

# Preprints: What, Why and How

Harlan M. Krumholz, MD

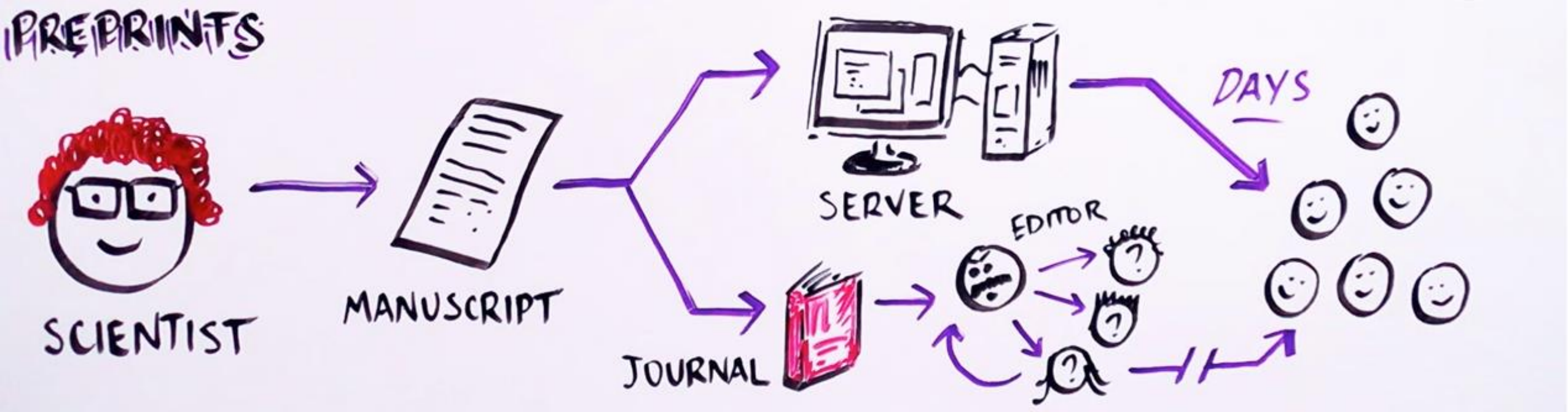
Yale School of Medicine

@hmkyale

# Disclosure

PCORI Board; United HealthCare and Aetna Advisory Committees; Grant from JNJ to distribute clinical trial data; grant from Medtronic, in collaboration with FDA, to improve device surveillance; contracts with CMS to develop performance measures; contract with Blue Cross for technology assessment; AHRQ and NIH grants; Founder, Hugo personal health information platform.

# Preprints



# Preprints



# Not everyone is a fan...

Opinion

EDITORIAL

## The Rush to Publication An Editorial and Scientific Mistake

Howard Bauchner, MD

**The world moves** at a far faster pace than even a decade ago. Instantaneous access to electronic communication via email and social media is available 24 hours a day, virtually anywhere in the world, on the ground and in the air, with video and audio on demand. Thus, no one ever needs to be—or ever is—disconnected from the world.

The speed of communication has clearly affected clinical and laboratory research. There appears to be an increasing rush to publish, or at least to make the results of studies immediately publicly available. It is unclear if flawed science is more common than in the past, but the number of accounts of serious problems with scientific reports appears to be increasing, with more high-profile

tation, and their own internal motivations, often request rapid review and publication by journals. Many journals acquiesce to these requests, in turn, placing more pressure on peer reviewers, most of whom are investigators, to complete review in a matter of days, and more pressure on journal staff and resources to expedite article preparation and distribution.

New interest in preprint servers in clinical medicine increases the likelihood of premature dissemination and public consumption of clinical research findings prior to rigorous evaluation and peer review. At *JAMA* and throughout the JAMA Network journals, the conclusions and interpretations of many research articles change substantially between the initially sub-

Progress in human health is measured in years, not days, weeks, or months. Major breakthroughs in clinical medicine are rare, with very few research findings likely to be implemented immediately. No drug or device, regardless of how effective, is likely to improve health outcomes more than many common and important clinical practices such as measuring blood pressure and treating hypertension with well-known drugs that have been proven to be safe and effective.

It usually takes years for interventions that improve patient outcomes to become part of routine practice and few novel interventions are likely to be more important than those already known to be effective. Improving the health of the world's population has little to do with the speed of publication (except in the case of major public health emergencies),<sup>12</sup> but rather with effective interventions that have been properly tested, appropriate implementation of known or new interventions, and sustainable improvements in health systems.

For most articles, public consumption of research findings prior to peer review will have little influence on health, but for some articles, the effect could be devastating for some patients if the results made public prior to peer review are wrong or incorrectly interpreted.

Sacrificing adequate and thoughtful peer review and editorial assessment is a mistake for research in medicine. Timely assessment and dissemination of medical research findings is certainly important, but for most articles, rushing to publication in days or weeks will not improve health outcomes.



