Preparing for clinical trial data sharing and re-use: the new reality for researchers.

Rebecca Li, PhD, Executive Director, Vivli

Frank W. Rockhold, PhD
Professor of Biostatistics and Bioinformatics
Duke University Medical Center

NIH Collaboratory Grand Rounds, September 27 2019

Duke Clinical Research Institute | FROM THOUGHT LEADERSHIP TO CLINICAL PRACTICE

Calls for Open Science

- Calls for greater transparency and ‘open data access’ in clinical research continue actively.
  - “Open science is the movement to make scientific research, data and dissemination accessible to all levels of an inquiring society”*
  - Open Science Project**: “If we want open science to flourish, we should raise our expectations to: Work. Finish. Publish. *Release.”*
- Specifically, open access to individual patient data from clinical trials is an critical tool for research in health care.

*https://www.fosteropenscience.eu/resources
**(http://openscience.org/):

Access to individual patient data (IPD) from clinical trials is important for future research

- There are certainly challenges, but question is not *whether* data should be shared, but rather *how and when* access should be granted.
- *Responsible* open access enables secondary analyses which:
  - Enhance reproducibility of clinical research
  - Honor the contributions of trial participants,
  - Improve the design of future trials
  - Generate new research findings
- This journey of making patient data available is part of an evolution in transparency and not a sudden awakening.

Open vs Transparent vs Access or Sharing

- What does it mean to be “open” or “transparent” and why is it important?
- Transparency and openness are strategy or belief systems
- Disclosure and access are actions which are necessary steps on that journey
- What is the difference between “access” and “sharing”?
- Disclosure or access without transparency, might check a regulatory box, but not help patients, healthcare practitioners or researchers.
- Transparency can only be achieved if people disclose in a manner digestible by the recipient

Enabling Open Science and IPD Access

- Some of the challenges are:
  - Patient privacy
  - Academic credit and commercial sensitivity
  - Data standards,
  - Resources (money and people)
- There should be room for researchers and patients alike to gain from this effort.
- Trialists, Patients, Statisticians and data scientists are essential elements in this effort.

2018: Numerous platforms in place!

- Clinical Study Data Request: multi-sponsor request site (13 companies), managed by the Wellcome Trust
- YODA: Yale Open data Access for two sponsors (Janssen/Medtronic)
- Project Data Sphere (CEO roundtable on cancer)
- INSPIIRE : Integrated System for Pfizer Investigator Initiated Research
- SOAR: Bristol Myers Squibb and Duke Data Strategic Initiative (DCRI)
- Celgene's Clinical Trial Data Sharing
- NIH BioLiNCC
- Vivli.org
- And many others in development
- So good news and in some ways but a fractured, disconnected approach

Spectrum of Data Sharing Models

Immune Tolerance Network- Trial share

- Open access to ITN data after registration and agreement to terms of use
- No further approval process
- Downloadable data

ClinicalStudyDataRequest.com

- Multiple industry sponsors; governance ranges by sponsor
- Secure interface, DUA, IRP
- IRP considers scientific relevance, COIs, and investigator expertise
- Some sponsors may review requests, and veto based on data specific considerations, competitive risk etc.

SOAR

- DUA, IRP
- Evaluates for COIs and research quality
- Requirement for detailed statistical analysis plan, evaluated for major design flaws
- Final analyses are reviewed by the IRC prior to publication

AHA Precision Medicine Initiative

- CV and stroke data
- Cloud-based, secure sharing environment
- Forum for collaboration
- Data access is granted in private workspaces by data contributor

Project data sphere (PDS)

- Oncology research
- Downloadable data
- DUAs
- Open to all
- **Control Group Only**

Vivli

- Attempting to harmonizing data sharing governance
- Secure interface, DUAs, IRP
- Review process considers the research plan, team, statistician, and COIs
- Contributors can veto requests, but number and reasons for rejections will be made public

YODA

- Generally, data is not downloadable
- DUA, IRP
- Data requests evaluated for scientific merit and COIs
- Restrictions to data access for legal or commercial purposes

Open Access

Restricted Access

Vivli

CENTER FOR GLOBAL CLINICAL RESEARCH DATA

Preparing for Clinical Trial Data Sharing and Reuse – The New Reality for Researchers and Institutions

Rebecca Li, Vivli Executive Director
Faculty Co-Director of Research Ethics, Harvard Center for Bioethics
Harvard Medical School

September 24, 2019

Agenda

1. Why should we share?
2. What are the key components of a data sharing program?
3. How should we think about sharing if we are:
   - an Institution
   - an Individual researcher or team
4. The Vivli Global Platform for Clinical Data sharing and Reuse

