Providing greater access to clinical trial data for further research

Robert Frost, Medical Policy
May 2015

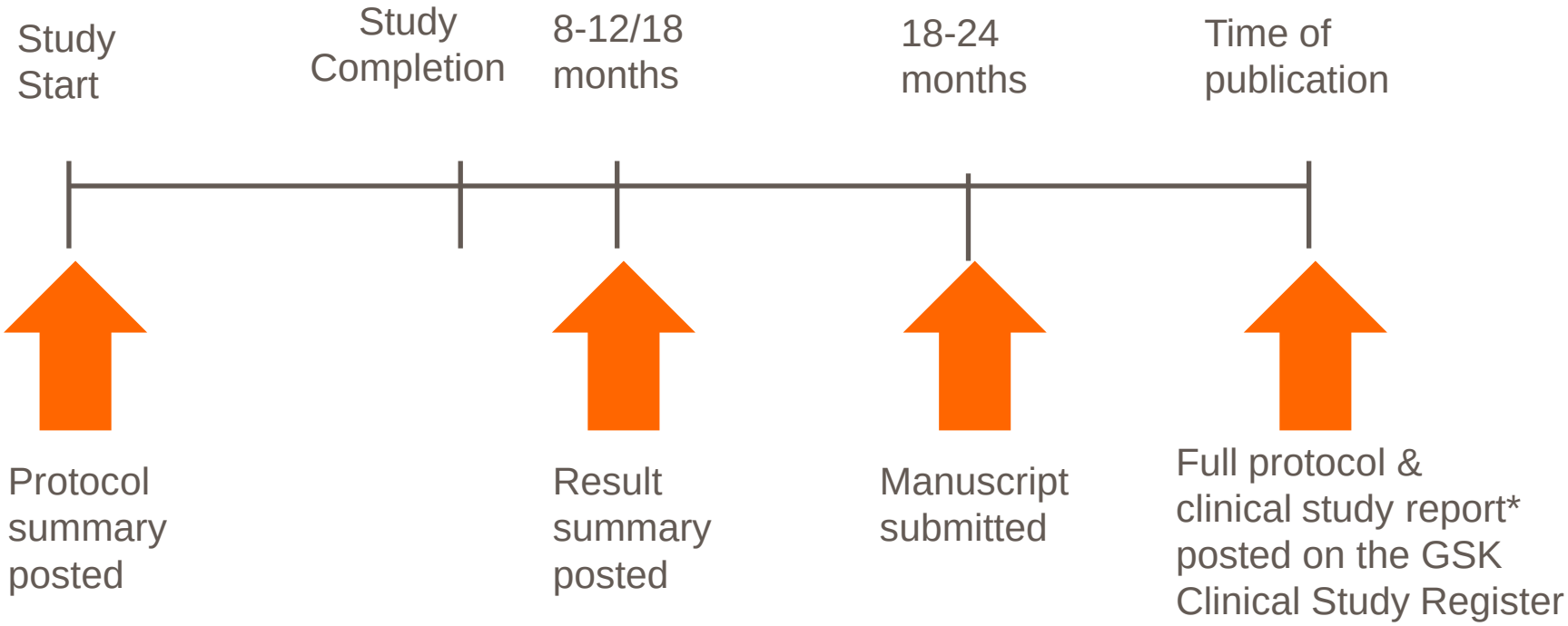
Content

1. Proving greater access to clinical trial data: benefits and challenges
2. Applying a range of safeguards
3. Delivering practical solutions for data anonymisation
4. Enabling further research to advance medical science and improve patient care

1. Providing greater access to data: benefits and challenges

Public Disclosure of Clinical Trial Information

All human subject research studies that evaluate investigational or approved medicinal products (phase I-IV, meta-analyses, observational studies)



* CSR posted after approval or termination of the medicine

So Why are We Doing More?

Result summaries and publications have limitations

- Summarise data from the study population with statistics to compare treatment groups.
- Do not include the data from each research participant.