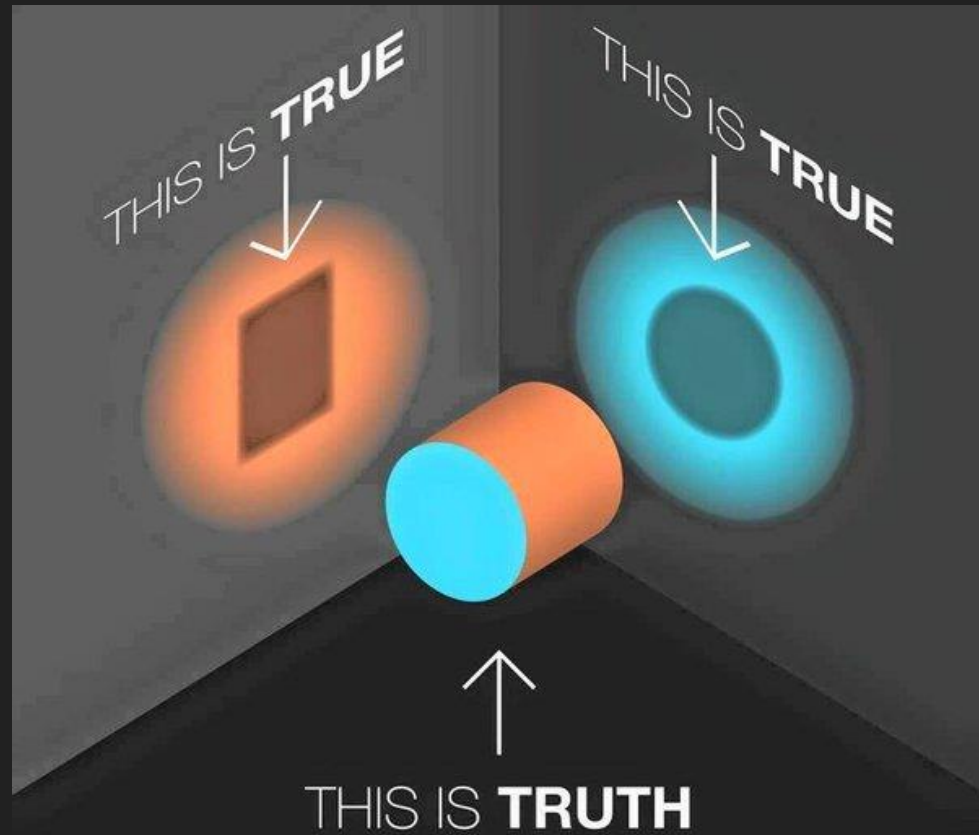
# Dual Nature of Medical Journals



dual

**PEER REVIEW  
MATTERS**

# Please try to keep an open mind



# What problem are we trying to solve?

- How do we easily and rapidly archive and share information with other scientists to accelerate research, enhance collaboration, reduce waste, increase transparency?

**BUT...** IS THERE A **BETTER** WAY?

IN **1991**:

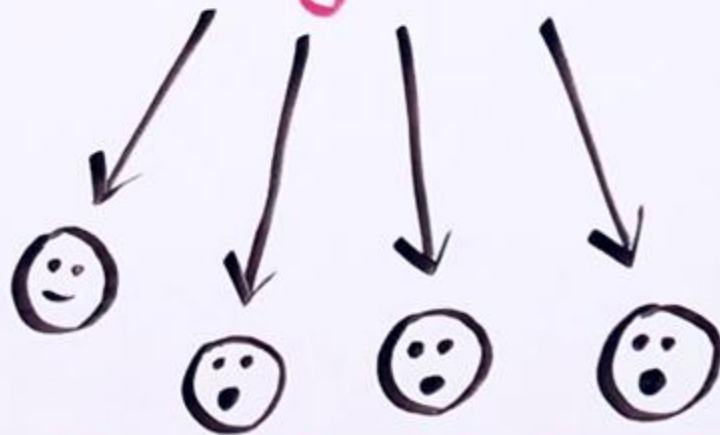
PAUL GINSPARG  
PHYSICIST



arXiv.org

A **PREPRINT**  
SERVER

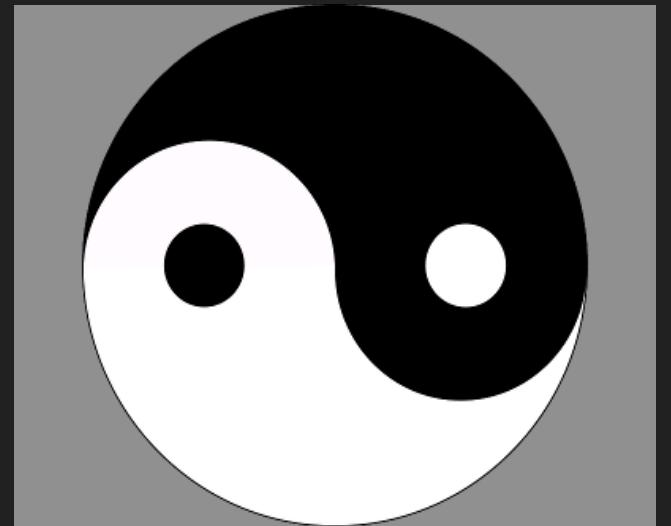
arXiv: > 100,000 papers  
EACH YEAR!