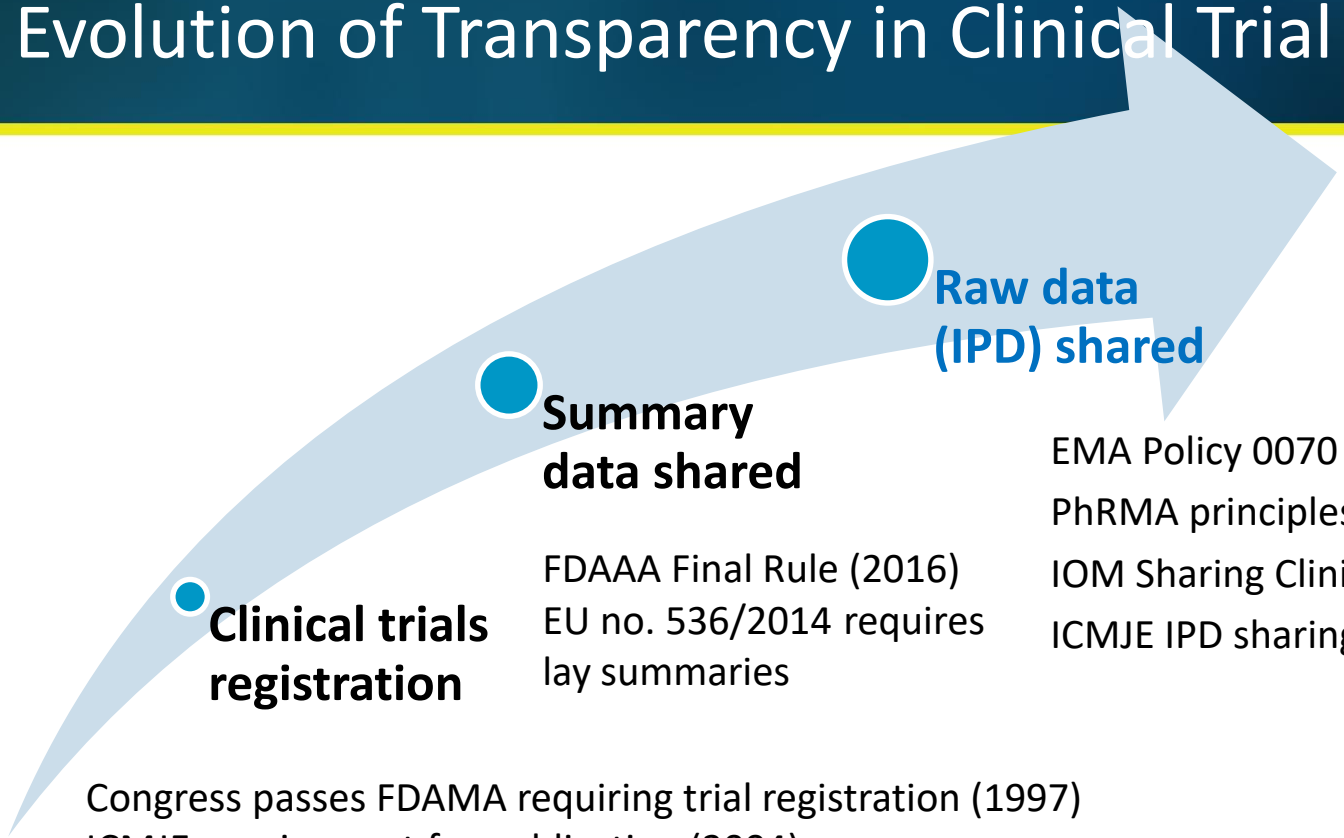
Vivli

1. Why Should We Share Our Clinical Research Data

- Funder requirements
- Journal requirements (new this year)
- 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

Vivli

Evolution of Transparency in Clinical Trial Data



Clinical trials registration

Congress passes FDAMA requiring trial registration (1997)
ICMJE requirement for publication (2004)

Summary data shared

FDAAA Final Rule (2016)
EU no. 536/2014 requires lay summaries

Raw data (IPD) shared

EMA Policy 0070 (2014), Policy 0043 (TBD)
PhRMA principles for data sharing (2014)
IOM Sharing Clinical Trial Data report (2015)
ICMJE IPD sharing statement (July 2018)

Vivli

What are Journals Requiring as of July 1, 2018?

The NEW ENGLAND JOURNAL of MEDICINE

EDITORIALS

Data Sharing Statements for Clinical Trials — A Requirement of the International Committee of Medical Journal Editors

The International Committee of Medical Journal Editors (ICMJE) believes there is an ethical obligation to responsibly share data generated by

explained at www.icmje.org/recommendations/browse/publishing-and-editorial-issues/clinical-trial-registration.html. If the data sharing plan

- Major journals including NEJM, JAMA, The Lancet, BMJ, Annals of Internal Med, PLoS Medicine, and hundreds of others (ICMJE)
- 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"

Taichman DB, et al. N Engl J Med 2017; 376:2277-2279

Vivli

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

Vivli

At Trial Registration (from Clinicaltrials.gov)

IPD Sharing Plan Description
Definition: If Plan to Share IPD is "Yes," briefly describe what specific individual participant data sets are to be shared (for example, all collected IPD, all IPD that underlie results in a publication). If the Plan to Share IPD is "No" or "Undecided," an explanation may be provided for why IPD will not be shared or why it is not yet decided.
Limit: 1000 characters.

If Plan to Share IPD is "Yes," provide the following information.

IPD Sharing Supporting Information Type
Definition: The type(s) of supporting information that will be shared, in addition to the individual participant data set and data dictionaries for the IPD itself. Select all that apply.

