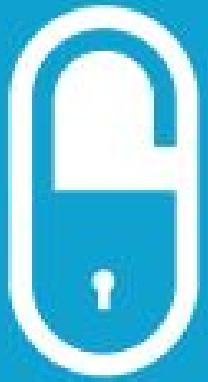


OpenTrials

All the Data, on All the Trials, Linked

Jessica Fleminger and Ben Meghreblian

1. introduction
2. we (still) have a problem
3. OpenTrials overview
4. techy stuff
5. data sources
6. user examples
7. demo
8. Q&A

opentrials@okfn.org

OPEN KNOWLEDGE

EBM Data Lab

University of Oxford

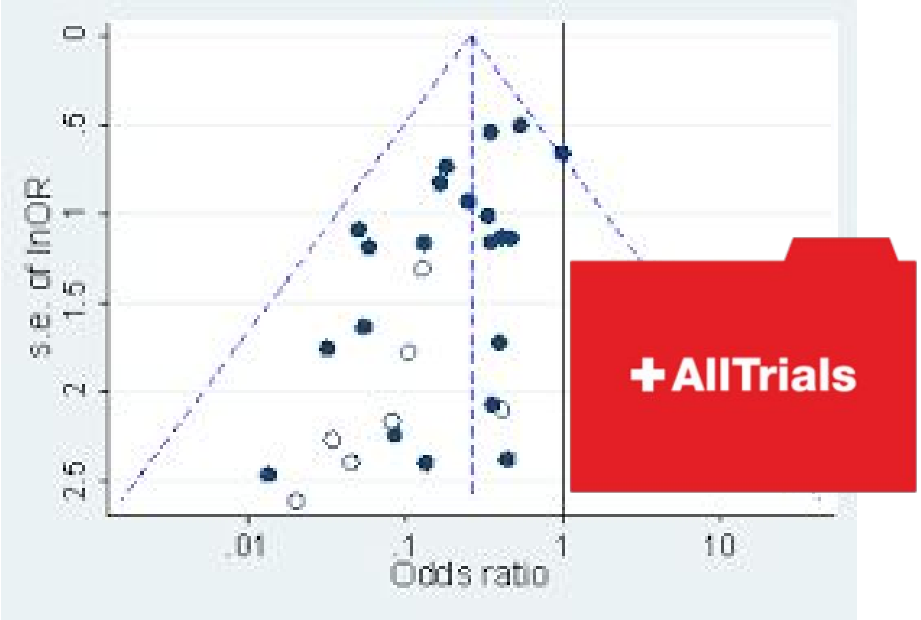
CENTER FOR OPEN SCIENCE

laura and john arnold foundation®

opentrials@okfn.org

audiinsperation (CC BY)

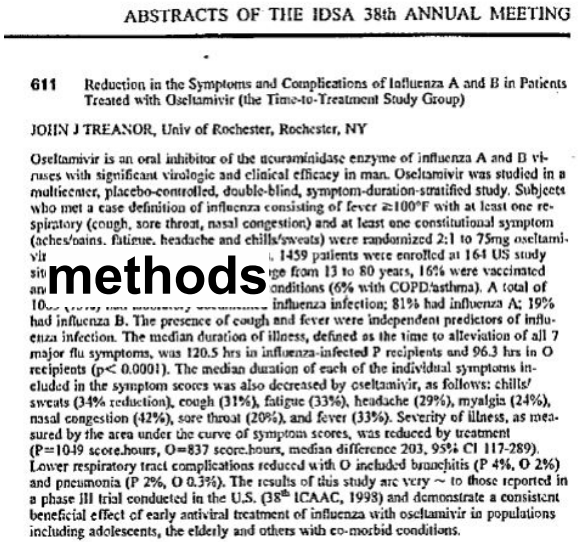
we have a **publication bias** problem



Anderson et al. N Engl J Med. 2015 Mar 12;372(11):1031-9. doi: 10.1056/NEJMsa1409364.
Prayle AP et al. BMJ. 2012 Jan 3;344:d7373. doi: 10.1136/bmj.d7373.

