Data Transparency – The Six W's

Natalie (Reynolds) Martinez
Eli Lilly And Company
PhUSE Upper Providence SDE
16-Jul-2015

Disclaimer

Note: The opinions in this presentation are those of the presenter and may not necessarily reflect the views of Eli Lilly and Company, PhUSE, or TransCelerate.

Abstract

Data Transparency and access to patient level data has come to the forefront with EMA's policy on Publication of Clinical Data and PhRMA/EFPIA's Principles for Responsible Clinical Trial Data Sharing.

This presentation discusses data de-identification and anonymization of Individual Patient Data in Clinical Studies, as well as discusses the corresponding data documentation provided by sponsor companies.

The Six W's of Investigation

*"If you can answer: **what**, **why**, **who**, **when**, **where** and **how**; you will have a clear and fundamental knowledge of the whole situation."*

- Stine Bergholtz, PhD

Source: http://www.brainspores.com/the-six-ws-of-investigation/

The Six W’s of Data Transparency

- What
- Why
- Who
- When
- Where
- How

What/When - **EMA Draft Policy 70**

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

1 24 June 2013
2 EMA/240810/2013
3 Executive Director

4 **Publication and access to clinical-trial data**

5

6 POLICY/0070
7 Status: Draft for public consultation

http://www.ema.europa.eu/docs/en_GB/document_library/Other/2013/06/WC500144730.pdf

What/When - PhRMA/EFPIA Commitment

July 2013 – PhRMA joined EFPIA (European Federation of Pharmaceutical Industries and Associations) in adopting joint **Principles for Responsible Clinical Trial Data Sharing**

The biopharmaceutical industry is firmly committed to enhancing public health

through responsible sharing of clinical trial data in a manner that is consistent with the following Principles:

- **Safeguarding the privacy of patients**
- **Respecting the integrity of national regulatory systems**
- **Maintaining incentives for investment in biomedical research**

1 Jan 2014 - Biopharmaceutical companies will make **publicly available**, at a minimum, **the synopses of clinical study reports (CSRs) for clinical trials** in patients submitted to the Food and Drug Administration (FDA), European Medicines Agency (EMA), or national competent authorities of EU Member States

Why – EMA Policy 70

- Published 02-Oct-2014
- Effective 01-Jan-2015

"A high degree of transparency will take regulatory decision-making one step closer to EU citizens, and promote better-informed use of medicines. In addition, the Agency takes the view that access to clinical data will benefit public health in future."

- Phase 1: Publication process for clinical reports
- Phase 2: Review aspects of providing individual patient data

Why – Industry

- Enhance Public Health
- Further legitimate scientific and medical research
- Increase goodwill

Who – Biopharmaceutical companies

Top 1-12 Pharmaceutical Companies, Ranked by 2013 Market Capitalization

Company	Clinical Trial Data Available for Sharing
Janssen Pharmaceuticals	Data from interventional trials for products and indications approved in both the EU and US (older studies may be difficult to access)
Pfizer	Global trials that ended after Sept 2007 Trials for which results are posted on ClinicalTrials.gov
Novartis	Jan 2014 – present Phase II – III trials for new medicines or indications approved by the EU and US regulatory agencies
Roche	Jan 1999 – present Trials from terminated programs Phase II – III trials, and some Phase IV trials, for new medicines that have been approved by the EU and US regulatory agencies
Sanofi	Jan 2014 – present Trials approved by EU and US regulatory agencies and accepted for publication
Merck	Sept 2007 – present Trials approved by EU and US regulatory agencies and accepted for publication Trials for which results are posted on ClinicalTrials.gov
GlaxoSmithKline	Dec 2000 – present All global interventional clinical studies
Bayer	Jan 2014 – present Data on new medicines and indications that have been approved by EU and/or US regulatory agencies without plans for further regulatory review or submissions CSR summaries are publically available back to 2005
Bristol-Myers Squibb	Jan 2008 – present Phase I- IV interventional trials for medicines and indications approved in the US and the EU
AbbVie	All data for medicines and indications approved in the US and the EU
Eli Lilly	1999 – present Interventional clinical studies for approved indications of medicines on the market in the US and EU
AstraZeneca	Not stated online

Source: http://www.forbes.com/sites/harlankrumholz/2014/06/02/time-to-assess-pharma-progress-in-data-sharing/

Where and How

- Many using multi sponsor request system: ClinicalStudyDataRequest.com

> Study sponsors who have committed to use this site are **Astellas, Bayer, Boehringer Ingelheim, Eisai, GSK, Lilly, Novartis, Roche, Sanofi, Takeda, UCB** and **ViiV Healthcare**.

- JnJ/Janssen using YODA (Yale Open Data Access)
- Bristol Myers Squib using DCRI (Duke Clinical Research Institute)

Where/How - Clinicalstudydatarequest.com

Content to be supplied:

- Anonymized Raw dataset
- Anonymized Analysis-ready dataset
- Protocols with any amendments
- Annotated case report form
- Reporting and analysis plan
- Dataset specifications
- Redacted Clinical study report

Sponsor Tools & Systems

Process:

- Pieces of package assembled in one location
- In SAS sharing process – will upload to SAS Multi Sponsor Environment

How: de-identify/anonymize datasets

- **De-identified** protected health information (PHI) is defined in the HIPAA Privacy Rule, Code of Federal Regulations, 45CFR164.514, as 'Health information that does not identify an individual and with respect to which there is no reasonable basis to believe that the information can be used to identify an individual is not individually identifiable health information.' Thus de-identifying the data includes removing or recoding identifiers, removing or redacting free text verbatim terms, and removing explicit references to dates. Participants' identification code numbers are de-identified by replacing the original code number with a new random code number.