- Study Protocol
- Statistical Analysis Plan (SAP)
- Informed Consent Form (ICF)
- Clinical Study Report (CSR)
- Analytic Code

Vivli

For Grant Submission

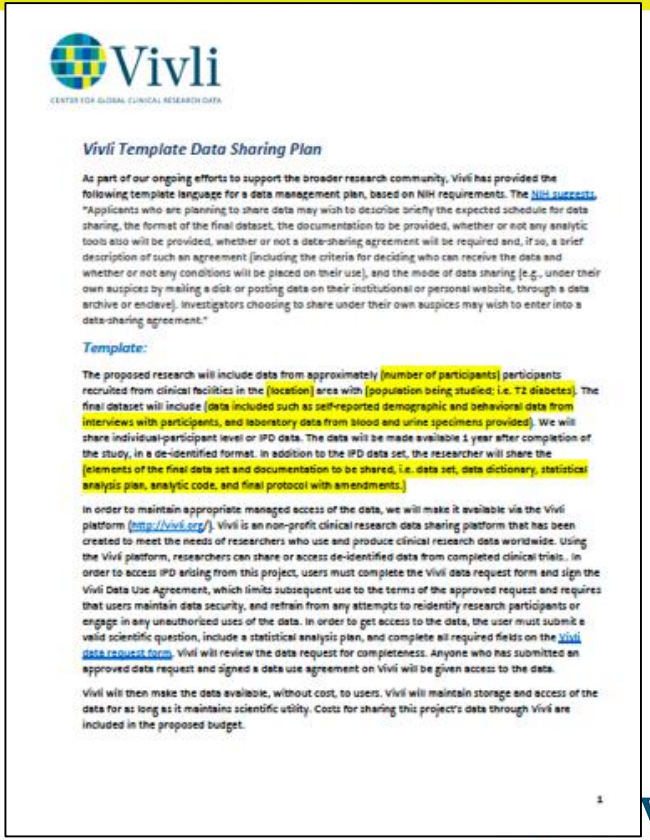
Funders increasingly requiring data sharing

Draft NIH Data Sharing and Management Policy is requiring

- IPD sharing plan for all grants
- sharing and managing of data according to approved plan

Data sharing costs should be part of the budget proposal

Vivli provides an NIH-compliant data sharing plan template



Vivli

CENTER FOR GLOBAL CLINICAL RESEARCH DATA

Vivli Template Data Sharing Plan

As part of our ongoing efforts to support the broader research community, Vivli has provided the following template language for a data management plan, based on NIH requirements. The NIH suggests, "Applicants who are planning to share data may wish to describe briefly the expected schedule for data sharing, the format of the final dataset, the documentation to be provided, whether or not any analytic tools also will be provided, whether or not a data-sharing agreement will be required and, if so, a brief description of such an agreement (including the criteria for deciding who can receive the data and whether or not any conditions will be placed on their use), and the mode of data sharing (e.g., under their own auspices by mailing a disk or posting data on their institutional or personal website, through a data archive or enclave). Investigators choosing to share under their own auspices may wish to enter into a data-sharing agreement."

Template:

The proposed research will include data from approximately [number of participants] participants recruited from clinical facilities in the [location] area with [population being studied; i.e. T2 diabetes]. The final dataset will include [data included such as self-reported demographic and behavioral data from interviews with participants, and laboratory data from blood and urine specimens provided]. We will share individual-participant level or IPD data. The data will be made available 1 year after completion of the study, in a de-identified format. In addition to the IPD data set, the researcher will share the [elements of the final data set and documentation to be shared, i.e. data set, data dictionary, statistical analysis plan, analytic code, and final protocol with amendments.]

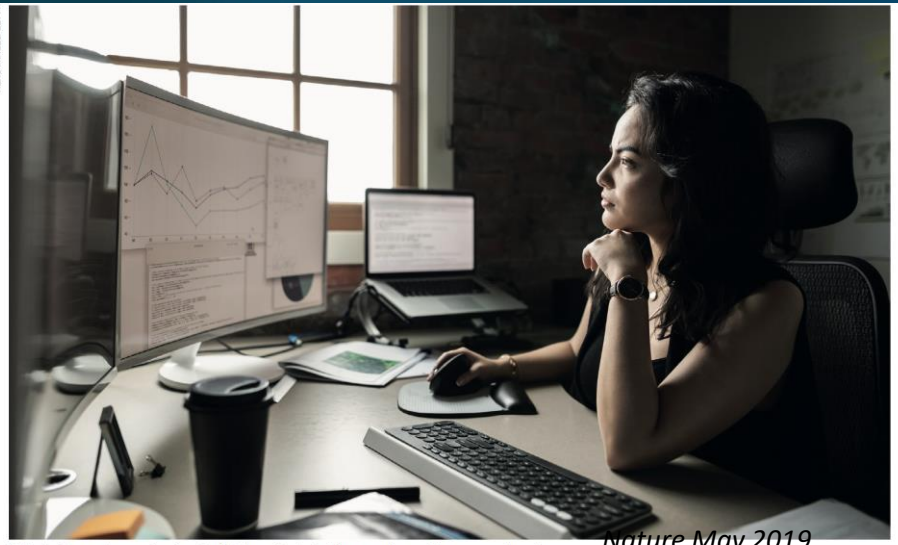
In order to maintain appropriate managed access of the data, we will make it available via the Vivli platform (http://vivli.org/). Vivli is an non-profit clinical research data sharing platform that has been created to meet the needs of researchers who use and produce clinical research data worldwide. Using the Vivli platform, researchers can share or access de-identified data from completed clinical trials. In order to access IPD arising from this project, users must complete the Vivli data request form and sign the Vivli Data Use Agreement, which limits subsequent use to the terms of the approved request and requires that users maintain data security, and refrain from any attempts to reidentify research participants or engage in any unauthorized uses of the data. In order to get access to the data, the user must submit a valid scientific question, include a statistical analysis plan, and complete all required fields on the Vivli data request form. Vivli will review the data request for completeness. Anyone who has submitted an approved data request and signed a data use agreement on Vivli will be given access to the data.