== **RAPID COMMUNICATION**

# Preprint Servers

- Complementary to, and not a replacement for, peer review journals.



# Science for Scientists



# DOI (Digital Object Identifier)

- DOI is unique alphanumeric string assigned by registration agency (**International DOI Foundation**) to identify content and provide a persistent link to its location on the Internet.





# Pre-peer review results already released...

- Major journals already allow pre-peer review release of information.

For example...

- Medical meetings
- Clinicaltrials.gov
- Press releases

But...

- Information often incomplete
- May not be citable
- May not be searchable

And...

- No opportunity for community comments/dialogue

# IMPROVE-IT Trial

Study Results Reported

Results First Announced:  
August 28, 2014

Trial First Presented:  
November 17, 2014

Paper published:  
June 3, 2015

ClinicalTrials.gov  
A service of the U.S. National Institutes of Health

Find Studies | About Studies | Submit Studies | Resources | About Site

Home > Search Results > Study Record Detail

Trial record 2 of 21 for: IMPROVE-IT  
Previous Study | Return to List | Next Study

**IMPROVE-IT: Examining Outcomes in Subjects With Acute Coronary Syndrome: Vytorin (Ezetimibe/Simvastatin) vs Simvastatin (P04103)**

This study has been completed.

Sponsor: Merck Sharp & Dohme Corp.  
Information provided by (Responsible Party): Merck Sharp & Dohme Corp.

ClinicalTrials.gov Identifier: NCT00202878  
First received: September 13, 2005  
Last updated: May 10, 2017  
Last verified: May 2017  
History of Changes

Full Text View | Tabular View | Study Results | Disclaimer | How to Read a Study Record

Results First Received: August 28, 2015

<b>Study Type:</b>	Interventional
<b>Study Design:</b>	Allocation: Randomized; Intervention Model: Parallel Assignment; Masking: Triple (Participant, Investigator, Outcomes Assessor); Primary Purpose: Treatment
<b>Conditions:</b>	Hypercholesterolemia Myocardial Infarction
<b>Interventions:</b>	Drug: ezetimibe/simvastatin Drug: simvastatin Drug: Placebo for simvastatin 40 mg Drug: Placebo for ezetimibe 10 mg/simvastatin 40 mg combination

Participant Flow  
Hide Participant Flow

Recruitment Details  
Key information relevant to the recruitment process for the overall study, such as dates of the recruitment period and locations  
Adult men and women presenting with non-ST segment elevation myocardial infarction (NSTEMI), STEMI, or hospitalized, documented unstable angina (UA) whom a percutaneous coronary intervention (PCI) was planned as management for the qualifying acute coronary syndrome (ACS) event were eligible for entry into the trial.

Pre-Assignment Details  
Significant events and approaches for the overall study following participant enrollment, but prior to group assignment  
Study continued until a minimum of 5,250 participants had a primary endpoint event and each participant was followed for a minimum of 2.5 years.

Reporting Groups

	Description
Ezetimibe/Simvastatin	One Ezetimibe 10 mg/simvastatin 40 mg combination tablet and two simvastatin 40 mg placebo tablets once per day.
Simvastatin	One simvastatin 40 mg tablet, one ezetimibe/simvastatin combination 10/40 placebo tablet and one simvastatin 40 mg placebo tablet once per day.

Participant Flow: Overall Study

	Ezetimibe/Simvastatin	Simvastatin
STARTED	907	907
COMPLETED	686	660
NOT COMPLETED	219	2217
Death	94	96
Only Vital Status Known	37	36
Lost to Follow-up	4	4
Site Closure	0	0
Withdrawal by Subject	79	80

# Lack options for less publishable products...

- Protocols/Technical reports
- Quality innovations
- White papers

# Unpublished Trial Data

- Many trials are never reported – or are delayed by years.

**We will work towards a timeframe of 12 months from primary study completion (the last visit of the last subject for collection of data on the primary outcome) as the global norm for summary results disclosure.**

#### **Registration of clinical trials**

Before any clinical trial is initiated (at any Phase) its details must be registered in a publicly available, free to access, searchable clinical trial registry complying with WHO's international agreed standards ([www.who.int/ictrp](http://www.who.int/ictrp)). The clinical trial registry entry must be made before the first subject receives the first medical intervention in the trial (or as soon as possible afterwards). Clinical trial registry records should be updated as necessary to include final enrolment numbers achieved, and the date of primary study completion (defined as the last data collection timepoint for the last subject for the primary outcome measure). If clinical trials are terminated, their status should be updated to note the date of termination, and to report the numbers enrolled up to the date of termination.

Completeness and accuracy of the clinical trial registry records can be a limiting factor for use of information from the registries, and it is encouraged that care is taken to ensure good quality registry entries.

