



CENTER FOR GLOBAL CLINICAL RESEARCH DATA

Responsible data-sharing to improve research integrity

November 2020

Julie Wood, Vivli

Vivli: The Difference Data Re-Use Can Make



Balancing Risks and Benefits

Openness



Maximizing the value of the data collected respects participants' contributions

Privacy



Protecting participant privacy, particularly when sharing individual patient-level data (IPD)

Approaches to Sharing Human Data

Type	Key Features	Examples
Open access	Anyone can access, simple account creation, simple on-line Data Use Agreement (DUA)	Health CAN, EMA, Project Datasphere
Managed access	<ul style="list-style-type: none">• for scientific purposes only (standard request form)• (independent) review process• secure environment for data access• clear legal framework	Vivli, CSDR, SOAR, VISTA
Restricted access	Invitation only, access only to those who provide data	DataCelerate, IBD Plexus

Challenges with Data Sharing

The Atlantic

Popular

Latest

Sections ▾

HEALTH

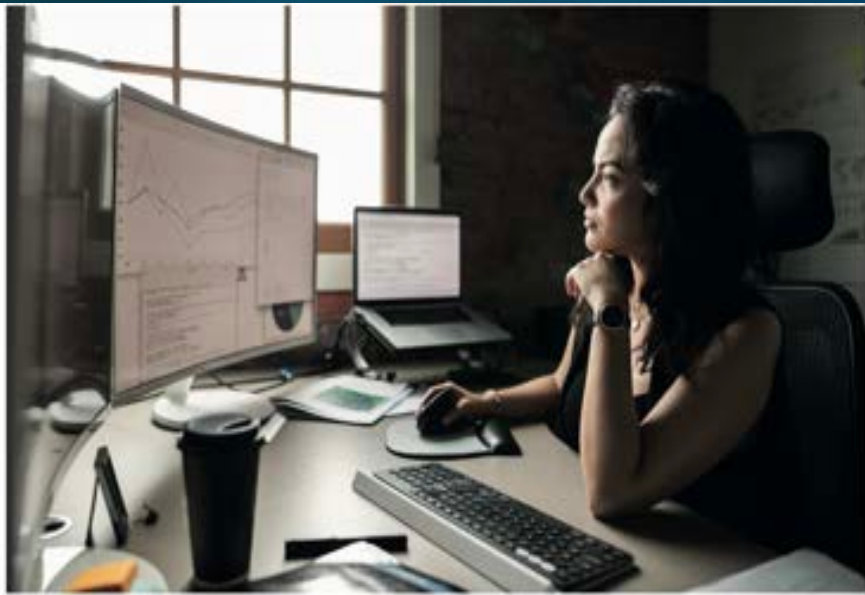
Scientists Have a Sharing Problem

Competition and disorganization within their disciplines prevent many researchers from making their data publicly available, which is stunting scientific progress.

MAGGIE PUNIEWSKA DEC 15, 2014

When it comes to sharing information, there seems to be quite a difference of opinion—across areas both trivial and serious—as to how much is enough.

Some people broadcast their lives on Facebook; others poke fun at the



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

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



Shutterstock/Rafal Rodzich

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

Challenges with Data Sharing

- “The secondary requester will be unfamiliar with the data set structure or analysis unduly alarm the public or hinder science rather than advance science”
- “What if they find something a mistake in my trial dataset?”
- Providing academic credit
- Clinical trials take years to design, conduct and analyze is this fair?
- “Won’t this stifle innovation and new science?”



Developing a data sharing program

1. Why should we share?
2. What are the key components of a data sharing program?
3. When should we begin a program?
4. How can we manage a data sharing program?
5. What is the role of a partner such as Vivli?

1. Why Should My Organization Share Its Data?

- Ethical obligations to trial participants
- Journal requirements
- BIO, EFPIA, PhRMA, IFPMA publicly stated commitments for members



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" [#NEJMDDataSummit](#) [#c](#)



7



11



Vinay Prasad MD MPH @VPplenarysesh · 4 Apr 2017

Patients listened to trialists fears for one day and to it's supposed to go [#nejmdatasummit](#)



2



5



Anna McCollister @annamcslipp · 4 Apr 2017

.@JeffDrazen -living in time where "trust me I'm a doctor" have data out there & let people see themselves [#I](#)



7



10



Sharon F. Terry @sharonferry · 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? [#NEJMDDataSummit](#)



2



2



4



P. F. Anderson @pfanderson · 4 Apr 2017

OUTCOMES of patient panel > Share early, often, with me, responsibly, understandably [#NEJMDDataSummit](#)



3



4



Aaron Eisman @aaroneisman · 4 Apr 2017

Patients incredulous that sharing data isn't the norm, speaking loud and clear: "share my data!" [#NEJMDDataSummit](#)



3



4



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

What Do Surveys Show Regarding Patient and Participant Preference Regarding Data Sharing?