Vivli will then make the data available, without cost, to users. Vivli will maintain storage and access of the data for as long as it maintains scientific utility. Costs for sharing this project's data through Vivli are included in the proposed budget.

1

vli

HERO IMAGES/GETTY



Data sharing can be complex for scientists to navigate, but the rewards are often career-enhancing.

Nature May 2019

OPEN SCIENCE

Setting your data free

As science becomes more open, researchers who share data are reaping the benefits.

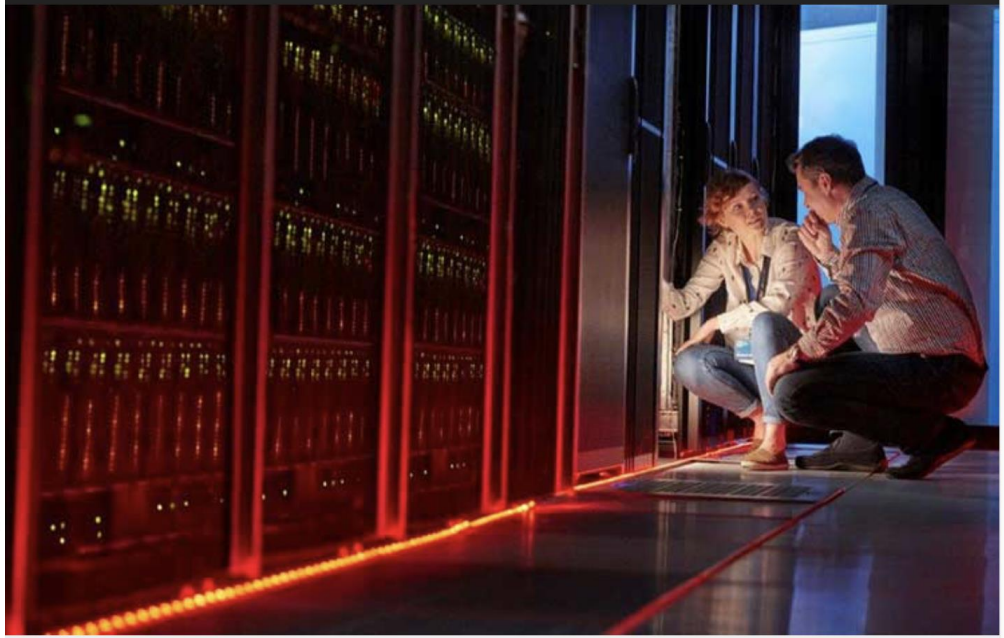
BY GABRIEL POPKIN

Ecologist Thomas Crowther knew that scientists had already collected a vast amount of field data on forests worldwide. But almost all of those data were sequestered in researchers' notebooks or per-

in CSV files (plain-text files that contain a list of data) on servers at Crowther's present laboratory at the Swiss Federal Institute of Technology in Zurich and on those of a collaborator at Purdue University in West Lafayette, Indiana; he hopes to outsource database storage to a third-party organization with expertise

current state of science: partly open, partly closed, and with unclear and inconsistent policies and expectations on data sharing that are still in flux. High-level bodies such as the US National Academies of Sciences, Engineering, and Medicine and the European Commission have called for science