opentrials@okfn.org

we have an **information dissemination** problem



ABSTRACTS OF THE IDSA 38th ANNUAL MEETING

611 Reduction in the Symptoms and Complications of Influenza A and B in Patients Treated with Oseltamivir (the Time-to-Treatment Study Group)

JOHN J TREANOR, Univ of Rochester, Rochester, NY

Oseltamivir is an oral inhibitor of the neuraminidase enzyme of influenza A and B viruses with significant virologic and clinical efficacy in man. Oseltamivir was studied in a multicenter, placebo-controlled, double-blind, symptom-duration-stratified study. Subjects who met a case definition of influenza consisting of fever ≥100°F with at least one respiratory (cough, sore throat, nasal congestion) and at least one constitutional symptom (aches/pains, fatigue, headache and chills/sweats) were randomized 2:1 to 75mg oseltamivir

methods

1459 patients were enrolled at 164 US study sites. Age from 13 to 80 years, 16% were vaccinated and 10[illegible] conditions (6% with COPD/asthma). A total of [illegible] influenza infection; 81% had influenza A; 19% had influenza B. The presence of cough and fever were independent predictors of influenza infection. The median duration of illness, defined as the time to alleviation of all 7 major flu symptoms, was 120.5 hrs in influenza-infected P recipients and 96.3 hrs in O recipients ($p < 0.0001$). The median duration of each of the individual symptoms included in the symptom scores was also decreased by oseltamivir, as follows: chills/sweats (34% reduction), cough (31%), fatigue (33%), headache (29%), myalgia (24%), nasal congestion (42%), sore throat (20%), and fever (33%). Severity of illness, as measured by the area under the curve of symptom scores, was reduced by treatment (P=1049 score.hours, O=837 score.hours, median difference 203, 95% CI 117-289). Lower respiratory tract complications reduced with O included bronchitis (P 4%, O 2%) and pneumonia (P 2%, O 0.3%). The results of this study are very ~ to those reported in a phase III trial conducted in the U.S. (38th ICAAC, 1998) and demonstrate a consistent beneficial effect of early antiviral treatment of influenza with oseltamivir in populations including adolescents, the elderly and others with co-morbid conditions.

vs.

Jefferson T et al. BMJ. 2014 Apr 9;348:g2545. doi: 10.1136/bmj.g2545.

globalismpictures (CC BY)

opentrials@okfn.org

we have a **discrepancy** problem

FDA

COMPARE

TRACKING SWITCHED OUTCOMES IN CLINICAL TRIALS

Turner EH et al. N Engl J Med 2008; 358:252-260 doi: 10.1056/NEJMsa065779.

myrfa (CC0)