Primary Efficacy Results: Total Population			
Emetic Episodes Day 1 To Day 5	**Dose 1**	**Dose 2**	**Dose 3**
Complete (0 Episodes)	7 (19)	8 (22)	10 (31)
Major (1-2 Episodes)	10 (28)	14 (39)	10 (31)
Minor (3-5 Episodes)	0	1 (3)	0
Failure (>5 Episodes/Rescued)	19 (53)	13 (36)	12 (38)
Dose 2 vs Dose 1	0.848		

Benefits

- Enables further research to benefit medical research and patient care.
- Enables the review of results from individual clinical trials to validate the results.
- Helps avoid duplication of research, unnecessarily enrolling patients into clinical trials and exposing them to possible risks.
- Strengthens trust in clinical research through enhanced openness and transparency.

Main Issues

- Protecting the privacy and confidentiality of research participants.
- Ensuring the data are used for valid scientific investigation.
- Practicalities of anonymising data and providing data in ways that enable external researchers to understand and navigate the information.

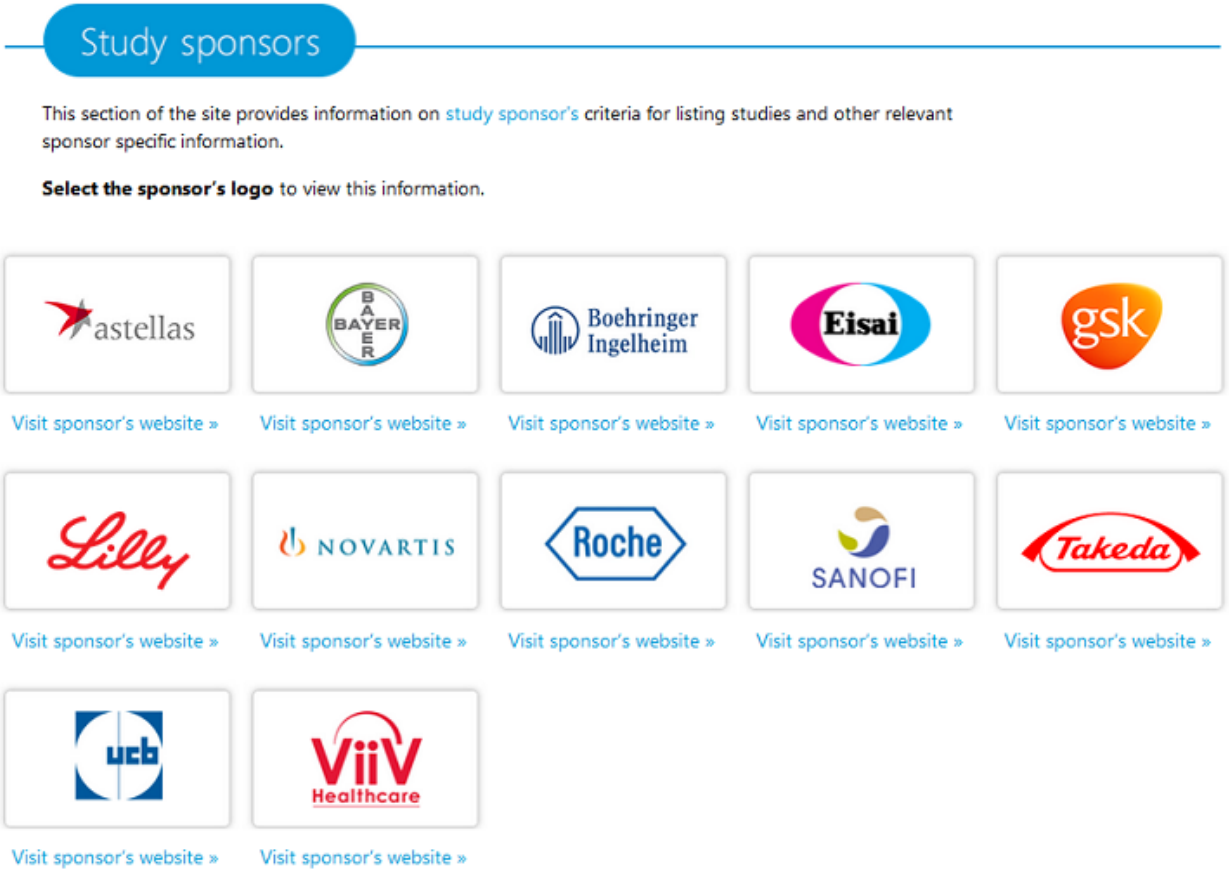
2. Protecting privacy and scientific rigour

Our Vision

The aim is to help realise a broad, independent solution to allow access to data from clinical trials conducted by multiple companies and organisations.