- High levels of support for data sharing; however, patients are reluctant to have their data “commodified” purely for commercial gain ¹
- If adequate safeguards were in place, trial participants are willing to share their data²

¹ Davidson S, McLean C, Treanor S, Aitken M, Cunningham-Burley S, Laurie G, et al. Public acceptability of data sharing between the public, private and third sectors for research purposes. Edinburgh: Scottish Government Social Research; 2013. [Google Scholar](#)

² Mello, Michelle M., Van Lieu, and Steven N. Goodman. "Clinical trial participants' views of the risks and benefits of data sharing." *New England Journal of Medicine* 378.23 (2018): 2202-2211.

Not only is it the right thing to do, it also helps increase citation rate

The screenshot shows the PLOS ONE website interface. At the top, there's a navigation bar with 'PUBLISH', 'ABOUT', and 'BROWSE' links, a search bar, and a 'SEARCH' button. Below the navigation bar, the article title 'Sharing Detailed Research Data Is Associated with Increased Citation Rate' is prominently displayed. The authors listed are Heather A. Pharo, Roger S. Day, and Douglas B. Fridman. The publication date is March 21, 2007, and the DOI is https://doi.org/10.1371/journal.pone.0000308. On the right side of the article header, there's a statistics box showing 733 Saves, 413 Citations, 72,867 Views, and 144 Shares. Below the article title, there's a tabbed interface with 'Article' selected. The article content includes an abstract, background, and principal findings. The principal findings section states: 'We examined the citation history of 85 cancer microarray clinical trial publications with respect to the availability of their data. The 48% of trials with publicly available microarray data received 85% of the aggregate citations. Publicly available data was significantly ($p=0.006$) associated with a 69% increase in citations, independently of journal impact factor, date of publication, and author country of origin using linear regression.'

PLOS ONE

PUBLISH ABOUT BROWSE

SEARCH advanced search

OPEN ACCESS PEER REVIEWED

RESEARCH ARTICLE

Sharing Detailed Research Data Is Associated with Increased Citation Rate

Heather A. Pharo, Roger S. Day, Douglas B. Fridman

Published: March 21, 2007 • <https://doi.org/10.1371/journal.pone.0000308>

Article Authors Metrics Comments Media Coverage

Abstract

Background

Sharing research data provides benefit to the general scientific community, but the benefit is less obvious for the investigator who makes his or her data available.

Principal Findings

We examined the citation history of 85 cancer microarray clinical trial publications with respect to the availability of their data. The 48% of trials with publicly available microarray data received 85% of the aggregate citations. Publicly available data was significantly ($p=0.006$) associated with a 69% increase in citations, independently of journal impact factor, date of publication, and author country of origin using linear regression.

Download PDF

Print Share

Check for updates

Included in the Following Collections

Open Access
Open Data
PLOS ONE 10 Year Anniversary Datasets

Reader Comments (8)

“The 48% of trials with publicly available microarray data received 85% of the aggregate citations. Publicly available data was significantly ($p = 0.006$) associated with a 69% increase in citations”

Trial Registration

- Data sharing plan is part of the ClinicalTrials.gov registration record

▼ 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.

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

- As of 1 January 2019, ICMJE requires registration of your data sharing plan at time of trial registration

What are Journals Requiring as of July 1, 2018?



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

- Major journals including NEJM, JAMA, The Lancet, BMJ, Annals of Internal Med, PLoS Medicine, and hundreds of others
- 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”

Providing Access to IPD Generates Value

- **Honors** the commitments of participants
- **Strengthens trust** in the clinical research enterprise
- **Prevents repetitive trials** and putting additional patients at risk
- **Enables new discovery** and scientific insights through by combining data from disparate sources

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

Developing a data sharing program

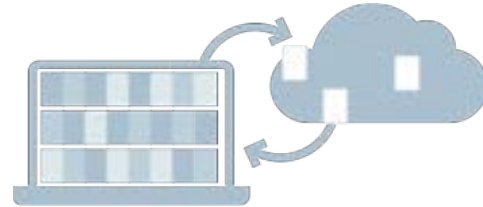
1. Why should we share?
2. **What are the key components of a data sharing program?**
3. When should we begin a program?
4. How can we manage a data sharing program?
5. What is the role of a partner?