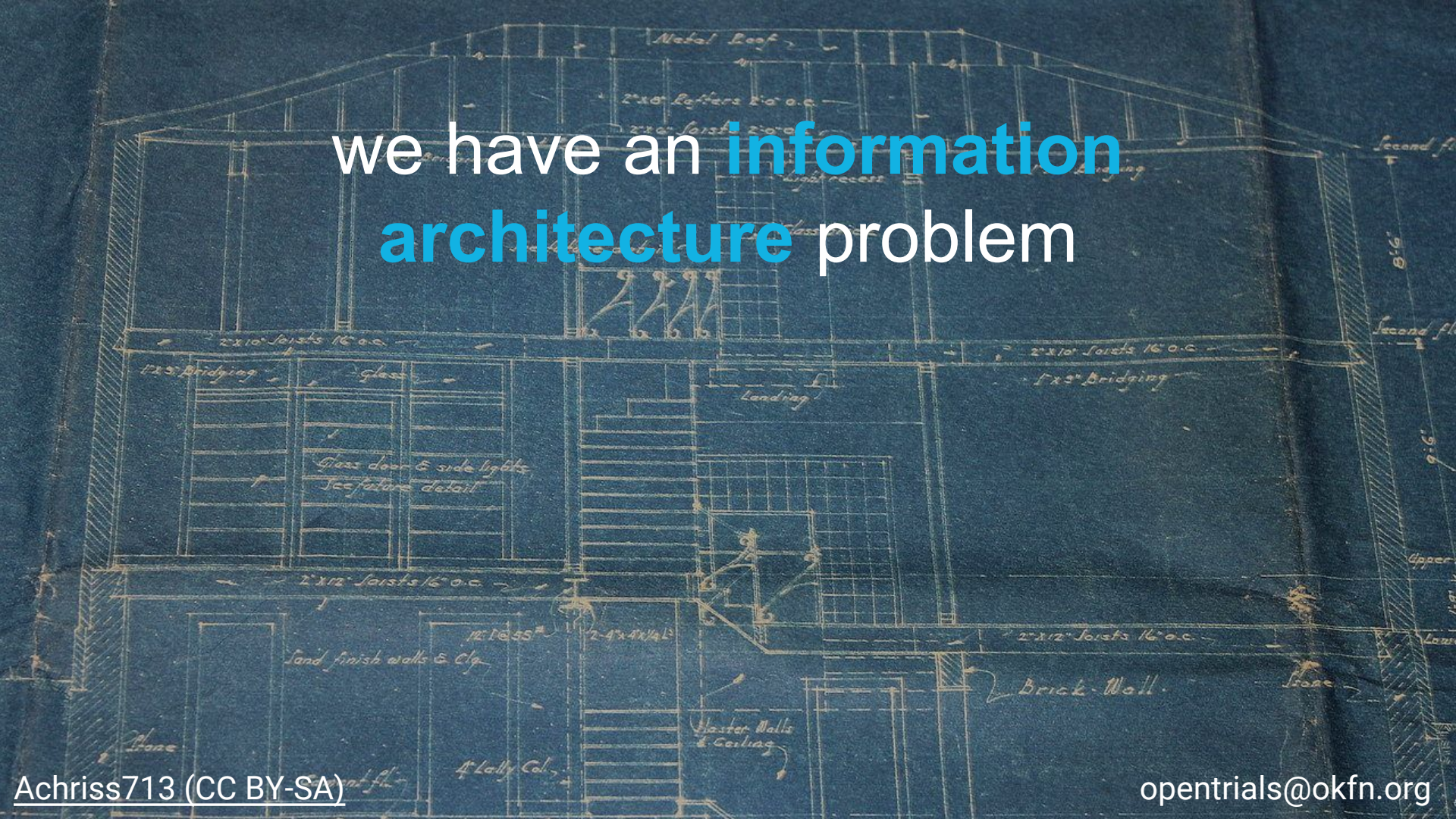
opentrials@okfn.org

we have a **publication bias** problem

we have an **information dissemination** problem