#### **Reporting timeframes for clinical trials**

We jointly agree that summary results of clinical trials should be made publicly available in a timely manner following primary study completion. There are two main modalities for this to occur. By posting to the results section of the clinical trial registry and by journal publication. We will work towards a timeframe of 12 months from primary study completion (the last visit of the last subject for collection of data on the primary outcome) as the global norm for summary results disclosure. As timelines for publication in a journal are not fully within the control of the sponsor or investigator, this joint statement focuses on use of registries – such as [clinicaltrials.gov](http://clinicaltrials.gov) and EU-CTR - to meet this results disclosure expectation. Publication in a journal is also an expectation, with an indicative timeframe of 24 months from study completion to allow for peer review etc. Access to a sufficiently detailed clinical trial protocol is necessary in order to be able to interpret summary results. Therefore we also encourage development of requirements that the protocols are made publicly available no later than the time of the summary results disclosure as part of the clinical trial registry summary results information (including amendments approved by ethics committees/institutional review boards, and either as uploaded electronic document formats such as pdfs or links to the pdf).

At the time of the initial grant submission, the plan for public disclosure of results should be included, including specific time bound commitments. Reasonable funds to enable compliance with these considerations is a cost eligible item in clinical trial budgets.

#### **Trial ID in clinical trial publication**

The Trial ID or registry identifier code/number should be included in all publications of clinical trials, and should be provided as part of the abstract to PubMed and other bibliographic search databases for easy linking of trial related publications with clinical trial registry site records. This is essential for linking journal publications with registry records.

#### **Registration and reporting of past trials**

Reporting of previous trials realises the value of funding; therefore the contribution made from reporting previous trials, whatever their results, will be considered in the assessment of a funding proposal. When a PI applies for new funding, they may be asked to provide a list of all previous trials on which they were PI within a specified timeframe and their reporting status, with an explanation where trials have remained unreported.



# Jeremy Farrar, Director Wellcome Trust

- “Not only will this help ensure that these research findings are more discoverable, but it will also reduce reporting biases, which currently favor publication of trials which have a positive outcome.”



# Trevor Mundel, Pres GH Gates Foundation

- "It's a 21st-century best practice – and an essential part of the social contract that underlies medical research – that clinical trial data should be made publicly available less than one year after a clinical trial's completion."





**In general, results information must be submitted no later than one year after the completion date of the applicable drug clinical trial.**



DEPARTMENT OF HEALTH AND HUMAN SERVICES

42 CFR Part 11

[Docket Number NIH-2011-0003]

RIN: 0925-AA55

Clinical Trials Registration and Results Information Submission

**AGENCY:** National Institutes of Health, Department of Health and Human Services.

**ACTION:** Final Rule.

**SUMMARY:** This final rule details the requirements for submitting registration and summary results information, including adverse event information, for specified clinical trials of drug products (including biological products) and device products and for pediatric postmarket surveillances of a device product to ClinicalTrials.gov, the clinical trial registry and results data bank operated by the National Library of Medicine (NLM) of the National Institutes of Health (NIH). This rule provides for the expanded registry and results data bank specified in Title VIII of the Food and Drug Administration Amendments Act of 2007 (FDAAA) to help patients find trials for which they might be eligible, enhance the design of clinical trials and prevent duplication of unsuccessful or unsafe trials, improve the evidence base that informs clinical care, increase the efficiency of drug and device development processes, improve clinical research practice, and build public trust in clinical research. The requirements apply to the responsible

party (meaning the sponsor or designated principal investigator) for certain clinical trials of drug products (including biological products) and device products that are regulated by the Food and Drug Administration (FDA) and for pediatric postmarket surveillances of a device product that are ordered by FDA.

**DATES:** These regulations are effective on January 18, 2017. Additional information on the effective date and the compliance date can be found in Section IV.F.

Only 29% of completed clinical trials conducted by the faculty at major academic centers were published within two years of completion and only 13% reported results on ClinicalTrials.gov

RESEARCH

OPEN ACCESS

## Publication and reporting of clinical trial results: cross sectional analysis across academic medical centers

Ruijun Chen,<sup>1</sup> Nihar R Desai,<sup>2,3</sup> Joseph S Ross,<sup>3,4,5,6</sup> Weiwei Zhang,<sup>3</sup> Katherine H Chau,<sup>1</sup> Brian Wayda,<sup>7</sup> Karthik Murugiah,<sup>8</sup> Daniel Y Lu,<sup>2</sup> Amit Mittal,<sup>8</sup> Harlan M Krumholz<sup>3,5,6</sup>