The Clinical Study Data Request System

- GSK took a first step in 2013 and in Jan 2014 the multi-sponsor site was launched: www.clinicalstudydatarequest.com



Scientific review and secure access to data

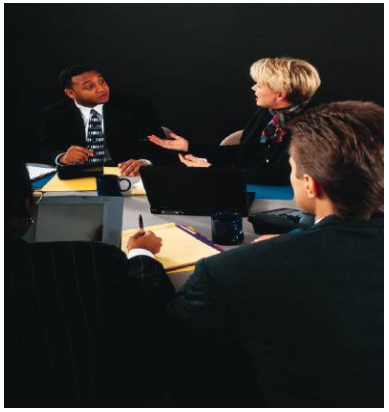
– Three components:

Multi-sponsor request site

1

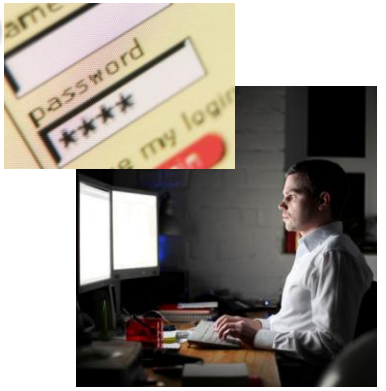
Independent review panel

2



Secure access system

3



– In 2015 the Wellcome Trust took over responsibility for managing the review of research proposals and the operation of the IRP.

Applying a range of safeguards

- **Independent review of research proposals and qualifications of research team**

- **Data sharing agreement:** which limits the purpose for which data can be used and prohibits the re-identification of individuals.

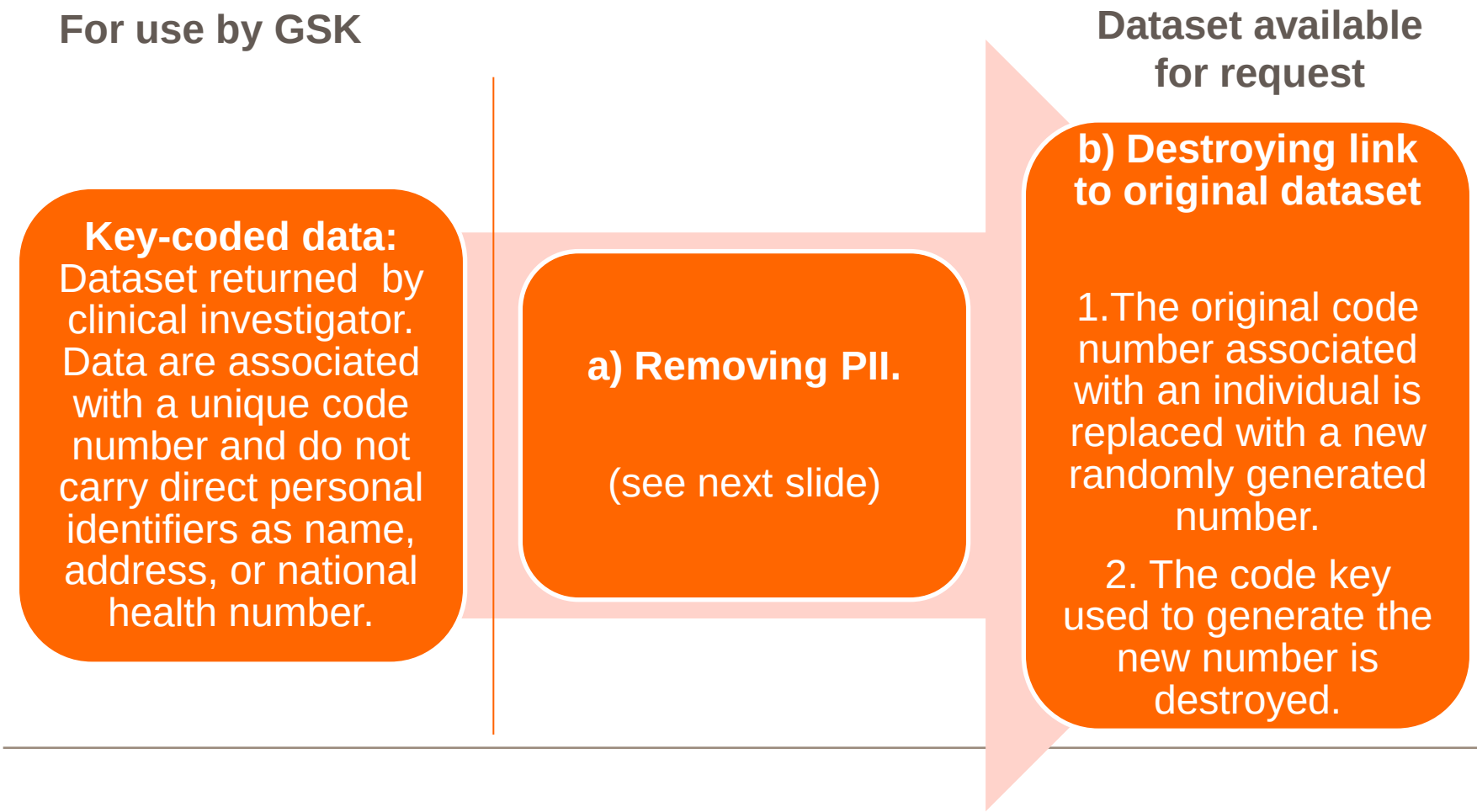
- **Data provided in secure environment:** Researchers are provided with access to anonymised data in a secure environment in which they conduct their analysis. Protections are in place to prevent the removal of raw data.