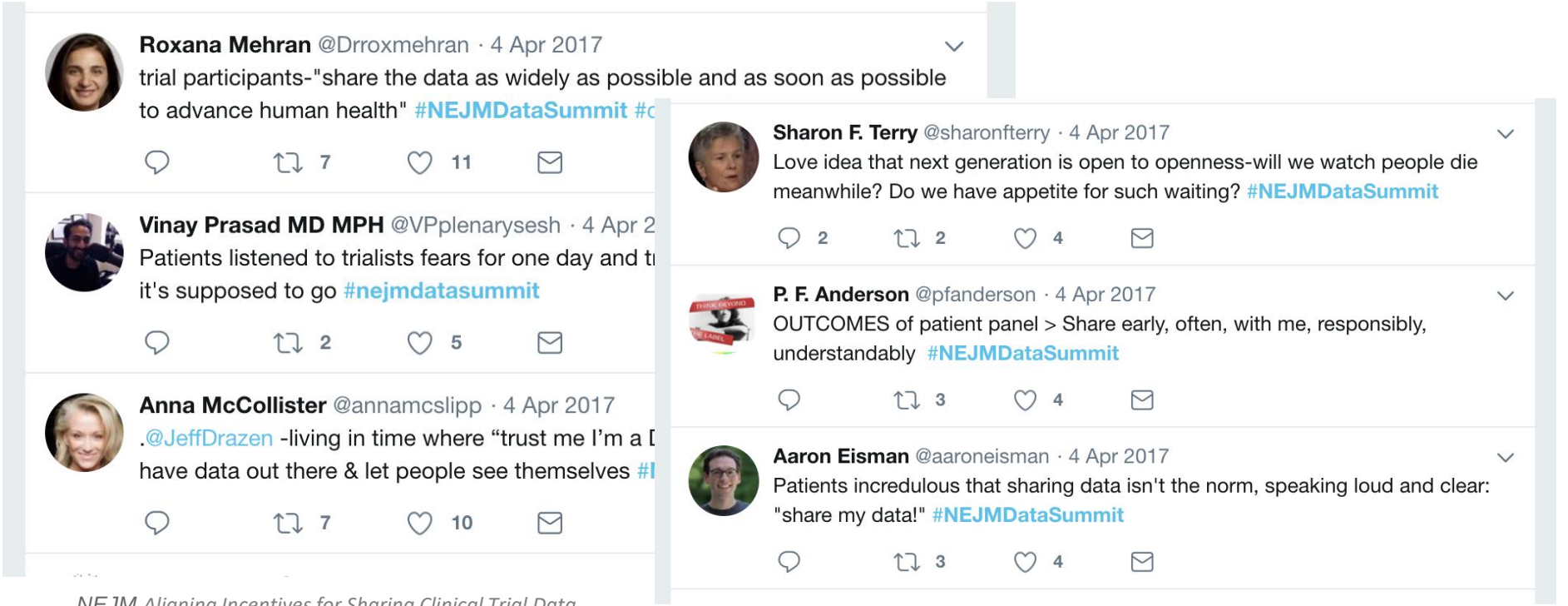
nature INDEX



Caiaimage/Rafal Rodzoch

"A love letter to your future self": What scientists need to know about FAIR data

Many Patients Expect Data Sharing and Reuse



Roxana Mehran @Drroxmehran · 4 Apr 2017
trial participants-"share the data as widely as possible and as soon as possible to advance human health" #NEJMDataSummit #c
7 11

Vinay Prasad MD MPH @VPplenarysesh · 4 Apr 2
Patients listened to trialists fears for one day and t
it's supposed to go #nejmdatasummit
2 5

Anna McCollister @annamcslipp · 4 Apr 2017
.@JeffDrazen -living in time where "trust me I'm a D
have data out there & let people see themselves #I
7 10

Sharon F. Terry @sharonfterry · 4 Apr 2017
Love idea that next generation is open to openness-will we watch people die meanwhile? Do we have appetite for such waiting? #NEJMDataSummit
2 2 4

P. F. Anderson @pfanderson · 4 Apr 2017
OUTCOMES of patient panel > Share early, often, with me, responsibly, understandably #NEJMDataSummit
3 4

Aaron Eisman @aaroneisman · 4 Apr 2017
Patients incredulous that sharing data isn't the norm, speaking loud and clear: "share my data!" #NEJMDataSummit
3 4

NEJM Aligning Incentives for Sharing Clinical Trial Data Summit, Boston, MA. April 2017

Vivli

Perhaps most importantly for participants if the data is not shared…

It is used only one time to answer one question (the primary endpoint) rather than leveraging participants' contributions to answer multiple scientific lines of inquiry thereby advancing science

Vivli

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
  - disease-specific data sharing platforms limit cross-disciplinary data discovery
  - limited range of analytic tools available

Agenda

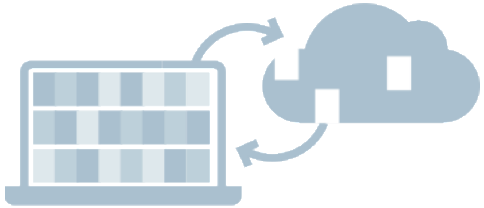
1. Why should we share?
2. **What are the key components of a data sharing program?**
3. How should we think about sharing if we are:
   - an Institution
   - an Individual researcher or team
4. The Vivli Global Platform for Clinical Data sharing and Reuse

Vivli

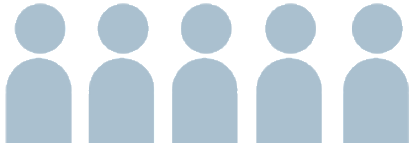
2. How to Share: 3 key elements to consider

POLICY

MECHANISM



RESOURCES



Vivli

Mechanisms for Sharing IPD Data Externally - Trial Data sharing platforms

Type	Key Requirements, features
Open access	No requirements /account creation, simple on-line DUA, data downloadable
Managed access	Intermediary, proposal process, specialized expertise, DUA, data available in the cloud /downloadable
Restricted access	Invitation-only, access to those that provide data

Vivli

Data sharing governance

Key considerations when formulating your policy

- In data sharing, transparent decision-making equals good public policy
- Data Sharing Policies vary based on an institution's current portfolio, experience with data sharing and risk tolerance

Vivli

Data sharing governance

Key considerations when formulating your institution's policy

1. Which studies will you share externally?
2. Are there exceptions to sharing?
3. Will there be a centralized review panel that will review requests or will this function be delegated?