**ABSTRACT**  
**OBJECTIVE** To determine rates of publication and reporting of results within two years for all completed clinical trials registered in ClinicalTrials.gov across leading academic medical centers in the United States.  
**DESIGN** Cross sectional analysis.  
**SETTING** Academic medical centers in the United States.  
**PARTICIPANTS** Academic medical centers with 40 or more completed interventional trials registered on ClinicalTrials.gov.  
**METHODS** Using the Aggregate Analysis of ClinicalTrials.gov database and manual review, we identified all interventional clinical trials registered on ClinicalTrials.gov with a primary completion date between October 2007 and September 2010 and with a lead investigator affiliated with an academic medical center.  
**MAIN OUTCOME MEASURES** The proportion of trials that disseminated results, defined as publication or reporting of results on ClinicalTrials.gov, overall and within 24 months of study completion.  
**RESULTS** We identified 4347 interventional clinical trials across 51 academic medical centers. Among the trials, 1005 (23%) enrolled more than 100 patients, 1216 (28%) were double blind, and 2169 (50%) were phase II through IV. Overall, academic medical centers disseminated results for 2892 (66%) trials, with 1560 (35.9%) achieving this within 24 months of study completion. The proportion of clinical trials with results disseminated within 24 months of study completion ranged from 16.2% (6/37) to 55.3% (57/103) across academic medical centers. The proportion of clinical trials published within 24 months of study completion ranged from 10.8% (4/37) to 40.3% (31/77) across academic medical centers, whereas results reporting on ClinicalTrials.gov ranged from 1.6% (2/122) to 40.7% (72/177).  
**CONCLUSIONS** Despite the ethical mandate and expressed values and mission of academic institutions, there is poor performance and noticeable variation in the dissemination of clinical trial results across leading academic medical centers.  
**Introduction** Randomized clinical trials are the ideal means for evaluating the efficacy and safety of medical drugs and devices. Timely dissemination of the findings from clinical trials is a prerequisite for ensuring that clinical decisions made by patients and physicians reflect the best scientific evidence, and that future scientific investigation benefits from previous inquiry. Dissemination is principally achieved through publication in peer reviewed biomedical journals as well as through public reporting of results on clinical trial registries.