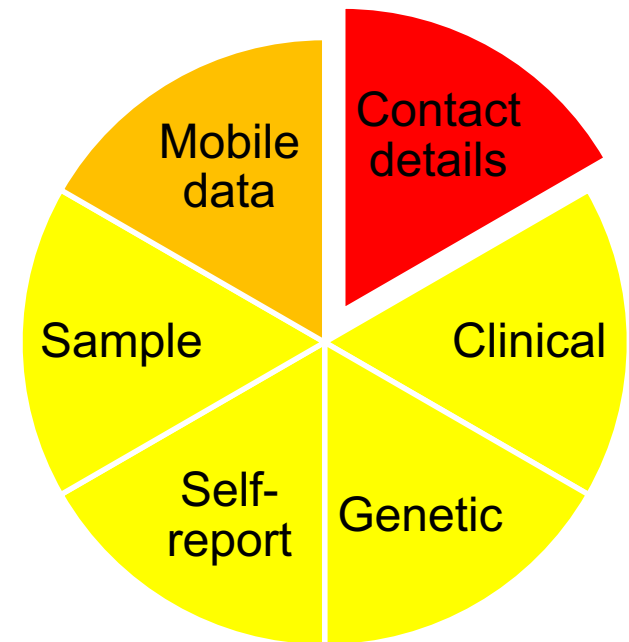
Loose ends

Where is data stored and who can see it?

- Contact details are seen by recruitment team e.g. at clinics by research nurses
- Other details are seen by other staff
- Data is linked only by a key

For NIHR BioResource:

- All staff have NHS research passports and training
- All but genetic data held in a NHS -approved secure data centre
- Genetic data (which is much bigger) is held in a secure research data centre



Loose ends 3

What's next?

Data from mobile apps

An example

MitoAction iPhone App

Now available for free in the App Store!



Here are some of the amazing things you can do with the MitoAction App:

- Track common Mito symptoms
- Record daily fluid and food intake
- Manage your Mito Cocktail and other medications
- Search for Mito terms in the comprehensive glossary
- Link to MitoAction's website and Facebook page



This app was developed through our generous partnership with David Aaron and his company, App Creatr LLC (www.appcreatr.me).

Useful resources

Policy and discussion sites:

[Understanding Patient Data](http://bit.ly/2ocNwbU) (data re-use advocates): <http://bit.ly/2ocNwbU>

[Wellcome Trust](http://bit.ly/2tP9Dd9) (medical research charity): <http://bit.ly/2tP9Dd9>

[MRC](http://bit.ly/2x1fSeq) (government-funded medical research): <http://bit.ly/2x1fSeq>

[PHG Foundation](http://bit.ly/1KbjsAd) (independent genetics think tank): <http://bit.ly/1KbjsAd>

Successful data re-use sites:

[UK Data Archive](http://bit.ly/2lq1krm) (social science): <http://bit.ly/2lq1krm>

[European Genome-phenome Archive](https://ega-archive.org/) (genomics): <https://ega-archive.org/>

[METADAC](https://www.metadac.ac.uk/) (mixed genetic and social science): <https://www.metadac.ac.uk/>

Technical initiatives to avoid moving data:

[Global Alliance for Genomics & Health](http://bit.ly/2lrNEMo) (international technology lead): <http://bit.ly/2lrNEMo>

[Repositive](https://repositive.io/) (data finder service): <https://repositive.io/>

Data sharing failures:

[PACE trial](http://bit.ly/2cMoq01): <http://bit.ly/2cMoq01>

[Google DeepMind](http://bit.ly/2siRYp4): <http://bit.ly/2siRYp4>

Questions?

