

Basic Concepts of Data Sharing

University College London & Yale University
Facilitating Data Access to Non-Industry Funded Research
Strategy Development Meeting
London
9 October 2015

Joseph S. Ross, MD, MHS

Section of General Internal Medicine, School of Medicine
Center for Outcomes Research and Evaluation, Yale-New Haven Hospital



CENTER FOR OUTCOMES
RESEARCH AND EVALUATION

Yale SCHOOL OF MEDICINE

Potential Conflicts of Interest

- **Research grant funding through Yale from:**
 - **Johnson & Johnson, Inc. and Medtronic, Inc.:** develop and promote clinical trial data sharing
 - **FDA:** medical device post-market surveillance
 - **Pew Charitable Trusts:** Drug regulation and safety
 - **CMS:** performance measure development

Why Share Data?

- **Positions research as a public good**
- **Respects contributions of participants, maximizing value of collected data while minimizing duplicative data collection**
- **Facilitates follow-up studies of secondary research questions using existing data**
- **Promotes transparency**
- **Promotes reproducibility, of sample, design, and analysis**

Many Psychology Findings Not as Strong as Claimed, Study Says

By BENEDICT CAREY AUG. 27, 2015

✉️ Email

📱 Share

🐦 Tweet

📁 Save

➔ More

The past several years have been bruising ones for the credibility of the social sciences. A star social psychologist [was caught](#) fabricating data, leading to more than 50 retracted papers. A top journal published [a study](#) supporting the existence of ESP that was widely criticized. The journal Science pulled a [political science paper](#) on the effect of gay canvassers on voters' behavior because of concerns about faked data.

Now, a painstaking yearslong effort to reproduce 100 studies published in three leading [psychology](#) journals has found that more than half of the findings did not hold up when retested. The analysis was done by research



Data Sharing Efforts across Fields

Public Health Emergencies



Archaeology



Astronomy



Chemistry



Genetics

ClinVar

ClinVar aggregates information about genomic variation and its relationship to human health.

Social Sciences Research



Biology



Medicine

The MMRF
CoMMpass StudySM



POND NETWORK
Province of Ontario Neurodevelopmental Disorders

Ecology



Geology



NHS

*National Institute for
Health Research*



National Institutes of Health
Turning Discovery Into Health

PhARMA

ljaf **efpia**
laura and john arnold foundation™



Sharing Clinical Trial Data
MAXIMIZING BENEFITS, MINIMIZING RISK

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

**BILL & MELINDA
GATES foundation**



Recommendations

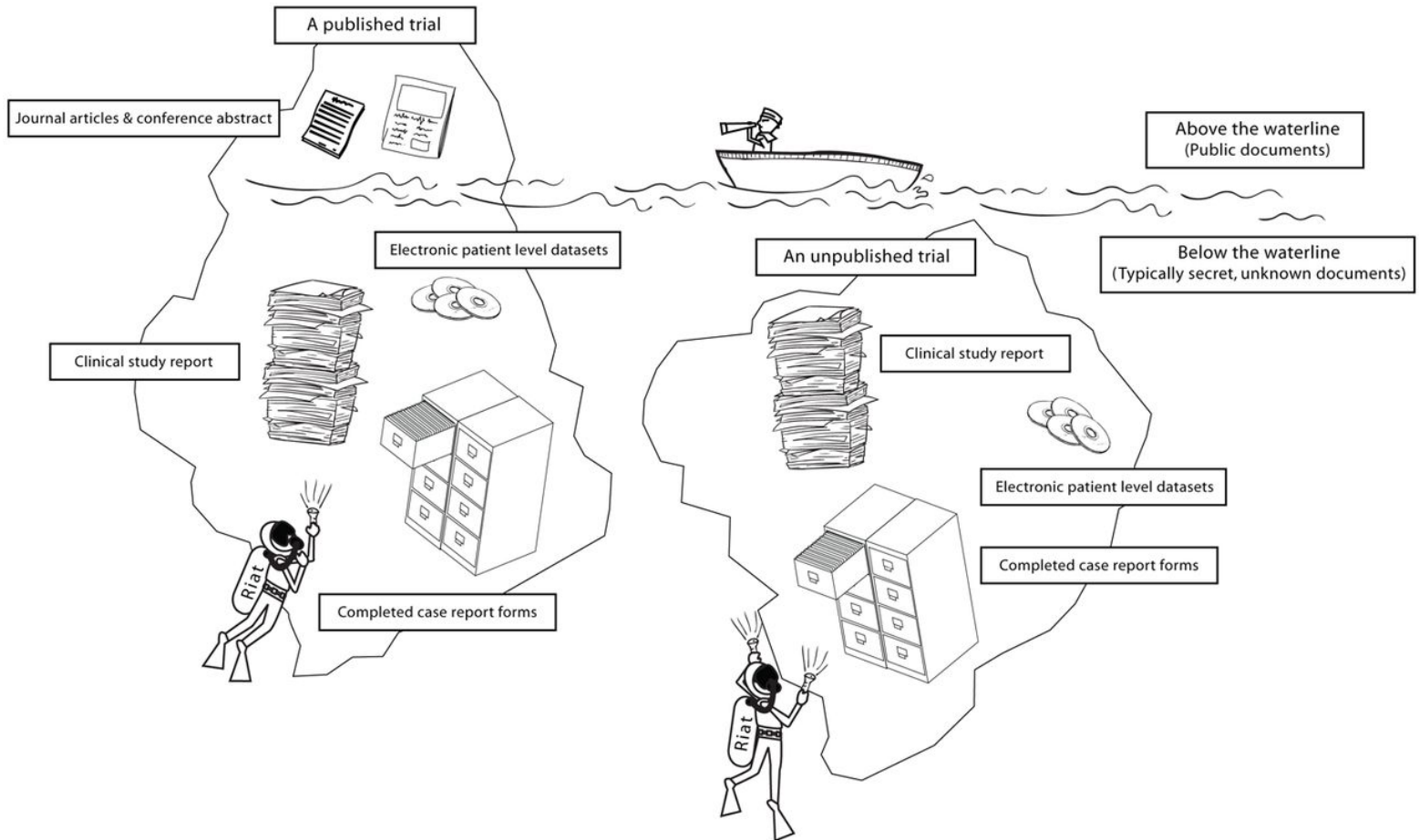
1. **“Stakeholders should foster a culture in which data sharing is the expected norm, and should commit to responsible strategies aimed at maximizing benefits, minimizing risks, and overcoming challenges of sharing trial data”**
2. **Clarified specific points in time by which various types of data are reported/shared**
3. **Data Holder strategies:**
 1. **Data Use Agreement**
 2. **Independent panels, including lay public**
 3. **Transparency**
 4. **Monitor “outcomes”**
4. **More work to be done, need to address key infrastructure, technological, sustainability, and workforce challenges ahead**

Sharing Data

- **Online, posted for general use**
- **Online, publicly accessible data repository**
 - Designated data storage for multiple databases
 - Can be subject-specific, topic-specific, or broad
 - Requires registration and meets accepted security criteria
- **Online secure data sharing platform**
 - A controlled site where access to data can be restricted
- **Non-secure file transfer service**
- **Secure file transfer service**
 - Data are transferred securely via email or CDs

Understanding the world of trial documents

What's above and below the waterline?



Types of Trial Data

- **Summary information**
 - Publications, conference abstracts, CT.gov
- **Clinical study reports**
- **Trial meta-data: data that describes and provides information about other data**
 - Data dictionary, statistical analysis plan, case report forms
- **Individual participant-level data**