<sup>1-4</sup> However, a large body of research found that between 29% and 50% of clinical trials remain unpublished, sometimes years after study completion.<sup>5,6</sup> Similarly, studies have shown that the results of many clinical trials are not reported promptly on ClinicalTrials.gov.<sup>7-10</sup> Academic medical centers play a critical role in the clinical trials research enterprise. However, studies suggest that academically based investigators perform suboptimally in publishing<sup>11</sup> and reporting trial results.<sup>12-16</sup>  
We carried out a comprehensive examination of the rates of publication and reporting of results within two years for all completed clinical trials registered in ClinicalTrials.gov across more than 50 academic medical centers in the United States with active clinical research programs.  
**Methods**  
**Data source and study sample** We used data from ClinicalTrials.gov through the Aggregate Analysis of ClinicalTrials.gov (AACT) database, reflecting data downloaded as of 27 September 2013, under the Clinical Trials Transformation Initiative. We identified all interventional clinical trials registered on

**WHAT IS ALREADY KNOWN ON THIS TOPIC**  
Timely dissemination of clinical trial results is required to honor the commitment of study participants, advance the research enterprise, and improve clinical care, but little is known about the performance of academic medical centers in this endeavor. Previous limited studies have shown that between 25% and 50% of clinical trials remain unpublished, sometimes years after completion, and the performance of academically based investigators in publishing and reporting of trial results is suboptimal.

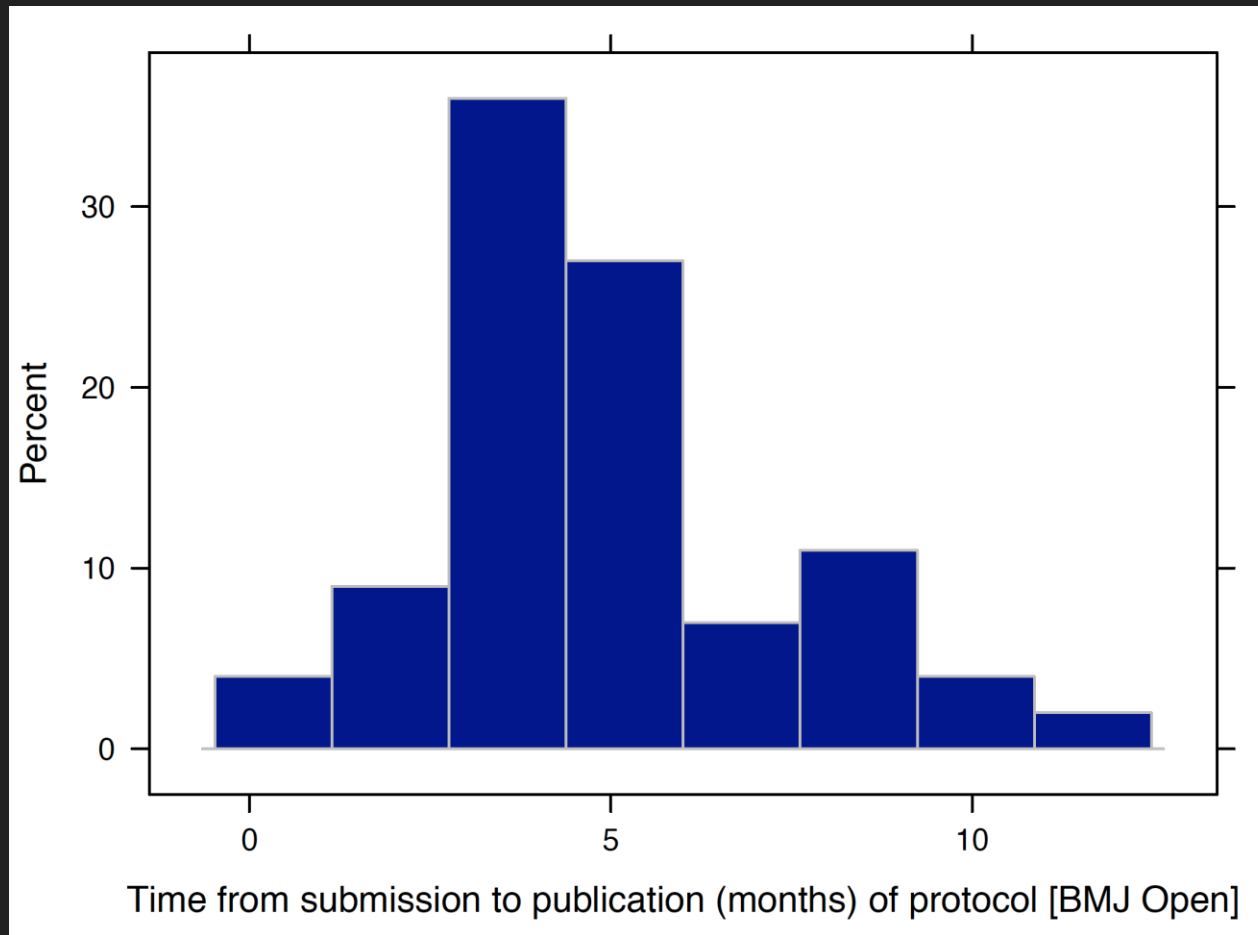
**WHAT THIS STUDY ADDS**  
Academic medical centers showed noticeable variation and poor performance in the dissemination of clinical trial results. Only 29% of completed clinical trials conducted by the faculty at major academic centers were published within two years of completion and only 13% reported results on ClinicalTrials.gov. Additional tools and mechanisms are needed to rectify this lack of timely reporting and publication, as they impair the research enterprise and threaten to undermine evidence based clinical decision making.