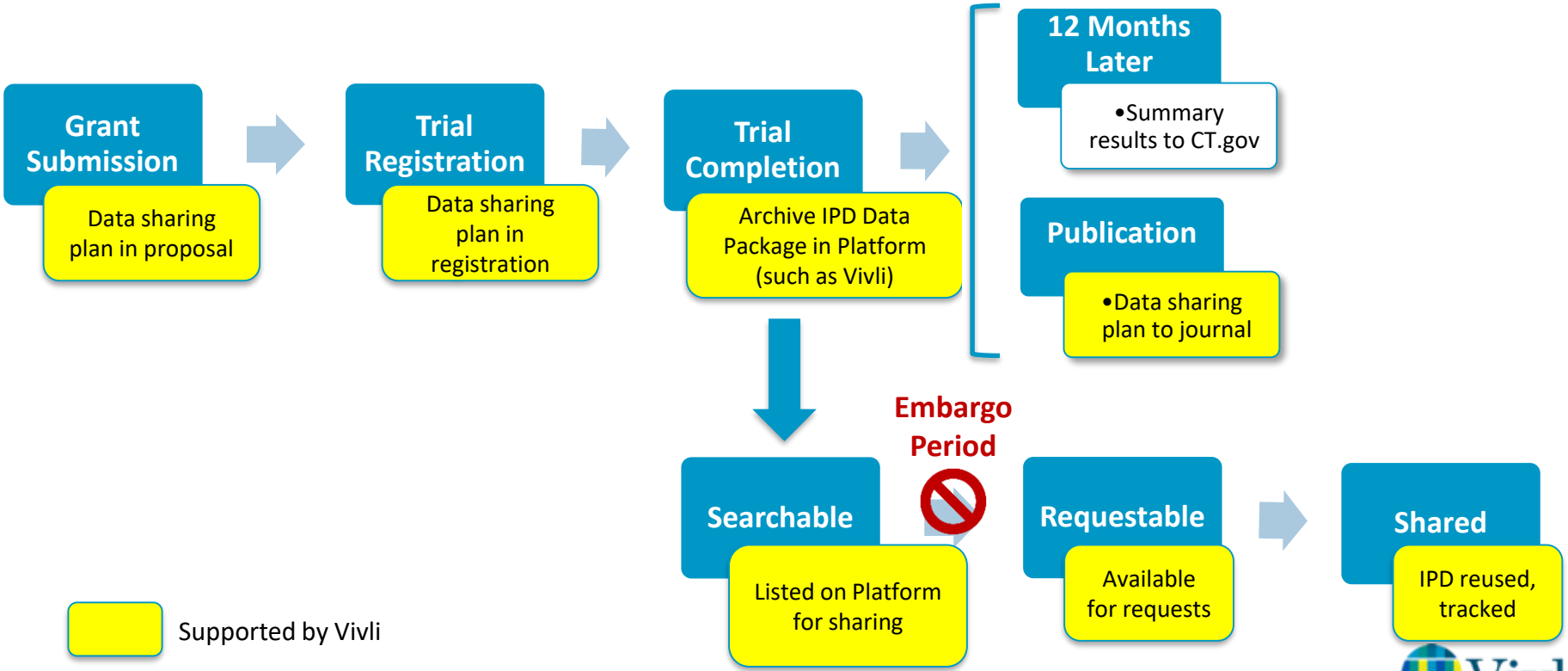
Vivli

Agenda

1. Why should we share?
2. What are the key components of a data sharing program?
3. **How should we think about sharing if we are:**
   - an Institution
   - an Individual researcher or team
4. The Vivli Global Platform for Clinical Data sharing and Reuse