- **Access is provided to anonymised data**. Anonymisation involves removing personally identifiable information from the dataset and destroying the link between the dataset that is provided and the original dataset. (detail in later slides).

3. Practical solutions to data anonymisation

Overview

- GSK's approach to data anonymisation is publically available:
  - https://www.clinicalstudydatarequest.com/Study-Sponsors-GSK-Details.aspx



Removing personally identifiable information

- The 18 identifiers defined by HIPAA are removed from the datasets (and related documentation). In addition any other PII that may be present is removed.

- **This involves the following steps**:
  - Recoding identifiers (or code numbers)
  - Removing any names e.g. Investigator names, laboratory names
  - Removing free text verbatim terms
  - Replacing date of birth with age.
    - With the exception of all ages above 89 which are aggregated into a single category of “90 or older”
  - Replacing all original dates relating to individual subjects with randomly generated offsets (+/- 360) which are then applied to create ‘dummy dates’.
  - Review and removing other PII.

Transcelerate

- GSK approach is consistent with the approached developed by Transcelerate, a non-profit organization of biopharmaceutical companies
  - http://www.transceleratebiopharmainc.com/assets/clinical-data-transparency/

- “Both the Safe Harbor and Expert Determination approaches start with a shared principle of identifying direct (e.g. ID number) and quasi (e.g. date of birth and date of death) identifiers, and applying de-identification techniques.”

- “In the context of providing data in a secure controlled access model where data requests are reviewed and subject to data sharing agreements, data providers may decide that a statistical assessment of the risk of re-identification (a key part of the Expert Determination approach) is not necessary in most cases”.

- “For this reason, the approach outlined in this paper is based primarily on the enhanced Safe Harbor method. For certain datasets or situations (e.g. rare diseases/small populations) sponsors may choose to employ expert determination methods to supplement that approach”

Approaches to data anonymisation

- Data providers, researchers and data protection authorities need to work together on simple solutions that can be routinely applied.

- Solutions need to:
  - Minimise risks and deliver benefits of greater access to data.

  - Be publically available in full so that they can be tested and independently applied.