thebmj | BMJ 2016;352:i637 | doi:10.1136/bmj.i637 1

And then there is speed...

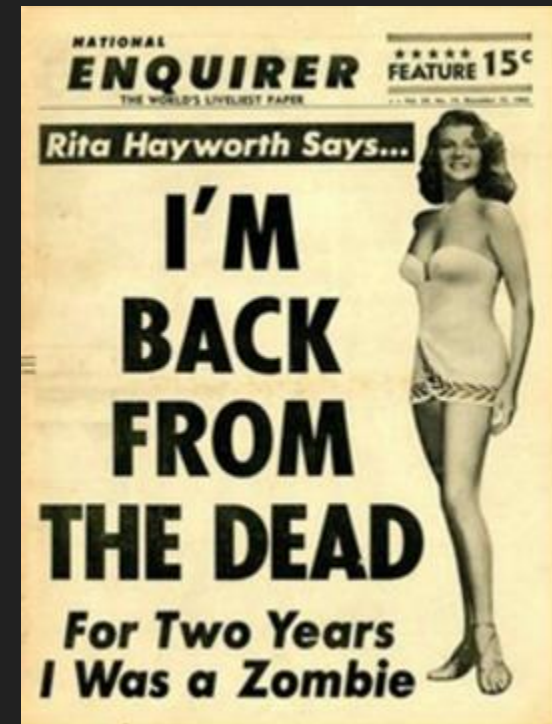


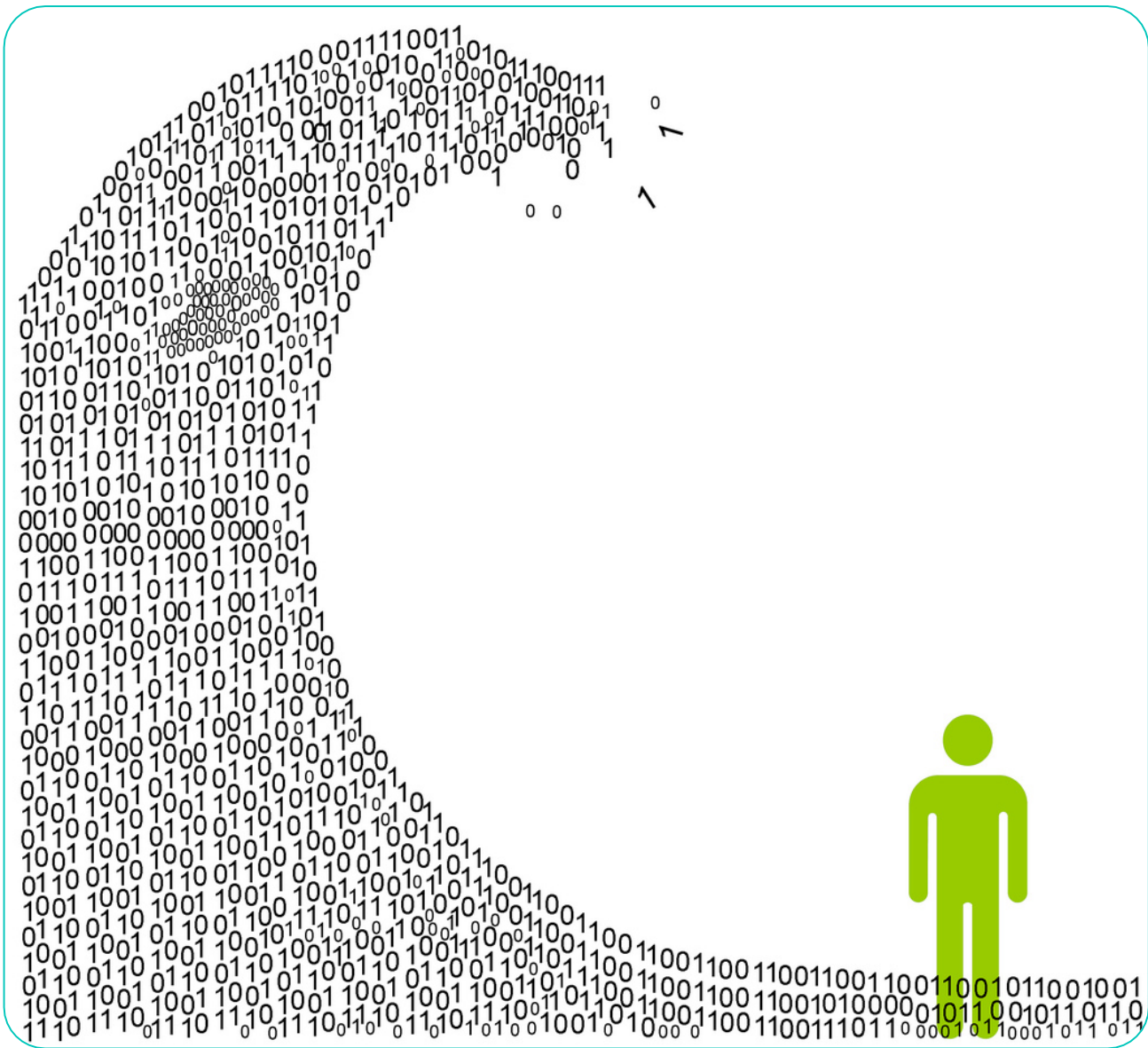
# Protocol Papers



# Concern: Fake News

- There is already too much fake information out there.





**Concern: There is already too much information.**



About  
volume...

“

**It's not information  
overload. It's filter  
failure.**

CLAY SHIRKEY

**Concern:  
Public may act  
prematurely.  
Harm may  
accrue.**



# Concern: May undermine Clinicaltrials.gov

*ClinicalTrials.gov*

A service of the U.S. National Institutes of Health

# Risk Mitigation

- High-level screen
- IRB-approved or exempted
- Require posting on [clinicaltrials.gov](https://clinicaltrials.gov)
- Corresponding author: ORCID
- Labeling and watermarks

# Concern: Publication

- Will this jeopardize peer-review publication?



# Journal Policies

Journal	Publisher	Policy type	Policy text
The BMJ (formerly British Medical Journal)	BMJ Publishing Group Ltd	Compatible	Preprint ("Original manuscript submitted to BMJ.") can be posted.

# Journal Policies

Nature Publishing  
Group

Compatible

The policy states "Neither conference presentations nor posting on recognized preprint servers constitute prior publication," and an editorial explains: "Nature never wishes to stand in the way of communication between researchers.[...] Communication between researchers includes not only conferences but also preprint servers. The ArXiv preprint server is the medium of choice for (mainly) physicists and astronomers who wish to share drafts of their papers with their colleagues, and with anyone else with sufficient time and knowledge to navigate it. [...] If scientists wish to display drafts of their research papers on an established preprint server before or during submission to Nature or any Nature journal, that's fine by us."



# PREPRINTS!

IMMEDIATE ACCESS!



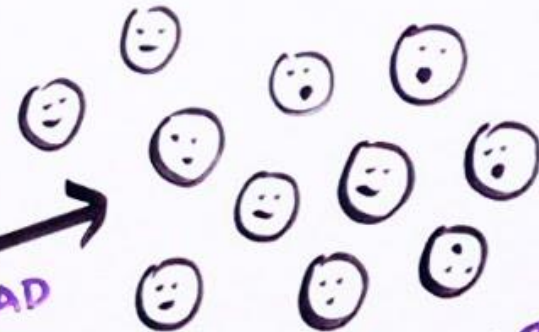
ACCELERATE SCIENTIFIC PROGRESS!



OPEN ACCESS!



DOWNLOAD



MORE FEEDBACK



PRIORITY OF WORK!



JOB\$

RECENT ACCOMPLISHMENTS!