- **Anonymization** is a step subsequent to de-identification that involves destroying all links between the de-identified datasets and the original datasets. The key code that was used to generate the new identification code number from the

How: de-identify/anonymize datasets

- TransCelerate white paper (May 2015):
  - Data De-identification and Anonymization of Individual Patient Data in Clinical Studies – A Model Approach

- PhUSE De-Identification Standard for SDTM 3.2 (May 2015)
  - Each of the domains and variables that can potentially hold "Personally Identifying Information" (PII) are:
    1. Rated in terms of impact on data privacy,
    2. Allocated rules to apply, and
    3. The rationale and impact on data utility is recorded

How: TransCelerate Model

Attribute	Approach to De-identification
Identification Number (eg, participant, investigator, site, laboratory)	Consistently replace these original identifiers with new randomly generated identifiers. In order to maintain the relationship between participants and investigators/sites, all participants from one investigator/site should be assigned the same random investigator/site identifier in the de-identified dataset(s). For identifiers other than participant numbers: If considered to jeopardize data privacy, can set to blank, set to '--redacted--', or aggregate.
Names (eg, participant, investigator, site, contractor, supplier, vendor, company staff)	Set to blank or remove.
Contact Information (eg, participant, investigator, site, vendor - Postal Address, Phone, Fax, Email)	Set to blank or remove.
Country	Retain as in the original dataset unless this is considered to jeopardize data privacy. If so, set to blank, set to '--redacted--', or aggregate (data driven decision).
Dates	Consistently apply either "Offset Date method" or "Relative Study Day" method.
Participant Date of Birth	Set to blank or remove.
Age	• Display exact age if <=89 years, and set to blank if >89 years • Display age category - either as '<=89' and '>89', or sub-categorizing '<=89' into multiple age bands eg, <25, 25-29, 30-34,…,85-89. (Age category '>89' must not be sub-divided).
Adverse Event / Medical History / Concomitant Medication / Procedure	Set verbatim level terms to blank or remove. Dictionary licenses allowing, provide coded terms eg, ideally at least the lowest level term (LLT) for MedDRA-coded adverse events and diseases; and trade name and

How: TransCelerate Model

Attribute	Approach to De-identification
	ingredients for medications coded using WHO Drug.
Other free text fields (eg, comment fields)	In general, set to blank or remove. If considered important for retention, set records (or parts thereof) containing privacy information to '--redacted--'.
Other indirect identifiers (eg, rare disease or rare adverse event, extreme values, unusual treatment)	Retain unless considered to jeopardize participant data privacy. If so, set to blank, set to '--redacted--', or aggregate as needed.

Where/How - Clinicalstudydatarequest.com

Datasets and documents provided	Where available, the following anonymised patient level data and information is provided for each clinical study. **Raw dataset.** This is the data collected for each patient in the clinical study. **Analysis-ready dataset.** This is the dataset used for Lilly's analysis. **Protocols with any amendments or addenda.** This describes the objectives, design, methodology, statistical considerations, and organisation of a clinical study. **Annotated case report form.** This is a blank case report form with descriptions of the data collected, metadata and how they are described in the dataset. If an annotated case report form is unavailable, a blank case report form may be provided. **Reporting and analysis plan.** This describes methods of analysis, procedures for data handling and data displays (figures and tables) Lilly used for the study. **Dataset specifications.** This is the meta-data which describes the datasets e.g., variable labels, variable descriptions, formats. **Clinical study report.** This is the report of efficacy and safety data from the study. It forms the basis of submissions to regulatory authorities such as the US FDA and the European Medicines Agency (EMA). Appendices which include patient level data are not included as these data are provided in the datasets Lilly provides. To protect research participants' privacy and confidentiality, case narratives are not routinely included. They may be provided where they are needed for a specific research proposal, provided research participants' privacy can be protected.

← Sponsor Page

Study Page:

Datasets and Documents Available for this Study

☑ Analysis-ready dataset ☑ Reporting and analysis plan ☑ Clinical study report ☑ Raw dataset

☐ Annotated case report form ☑ Dataset specifications ☑ Protocol with any amendments

Data Documentation Summary Sponsor page

Company	Annotated CRF	Dataset Specifications	SAP (or equivalent)
Astellas	X	X	X
Bayer	X	X	X
Boehringer Ingelheim	X	X	X
Eisai	X	X	X
Eli Lilly	X	X	X
GlaxoSmithKline	X	X	X
Novartis	X	X	X
Roche	X	X	X
Sanofi	X	X	X
Takeda	X	X	X
UCB	X	X	X
ViiV	X	X	X

Source: www.clinicalstudydatarequest.com

Data Documentation Summary
Study Page (box checked only if all CSDR studies provided the data documentation artifact)

Company (# studies)	Annotated CRF	Dataset Specifications	SAP (or equivalent)
Astellas (11)	X	X	X
Bayer (0)	n/a	n/a	n/a
Boehringer Ingelheim (266)			
Eisai (0)	n/a	n/a	n/a
Eli Lilly (164)	blank CRF, if available	X	
GlaxoSmithKline (1371)			
Novartis (6)	X	X	X
Roche (119)		X	
Sanofi (2)	X	X	
Takeda (353)	X	X	X
UCB (21)		X	X
ViiV (41)	X	X	X

Source: www.clinicalstudydatarequest.com, June 2015

Conclusion

Recommendation:

1. Provide as much documentation to researchers as possible to avoid future questions
   - Annotated or blank CRF
   - Complete Define document (define.xml)
   - Analysis Plan
2. Create documentation at end of study
   - Supply for both FDA (original) and EMA (transparency) submissions

Questions?

- Thank you PhUSE for the opportunity to present at the 2015 Upper Providence SDE!

Contact information

Natalie Martinez

nmartinez@lilly.com