we have a **discrepancy** problem

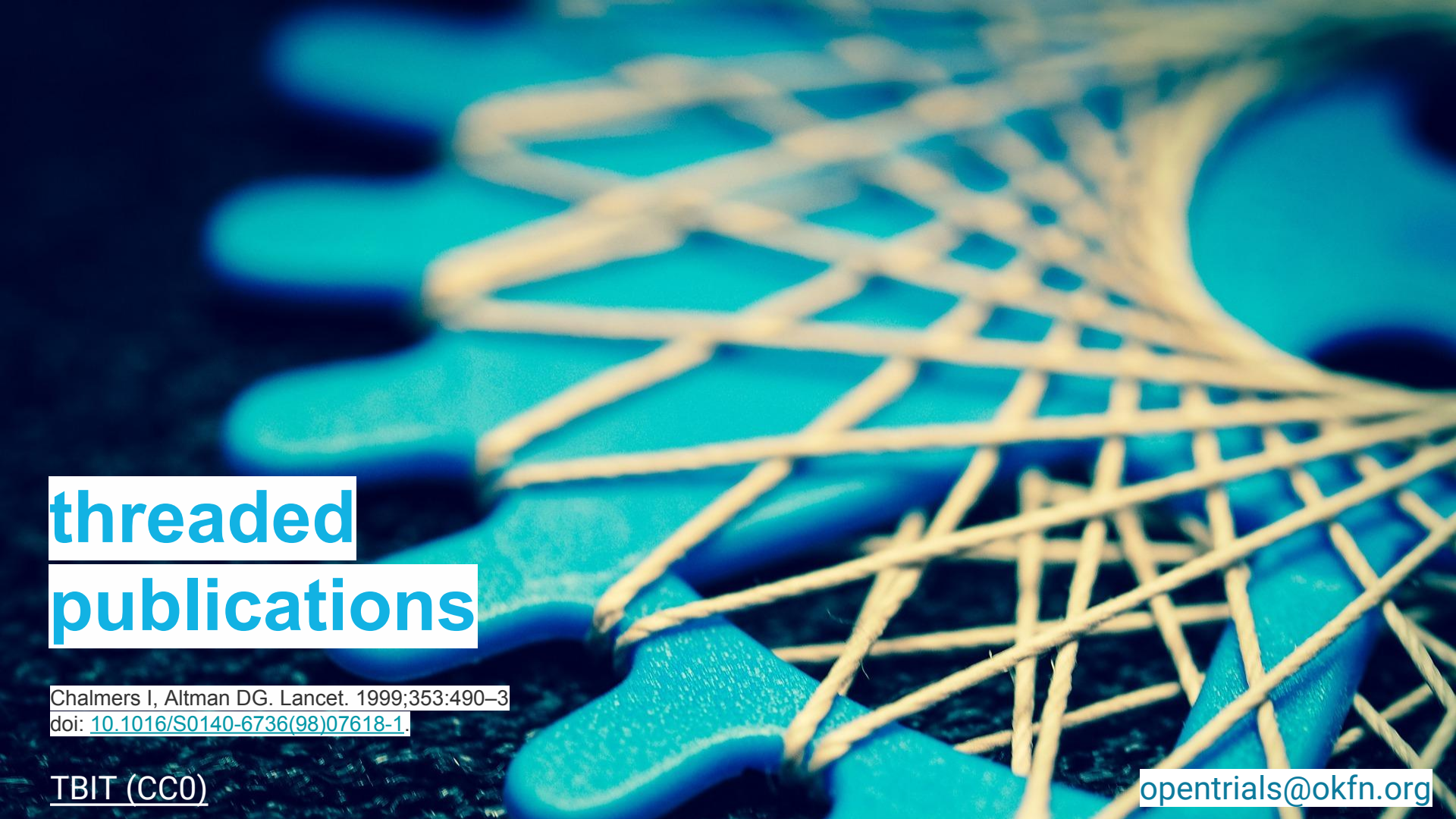
opentrials@okfn.org



we have an **information architecture** problem

Achriss713 (CC BY-SA)