# NIH encourages the use of preprints

Preprints and other interim research products "can be

[Home](#) > [Funding for researchers](#) > [Research features](#) > We accept preprints in grant applications: new guidance for researchers

## We accept preprints in grant applications: new guidance for researchers

Category: [Research Feature](#)

 30 May 2017

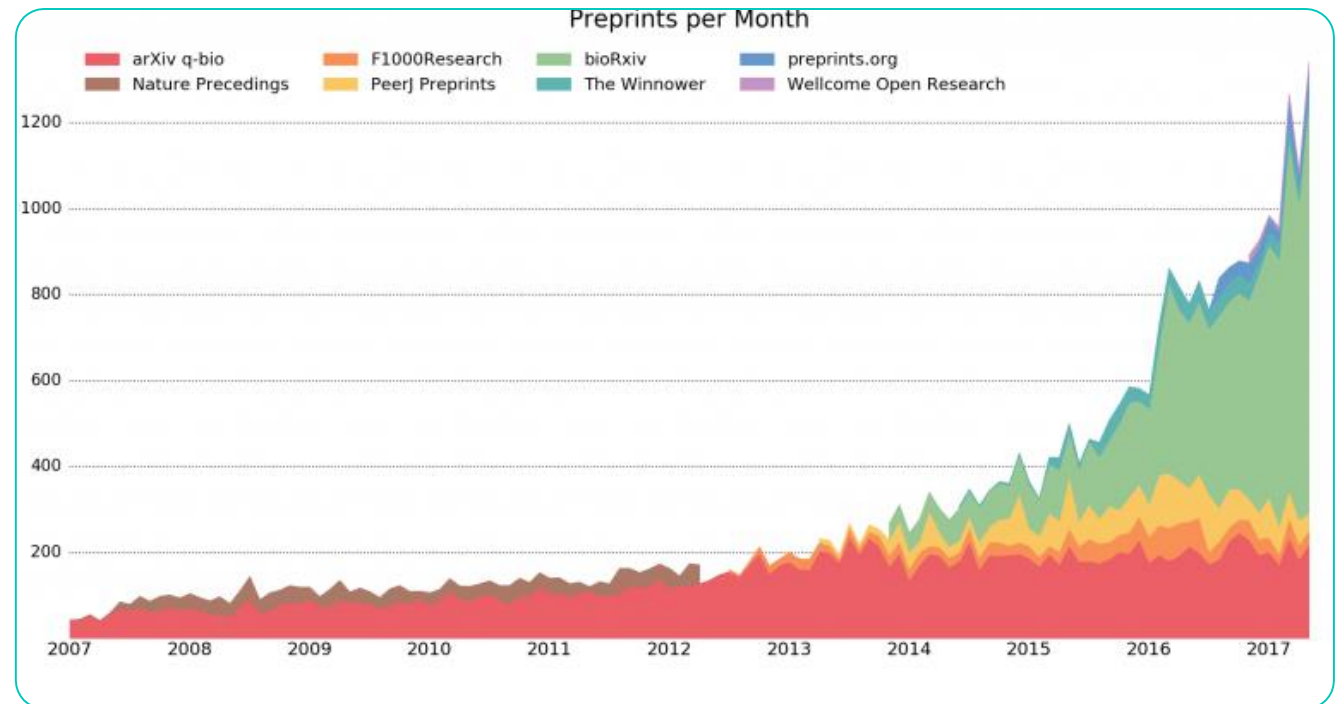
 Cancer Research UK



Cancer Research UK encourages the use of preprints



# Growing use in biology: bioRxiv



# Yale Data Open Access (YODA) Project

the  
**YODA**  
PROJECT

Forging a unified  
scientific community

f t

ABOUT REQUEST TRIALS FAQs LOG IN

The real voyage of discovery consists not  
in seeking new landscapes, but in having  
new eyes.

*Marcel Proust*

## OUR MISSION

The Yale University Open Data Access (YODA) Project's mission is to advocate for the responsible sharing of clinical research data, open science, and research transparency. The Project is committed to supporting research focused on improving the health of patients and informing science and public health. The YODA Project can only improve with your feedback. Please share your comments and ideas.

CONTACT US

## REQUEST DATA

Are you ready to request data? To date, 191 trials have been identified as available. The YODA Project and partnered Data Holders continue to identify and add more.

GET STARTED

## OUR MODEL

The YODA Project seeks mutually beneficial partnerships with Data Holders, promoting independence, responsible conduct of research, good stewardship of data, and the generation of knowledge in the best interest of society. To participate, each Data Holder must transfer full jurisdiction over data access to the YODA Project.

LEARN MORE

## SHARE FINDINGS

The YODA Project maintains MedArXiv, a free preprint service for the medicine and health sciences, to accelerate the scientific enterprise by facilitating results reporting, documenting the provenance of ideas, and fostering scientific communication.

LEARN MORE

“The scientific community [is] arguably the most powerful collective enterprise in human history.”

*Atul Gawande*



**When is the right time?**

When is the  
right time?



It's Time

# Medical Preprint Server

- Community resource
- Stewards not owners
- Iterative learning



# Goal

- Not to establish a server
- But to improve and accelerate science, promote collaboration, enhance transparency, reduce waste
- Improve over time

# Ask...

- Journals should allow co-existence of pre-print servers and not penalize scientists who use them.
- Is there really a difference from presenting at a meeting?

Tell us what you think...

[medarxiv@yale.edu](mailto:medarxiv@yale.edu)