3. How to Share: 3 key elements to consider

POLICY



MECHANISM



RESOURCES



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 a sponsor or funders current portfolio, experience with data sharing and risk tolerance

Data sharing governance

Key considerations when formulating your policy

1. Which studies will you share?
2. Are there exceptions to sharing?
3. On a specific request, who makes the final decision on whether to share?

Data sharing
governance

Key considerations
when formulating
your policy

1. What studies will you share?

- **From when? (Date)**
- **Which Phases?**
- **Both Submitted and Approved products?**

Data sharing
governance

Key considerations
when formulating
your policy

1. What studies will you share?

- **From when?**
- **Which Phases?**
- **Both Submitted and Approved products?**

Example 1: we will share interventional clinical trials conducted in patients (Phase I-IV) for products and indications submitted and approved, since 2002 will be shared.

Example 2: Company-sponsored studies supporting indications approved in both the United States and European Union after January 1, 2013 will be considered for sharing.

Data sharing
governance

Key considerations
when formulating
your policy

2. Exceptions to sharing

- **Practical constraints**
- **Legal or contractual constraints**
- **Language**
- **Anonymization**

Data sharing governance

Key considerations when formulating your policy

2. Exceptions to sharing

- **Practical constraints**
- **Legal or contractual constraints**
- **Language**
- **Anonymization**

Examples:

- Studies where there is reasonable likelihood that patients' anonymity cannot be maintained. For example in very rare diseases, studies with very low patient numbers or studies performed at a single center.
- There are practical constraints to providing the data (for example, issues related to the format of the databases, and/or resources (costs) are considerable to retrieve data)
- Study documentation is not in English.

Data sharing governance

Key considerations
when formulating
your policy

1. What studies will you share?
2. Are there exceptions to sharing?
3. **On a specific request, who makes the final decision on whether to share?**

Data sharing governance

Key considerations when formulating your policy

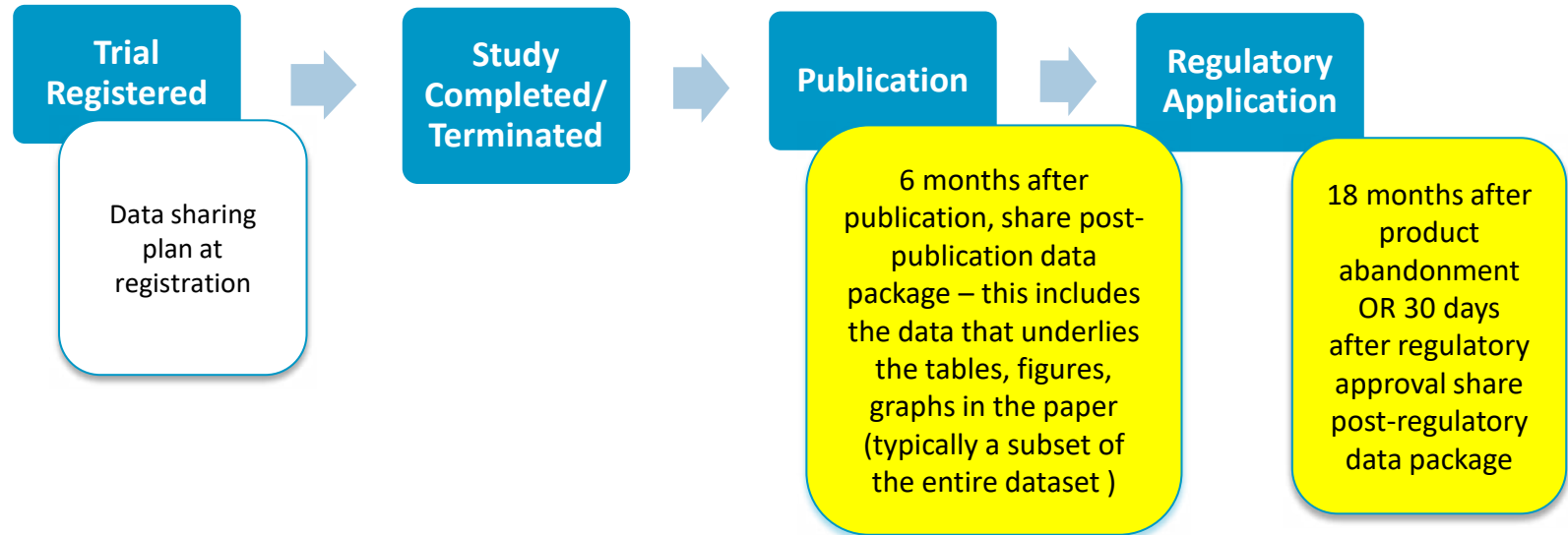
3. On a specific request, who makes the final decision on whether to share?