  - Be proportionate and scalable:
    - So they can be applied independently by a small, academic group with limited internal expertise and resource.
    - So they can be embedded as a routine part of the clinical trial process for a large organisation wanting to make data available from 1000s of trials.

4. Enabling further research

Facilitating further research

- https://www.clinicalstudydatarequest.com/Metrics.aspx

- No. of Research Proposals submitted up to 31 March 2015 = **119**
- Met requirements = **91**

Independent Panel review

- Approved or approved with conditions = **83**
- Rejected or advised to re-submit **= 6**

Facilitating new research

- Increasing number of ‘joint’ requests involving data from multiple companies.

- Novel research includes:
  - **Risk factors and Biomarkers** e.g impact of tumour size on survival.
  - ‘**Methodological’** e.g. development of toxicology predictive tools.
  - **Comparing treatment regimens** e.g. meta-analysis of epilepsy treatments
  - **Optimising treatments** e.g. optimising activated clotting time to avoid bleeding complications
  - **Treatment or patient stratification** e.g. differential effects in ethnic groups
  - **Co-morbidities/Lifestyle factors** - relationship between flu vaccination and lifestyle factors
  - **Re-analysis of GSK study** – effectiveness of Paxil on treatment of adolescent depression
  - **Informing future research design** – understanding expected frequency of Severe Adverse Events before designing a new trial in the developing world.

Conclusion

Meeting expectations and delivering practical solutions

- All Trials campaign: ~ 85,000 people and 570 organisations calling for greater trial transparency [“...not calling for individual patient data to be made publicly available”]

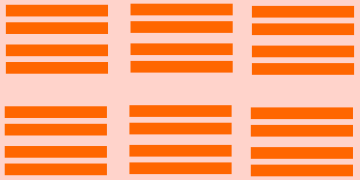
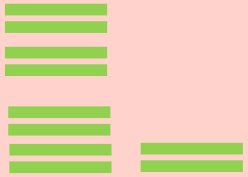
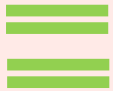
- **European Patients Forum** “We are committed to continuing our collaboration with EMA to find a solution [to access to individual level data] that benefits transparency and is acceptable to patients (Oct. 2014)]

- Trend towards **journals** mandating that researchers submitting clinical trials for publication must make anonymised patient level data available on reasonable request (e.g. BMJ).

- Independent reports including the **Institute of Medicine: “**...foster a culture in which data sharing is the expected norm...responsible strategies aimed at maximizing the benefits, minimizing the risks, and overcoming the challenges of sharing clinical trial data...”

What is the right balance?

- Each data provider (e.g. pharma, biotech, academia) may have a certain resource that can be taken from other activities to implement data sharing and data anonymisation.
- Methods need to minimise risk but also deliver benefits of providing access to data.

Anonymisation method	Resource to prepare data	Minimise privacy risks	Trials with anon. data available for external research	Scientific and medical benefits
One				
Two				
Three				

Conclusion

- Approaches to protect privacy need to consider a range of safeguards, above and beyond how data are anonymised.
- Solutions to data anonymisation need to be practical, methods need to be publically available and appropriate for use across the research community.
- Providing greater access to data can enable further research to advance medical science and improve patient care.
- We have adopted an approach to provide access to clinical trial data where risks to individual privacy are minimised and the benefits of providing access to data for further research are maximised.

Thank you

robert.x.frost@gsk.com

Background slides

GSK's commitment

- All interventional trials started from January 1 2013.
- Global interventional trials that were ongoing or started after the formation of GSK in 2000. This listing will complete in 2015.
- Trials are listed after the primary manuscript has been accepted for publication AND the medicines has been approved / terminated.
- Currently over 1300 GSK trials are listed for request.
- Researchers can enquire about the availability of data from other non-listed studies.

What information do we make available?

The following anonymised data and documents (with personally identifiable information redacted) are made accessible in the secure system:

- Raw dataset
- Analysis-ready dataset
- Protocols with any amendments
- Annotated case report form
- Reporting and analysis plan
- Dataset specification
- Redacted Clinical Study Report including modular appendices
  (potentially identifiable information, including patient level data and patient narratives are removed)