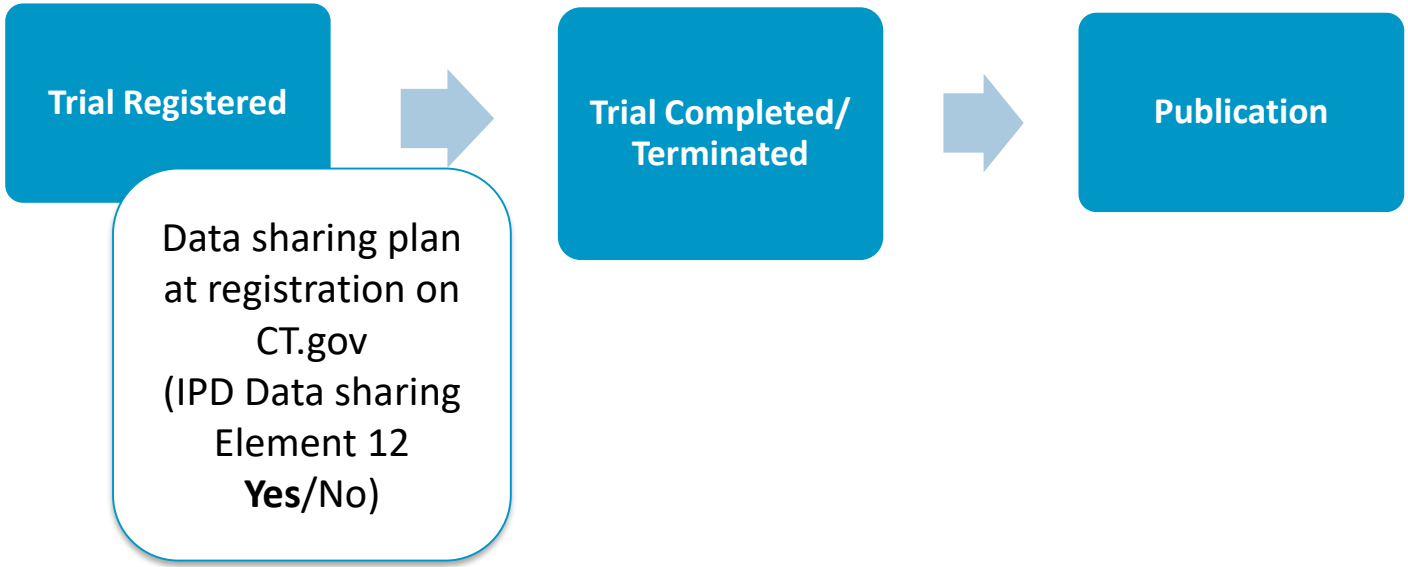
Vivli

Data Sharing Overall Timeline



Vivli

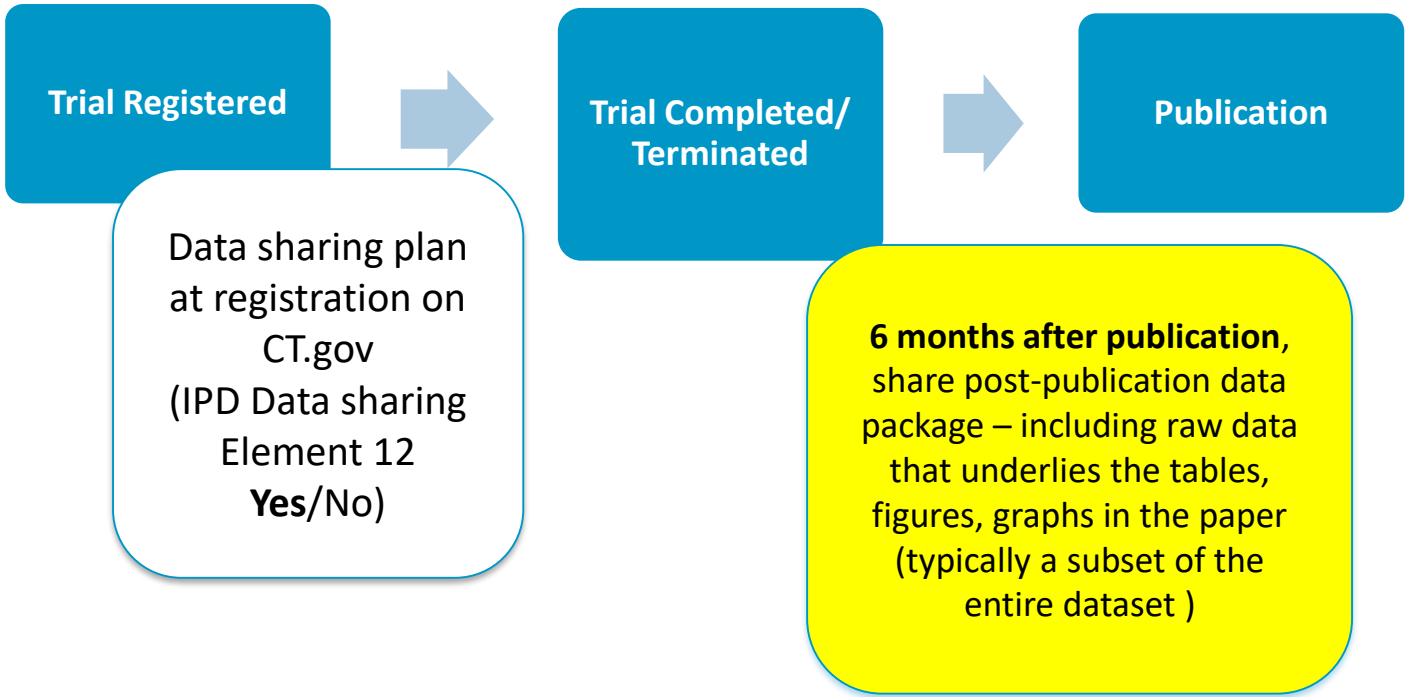
3. Recommendation - When to share what



Recommendations based upon Institute of Medicine report *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk*, Jan. 2015