- Internal Approving Entity
- External Review Panel (ERP)
- Independent Review Panel (IRP)

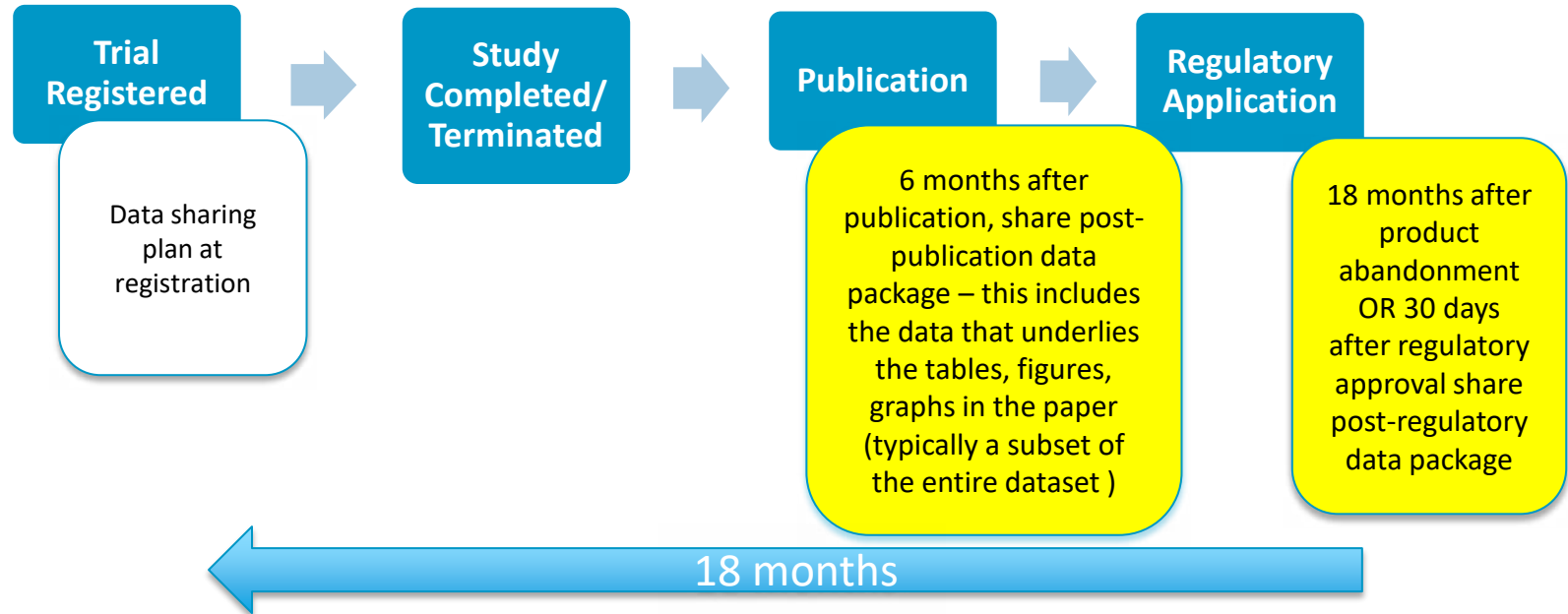
Agenda

1. Why should we share?
2. What are the key components of a data sharing program?
- 3. When should we begin a program?**
4. How can we manage a data sharing program?
5. What is the role of a partner?

3. Recommendation - When to share what



3. Recommendation - When to share what



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

What data will be shared?