opentrials@okfn.org

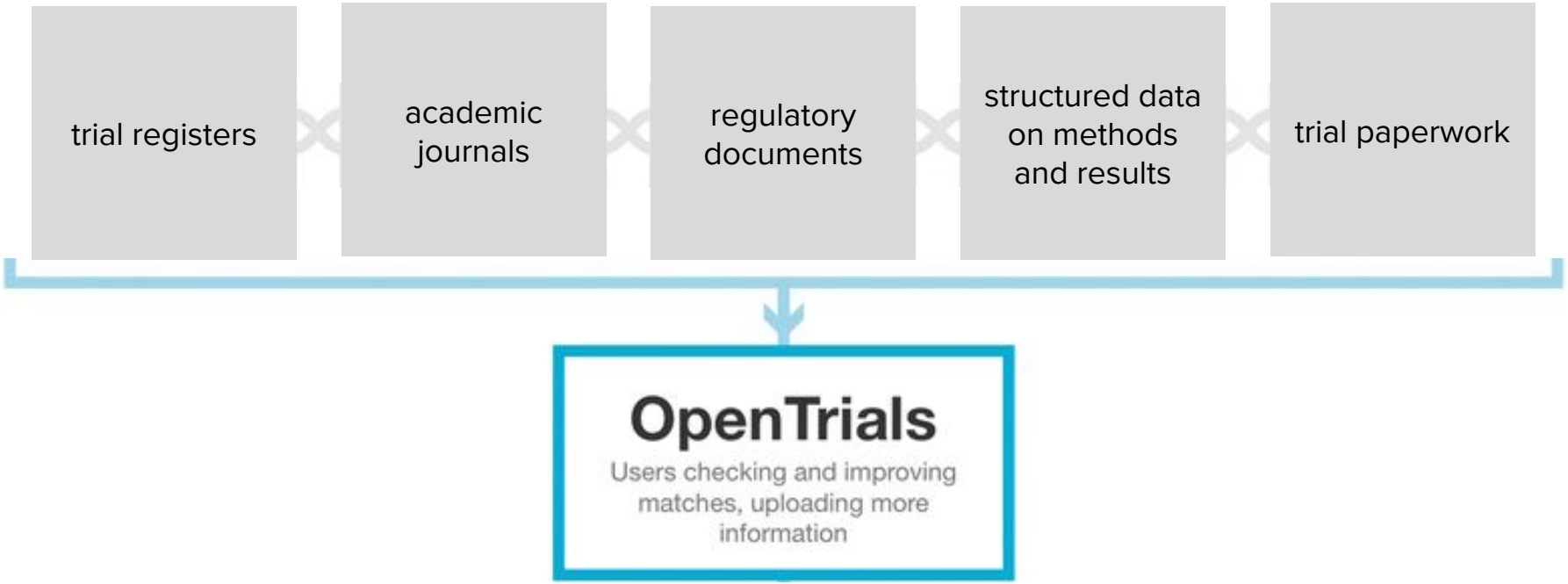


threaded publications

Chalmers I, Altman DG. Lancet. 1999;353:490–3
doi: 10.1016/S0140-6736(98)07618-1.

TBIT (CC0)