Vivli

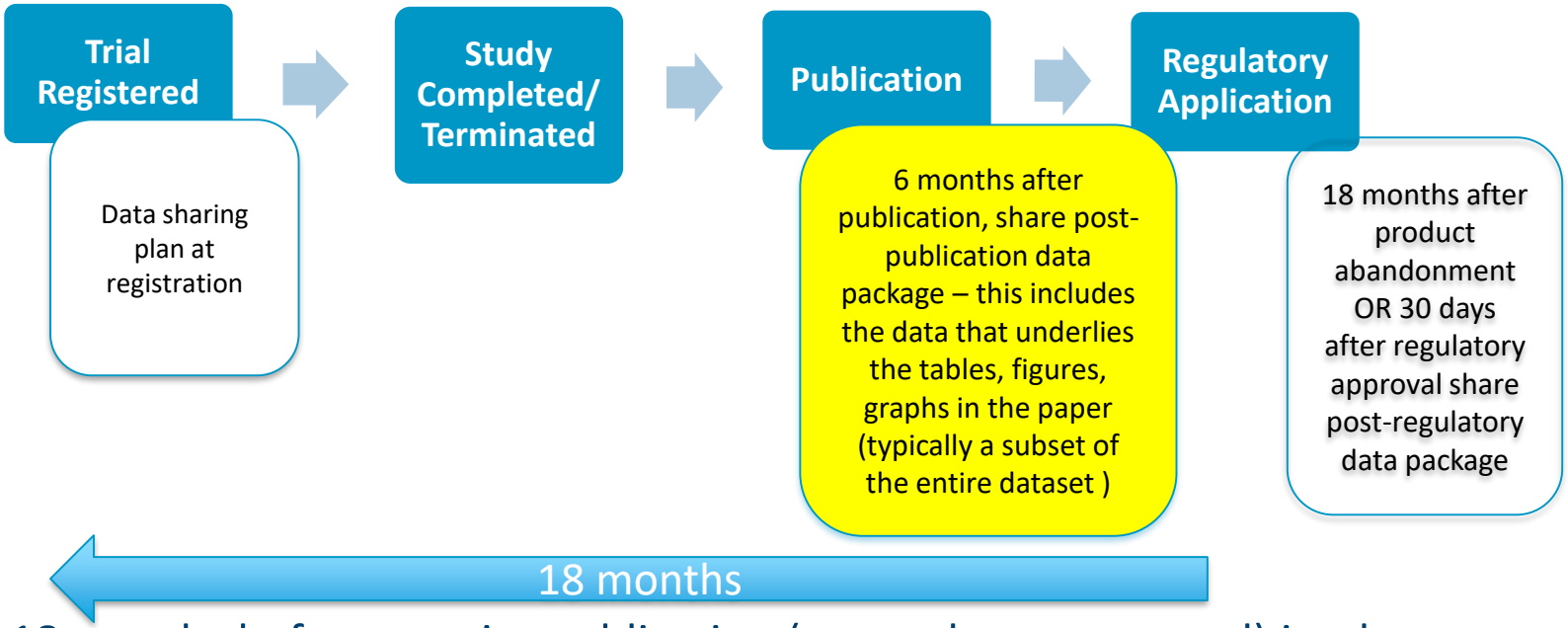
3. Recommendation - When to share what



Recommendations based upon Institute of Medicine report *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk*, Jan. 2015

Vivli

3. Recommendation - When to share what



At least 18 months before a major publication (or regulatory approval) is when teams or institutions should begin their data sharing program planning

Recommendations based upon IOM report *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk*, Jan. 2015

Vivli

What data will be shared?

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

NOTE: *this is a subset of the entire full data package and includes the data that underlies the publication findings (tables, figures)

VIVLI

How can we manage a data sharing program?

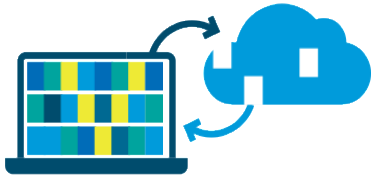
- **Manage in-house:**
  - *Mechanism for sharing* – build, management and updating of a platform
  - *Team* – internal resources to maintain the platform; negotiate legal agreements; user queries, generate metrics, data anonymization and data preparation
  - *Policy* – draft and manage data sharing policies

- **Or Consider a partnerships to manage and assist with:**
  - Mechanism
  - Team
  - Policy

Agenda

1. Why should we share?
2. What are the key components of a data sharing program?
3. How should we think about sharing if we are:
   - an Institution
   - an Individual researcher or team
4. **The Vivli Global Platform for Clinical Data sharing and Reuse**

Vivli

YOUR NEED		VIVLI SOLUTION
POLICY		We provide expertise in policy development and harmonized agreements
MECHANISM		Use the platform to securely share your data
RESOURCES		Our team manages researchers' queries

Vivli

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

Vivli Solution Models for Data Sharing

Institutional Sharing

- Institutional membership
- Ensures all researchers at an institution or division have access to a central sharing resource
- DOI minted for credit and citation

Individual Researcher/ Team Sharing and Reuse

- Covers single publication or trial
- Recognizes life cycle of grant is not the same as life cycle of sharing
- DOI minted for credit and citation

Vivli

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

Vivli

Vivli

CENTER FOR GLOBAL CLINICAL RESEARCH DATA

How to Access Data in Vivli?

Secure Environment Bridges Multiple Platforms

BioLINCC
Biologic Specimen and Data Repository Information Coordinating Center

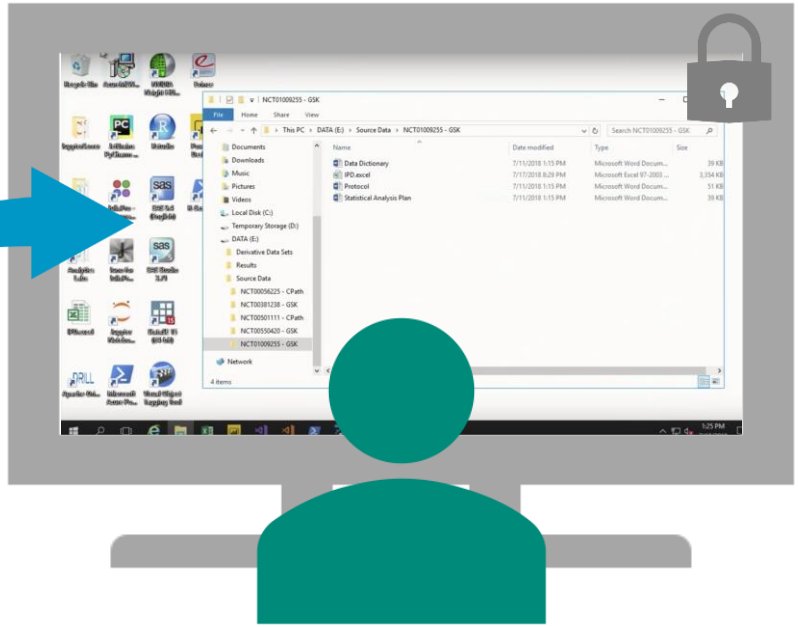
Johnson & Johnson

ImmPort
Bioinformatics For The Future Of Immunology

Project Data Sphere®

gsk
do more
feel better
live longer

Vivli Secure Environment



- STATA
- MS Office
- R
- Jupyter Notebook
- Python
- SAS

Vivli

Vivli is a Global Data Platform – Agnostic to Disease, Funder or Data Contributor

Vivli

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.

Vivli

Log on to

Vivli.org

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

Vivli