Revealing the hidden challenges of making health research reusable
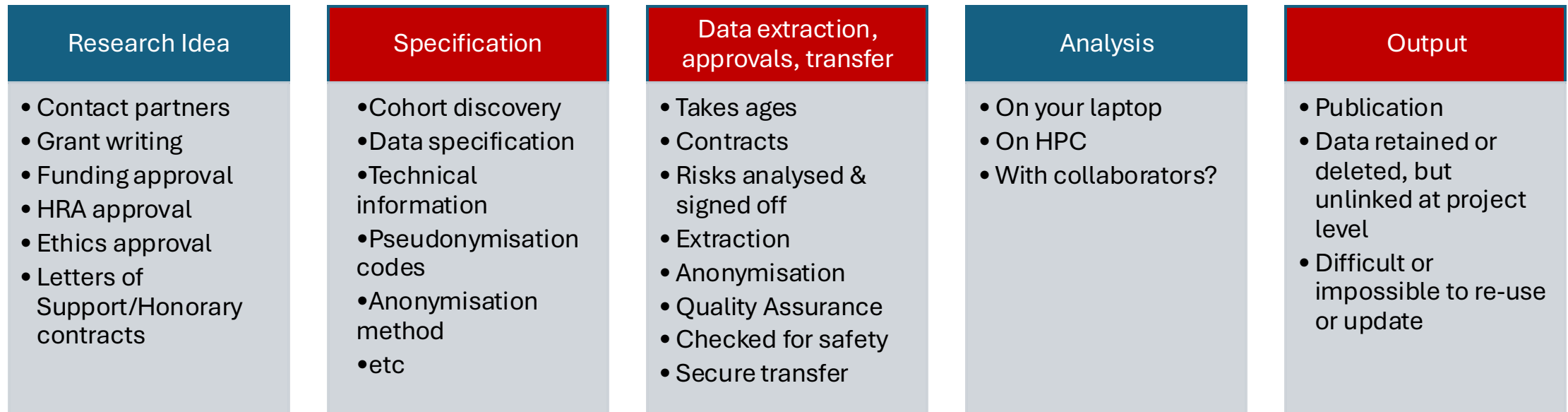
Will Chapman, Jean Golding Institute

BRC funded, JGI and SWSDE

Multiple pain points threaten research project timelines and later data reuse

Research Idea	Specification	Data extraction, approvals, transfer	Analysis	Output
• Contact partners • Grant writing • Funding approval • HRA approval • Ethics approval • Letters of Support/Honorary contracts	•Cohort discovery •Data specification •Technical information •Pseudonymisation codes •Anonymisation method •etc	• Takes ages • Contracts • Risks analysed & signed off • Extraction • Anonymisation • Quality Assurance • Checked for safety • Secure transfer	• On your laptop • On HPC • With collaborators?	• Publication • Data retained or deleted, but unlinked at project level • Difficult or impossible to re-use or update

Multiple pain points threaten research project timelines and later data reuse

Research Idea	Specification	Data extraction, approvals, transfer	Analysis	Output
• Contact partners • Grant writing • Funding approval • HRA approval • Ethics approval • Letters of Support/Honorary contracts	•Cohort discovery •Data specification •Technical information •Pseudonymisation codes •Anonymisation method •etc	• Takes ages • Contracts • Risks analysed & signed off • Extraction • Anonymisation • Quality Assurance • Checked for safety • Secure transfer	• On your laptop • On HPC • With collaborators?	• Publication • Data retained or deleted, but unlinked at project level • Difficult or impossible to re-use or update

Solutions are reinvented for each project and often person-specific.

Purpose of this session

- We want to understand the current process(es) for getting routine data for research from our local NHS
  - "Easy data" that is fast to specify and extract
  - Structured data that needs transformation before transfer
  - Multi-modal data that can't be created through the right query, eg images, vitals

- We want to understand what it took to get (routine) data for health research via the Trusts
  - Which parts of the process **went well**, with good support from UoB/NHS?
  - Which parts of the process were **lengthy or painful**?

- This will feed into a project to streamline health data pipelines for both the SDE and our usual research platforms

Purpose of this session

- We want to understand the current process(es) for getting routine data for research from our local NHS
  - "Easy data" that is fast to specify and extract
  - Structured data that needs transformation before transfer
  - Multi-modal data that can't be created through the right query, eg images, vitals

- We want to understand what it took to get (routine) data for health research via the Trusts
  - Which parts of the process **went well**, with good support from UoB/NHS?
  - Which parts of the process were **lengthy or painful**?

- This will feed into a project to streamline health data pipelines for both the SDE and our usual research platforms

Purpose of this session

- We want to understand the current process(es) for getting routine data for research from our local NHS
  - "Easy data" that is fast to specify and extract
  - Structured data that needs transformation before transfer
  - Multi-modal data that can't be created through the right query, eg images, vitals

- We want to understand what it took to get (routine) data for health research via the Trusts
  - Which parts of the process **went well**, with good support from UoB/NHS?
  - Which parts of the process were **lengthy or painful**?

- This will feed into a project to streamline health data pipelines for both the SDE and our usual research platforms

Purpose of this session

- We want to understand the current process(es) for getting routine data for research from our local NHS
  - "Easy data" that is fast to specify and extract
  - Structured data that needs transformation before transfer
  - Multi-modal data that can't be created through the right query, eg images, vitals

- We want to understand what it took to get (routine) data for health research via the Trusts
  - Which parts of the process **went well**, with good support from UoB/NHS?
  - Which parts of the process were **lengthy or painful**?

- This will feed into a project to streamline health data pipelines for both the SDE and our usual research platforms

Questions for discussion:

Have you requested routine structured or unstructured data from your local NHS trust? More than once?

At which stages did you need the most help?
- Proposal development
- Data specification
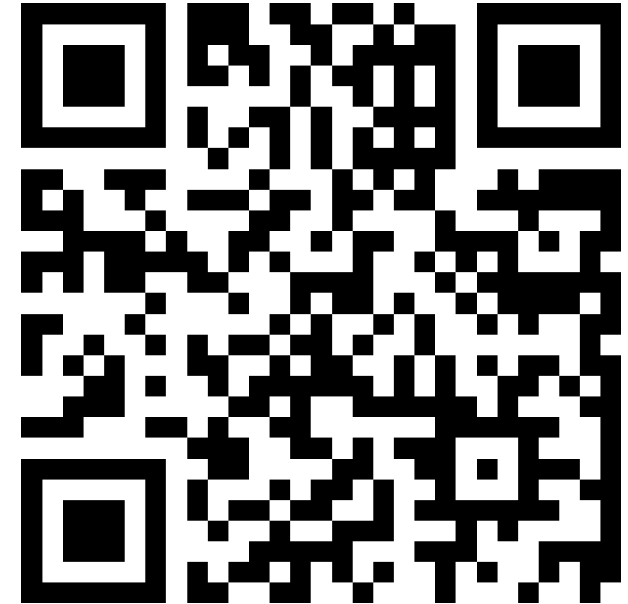- Data extraction and transfer
- Analysis and outputs

What kind of assistance did you need?
- Clinical knowledge
- Technical/systems knowledge
- Organisational knowledge

What is essential/desirable for a reproducible process?
- People as well as systems

What or who made it easier?

Sli.do #3443597