Clinical Trial Data Sharing and Reuse
A New Reality for Researchers

Ida Sim, MD, PhD, Professor of Medicine, UCSF and Vivli Co-Founder
Rebecca Li, Vivli Executive Director
Faculty Co-Director of Research Ethics, Harvard Center for Bioethics
Harvard Medical School

October 2019
Agenda

1. Why share data?
2. What is the role of data standards in data sharing and reuse?
3. The Vivli Global Platform for Clinical Data sharing and Reuse as a use case
1. Why Should We Share Our Clinical Research Data

- Funder requirements
- Journal requirements (new as of 2018)
- Publicly stated commitments for industry (BIO, EFPIA, PhRMA)
- Drive new science (integrate data to drive new insights faster)
- Ethical obligations to trial participants
- Enhance and advance your career
What are Journals Requiring as of July 1, 2018?

- Trial manuscripts must be submitted with a data sharing statement
  - Must describe how you will share your Individual participant-level data (IPD), including who, what, when, where, and why
- IPD sharing is not (yet) required but “editors may take into consideration data sharing statements when making editorial decisions”

Declaring Your Data Re-use Plans as part of the Trial Registration Record... before the 1st patient is enrolled

- Data sharing plan is part of the ClinicalTrials.gov registration record
- As of 1 January 2019, ICMJE requires registration of your data sharing plan at time of trial registration.

**12. IPD Sharing Statement**

**Plan to Share IPD**
Definition: Indicate whether there is a plan to make individual participant data (IPD) collected in this study, including data dictionaries, available to other researchers (typically after the end of the study). Select one.
- Yes: There is a plan to make IPD and related data dictionaries available.
- No: There is not a plan to make IPD available.
- Undecided: It is not yet known if there will be a plan to make IPD available.

Note: Undecided is not allowed as a choice for the ICMJE but is a choice in CT.gov
Many Patients Expect Data Sharing and Reuse

NEJM Aligning Incentives for Sharing Clinical Trial Data Summit, Boston, MA. April 2017
“Patients don’t have patience”

Why is data used only one time to answer one question (the primary endpoint) rather than leveraging participants’ contributions to answer multiple questions to understand disease and find treatments?
Barriers to Data Sharing (IPD) for Academics

• For most academic trialists (Data Contributors)
  - secure data hosting and sharing platforms not available or limited to within the institution
  - no standard data use agreements
  - no independent review process available to adjudicate data requests
  - cost and difficulty of de-identifying IPD and making it available
  - All this makes it difficult to meet data sharing requirements

• For Data Users
  - difficult to discover what IPD is available for sharing
  - combining datasets from different platforms is resource- and time-intensive
  - different data standards, data requirements, security standards, policies
The Role of Data Standards in Data Sharing and Reuse

• Increasing the number of datasets that use data standards will allow meta-analysts to more easily harmonize and integrate studies and perform analyses such as:
  - Conducting new subgroup analyses
  - Model prognostic or diagnostic data
  - Aggregate studies with time to event outcomes such as survival
# What Data is Shared Through Vivli?

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recommended Set</strong></td>
<td></td>
</tr>
<tr>
<td>Study protocol</td>
<td>Final protocol with all amendments</td>
</tr>
<tr>
<td>Data dictionary</td>
<td>Detailed descriptions of each variable in the dataset, including the definition, source, coding, etc. of the variable</td>
</tr>
<tr>
<td>Statistical Analysis Plan</td>
<td>Description of the principal features of the analyses described in the protocol</td>
</tr>
<tr>
<td>Clinical Study Report (CSR)</td>
<td>Report that summarizes the efficacy and safety data from the study (after regulatory decision)</td>
</tr>
<tr>
<td>IPD dataset</td>
<td>Final cleaned individual participant-level data, de-identified/anonymized</td>
</tr>
<tr>
<td><strong>Optional</strong></td>
<td></td>
</tr>
<tr>
<td>Analytic code</td>
<td>Software code used to carry out prespecified and additional analyses</td>
</tr>
<tr>
<td>Analysis ready IPD dataset</td>
<td>Dataset in a format used to carry out a sponsor’s analyses</td>
</tr>
<tr>
<td>Case report forms</td>
<td>Forms used to collect the data that is described in the protocol for each trial participant</td>
</tr>
</tbody>
</table>

NOTE: *this is a subset of the entire full data package and includes the data that underlies the publication findings (tables, figures)*
Datasets with variations of Biomarker X mapped to e.g., a common CDISC variable

Agenda

1. Why share data?
2. What is the role of data standards in data sharing and reuse?
3. The Vivli Global Platform for Clinical Data sharing and Reuse as a use case
Introducing Vivli

THE ENTITY

- Non-profit organization
- Convening function
  - Biomedical industry (pharma, bio, device)
  - Academia
  - Non-profit funders and foundations
  - Government (funders and regulators)
  - Patient/patient advocates
- Governance and policy
  - Harmonizing language & agreements
  - Move culture of data sharing
- Advocacy
  - Lowering barriers
  - Promoting incentives
- Oversight of implementation

THE PLATFORM

- A user-friendly, secure, state-of-the-art data sharing and computing platform
- Serving the international community, including trials from any disease, country, sponsor, funder, or investigator
  - Open search
  - Robust security
  - Modern tools and technologies
<table>
<thead>
<tr>
<th>Institutional Sharing</th>
<th>Individual Researcher/Team Sharing and Reuse</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Institutional membership</td>
<td>• Covers single publication or trial</td>
</tr>
<tr>
<td>• Ensures all researchers at an institution or division have access to a central sharing resource</td>
<td>• Recognizes life cycle of grant is not the same as life cycle of sharing</td>
</tr>
<tr>
<td>• DOI minted for credit and citation</td>
<td>• DOI minted for credit and citation</td>
</tr>
</tbody>
</table>
Vivli Diverse Membership
Summary: Benefits of Sharing through Vivli

- **Ease of sharing** - Sharing de-identified data is facilitated through either institutional memberships in Vivli or individually per dataset.
- **Citation** – DOIs allow for citation and credit of your research data.
- **Metrics** – Yearly metrics on number of data requests, resulting publications, etc.
- **Long-term archiving** – Archive your trials on Vivli (at least 25 years).
- **Post-grant data sharing** – Management of IPD sharing that continues even after grant funds end.
- **Funder and journal mandates** – Easily fulfill requirements for data sharing plans.
How to Access Data in Vivli?

4600+ Trials

2M Participants from 109 countries
Secure Environment Bridges Multiple Platforms

- STATA
- MS Office
- R
- Jupyter Notebook
- Python
- SAS
Vivli is a Global Data Platform – Agnostic to Disease, Funder or Data Contributor
Data Request and Access Process

**SEARCH**

Search Vivli platform for information about available studies.

**REQUEST**

Request IPD data package.

Each Data Request reviewed according to contributor’s publicly stated requirements.

**ACCESS**

Access data from approved requests in Vivli’s secure research environment or downloaded with permission.

**ANALYZE**

Use robust analytical tools to combine and analyze multiple data sets.

**DISSEMINATE**

Completed research results assigned a DOI. Researchers may use the Vivli platform to meet their publication requirements.
Log on to Vivli.org

- Explore the ~thousands of trials available via the Vivli platform
- Begin your search
- Contact support@vivli.org with questions