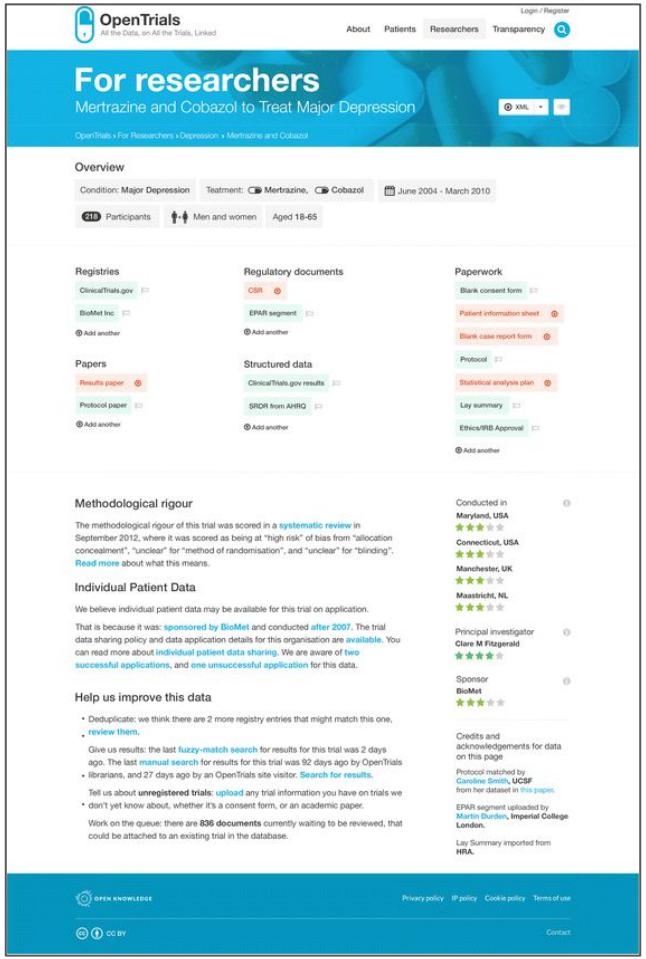
opentrials@okfn.org



opentrials@okfn.org

release early
release often

opentrials@okfn.org



opentrials@okfn.org

audit publication bias

myrfa (CC0)

opentrials@okfn.org

host/link important trial documents

Pexels (CC0)

opentrials@okfn.org

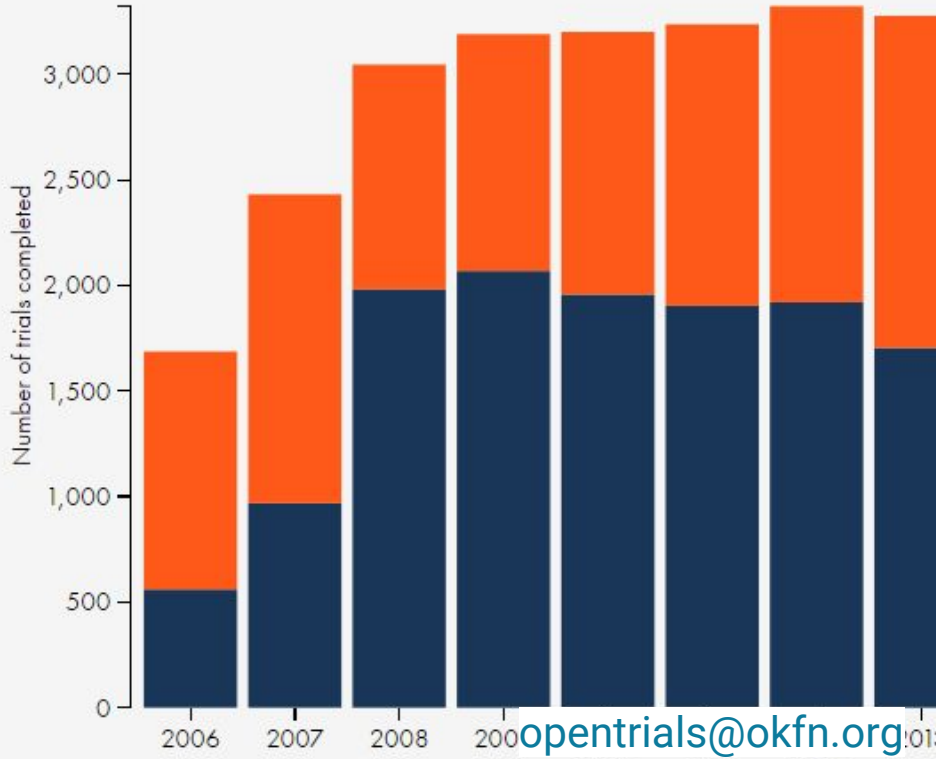
trial sponsors with the most unreported trials registered on
n a sponsor's name to find out whether it's getting better at
s - or worse.