Item	Description
<i>Recommended Set</i>	
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
<i>Optional</i>	
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

4. 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

- **Trusted partner to manage and assist with:**

- Mechanism
- Team
- Policy

5. What can partners like Vivli do for us?



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

CONFIDENTIAL - Not for distribution

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 by the numbers ...TODAY



Vivli Members

abbvie

AstraZeneca

Boehringer
Ingelheim

Biogen

BioLINCC
Biology Specimen and Data Repository Information Coordinating Center

Celgene

CRITICAL PATH
INSTITUTE

Cure
Duchenne

Daiichi-Sankyo

DD
DONOR DUKES
CHARITABLE FOUNDATION

Duke UNIVERSITY

gsk
do more
feel better
live longer

HARVARD
UNIVERSITY



THE LEONA M. AND HARRY B.
HELMSLEY
CHARITABLE TRUST

IMMPORT
IMMUNOMODULATORY FOR TUMOR FIGHTING FOR IMMUNOTHERAPY

JOHNS HOPKINS
UNIVERSITY

Johnson & Johnson

FAMILY OF COMPANIES

Lilly
Lundbeck



Mitsubishi Tanabe Pharma

Pfizer

Project Data
Sphere

REGENERON

Roche

TAIHO

TEMPUS

Takeda



Inspired by patients.
Driven by science.

UCSF

University of California
San Francisco

Vivli

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

Irritable Bowel Syndrome
Bacterial Peritonitis Glaucoma Endometriosis
Kidney cancer Non Hodgkins Lymphoma Epilepsy HIV
Breast cancer Cystic Fibrosis Diabetes Mellitus Insomnia
Coronary Artery Bypass Surgery Schizophrenia Bariatric Obesity
Atrial Fibrillation Fibromyalgia Cancer Traumatic Brain injury Trauma
Influenza Crohn's Diabetes Huntington's Disease Dabigatran
Atorvastatin Disease Hypertension Hepatitis CHepatitis Autism
Hidradenitis Depression Myocardial Arthritis
Psoriasis Statin Endometriosis Interleukin-6 Zoloft
Tysabri Tuberculosis Heart-Failure
Bipolar disorder Cannabinoids Asthma Lung cancer Lymphoma
Multiple Sclerosis Sickle Cell disease Atopic Dermatitis
Tumor burden Vitamin D Total Joint Replacement Cancer
Vedolizumab Pulmonary Arterial Hypertension Infarction
Hemophilia Sleep Apnea Edoxaban Type 1 Diabetes Mellitus
HPV Humira Colorectal Cancer Osteoarthritis
Lymphoma Stroke Ulcerative Colitis Vitiligo

How Vivli works

SEARCH

Search Vivli platform for information about available studies.



REQUEST

Request IPD Data sets.
Each Data Request will be **reviewed** according to contributors' publicly stated requirements.



ACCESS

Data from approved requests can be **accessed** in Vivli's secure research environment or **downloaded** with permission.



ANALYZE

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



DISSEMINATE

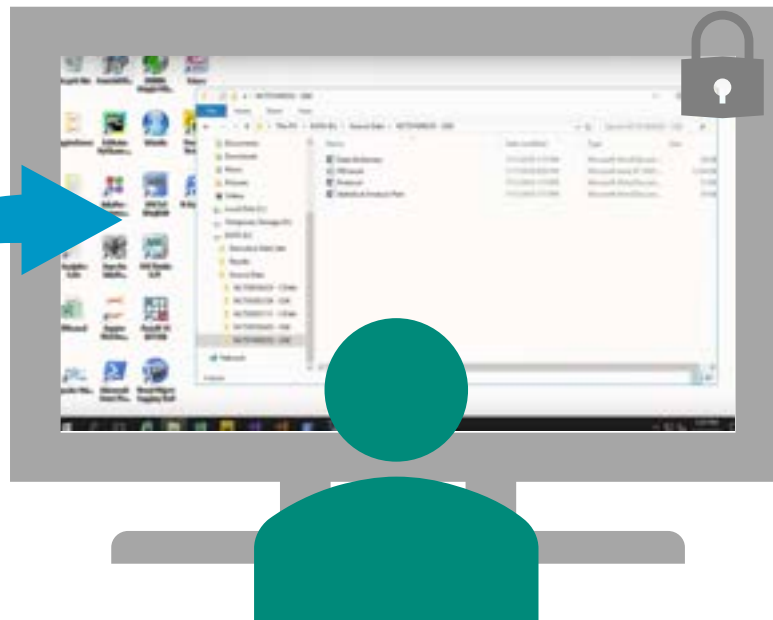
Completed **research results** will be assigned a DOI.
Researchers may use the Vivli platform to meet their **publication** requirements.



Vivli's Secure Environment Bridges Multiple Platforms



Vivli Secure Environment



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

Explore Vivli

Log on to

Vivli.org

To explore the ~5,800 trials available

Thank you!