r	Trials missing results	Total eligible trials	Percent missing
	285	435	65.5%
aceuticals	201	534	37.6%
Institute (NCI)	194	558	34.8%
que - Hôpitaux de Paris	186	292	63.7%
e	183	809	22.6%
y			
Unive			63.3%
ical Trials in Oncology	129	160	80.6%

track sponsors and organisations

Trials by year

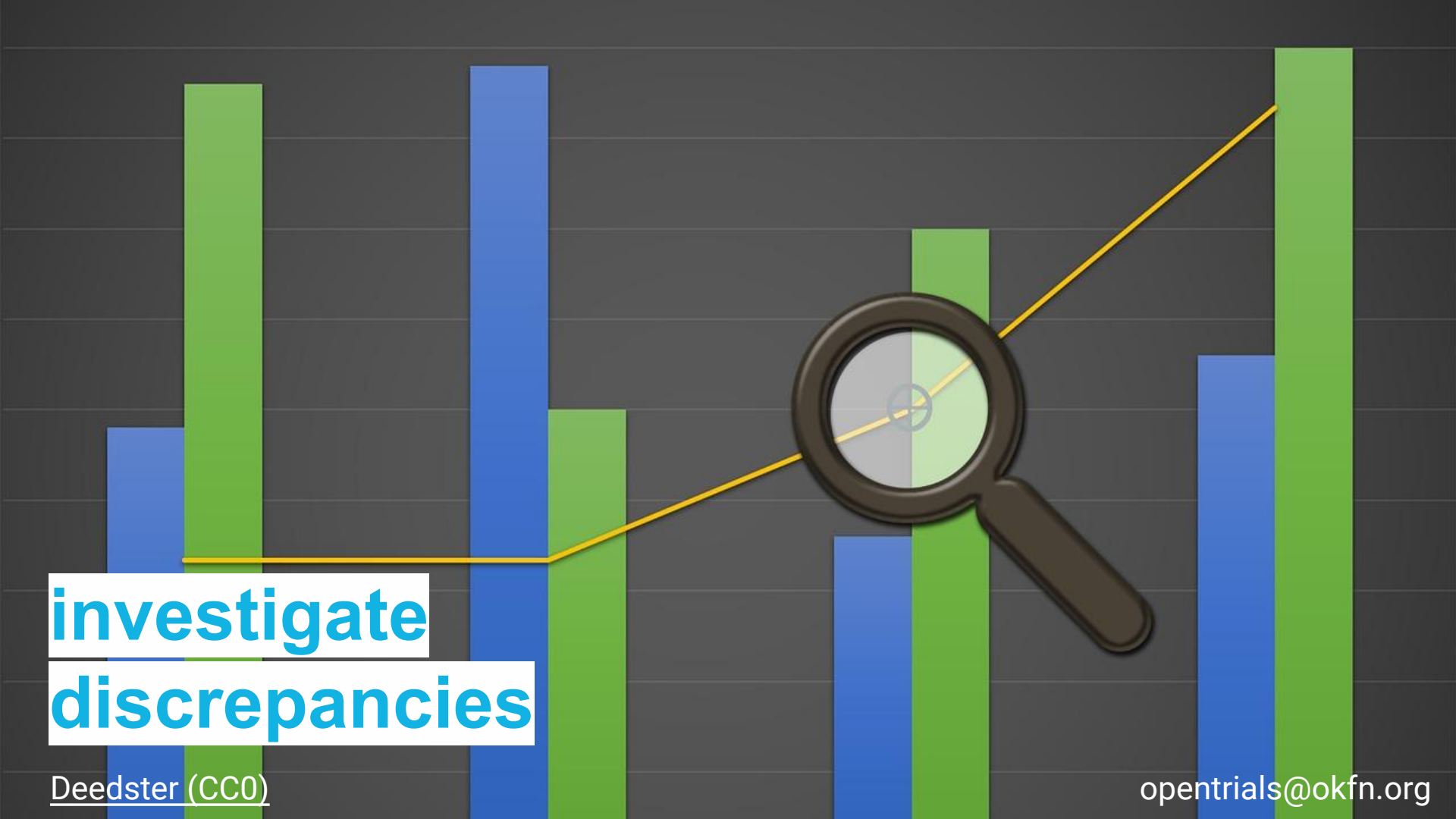
Since Jan 2006, **all major trial sponsors** completed 25,927 eligibl
haven't published results for 11,714 trials. That means 45.2% of
missing results.



opentrials@okfn.org

search functionality

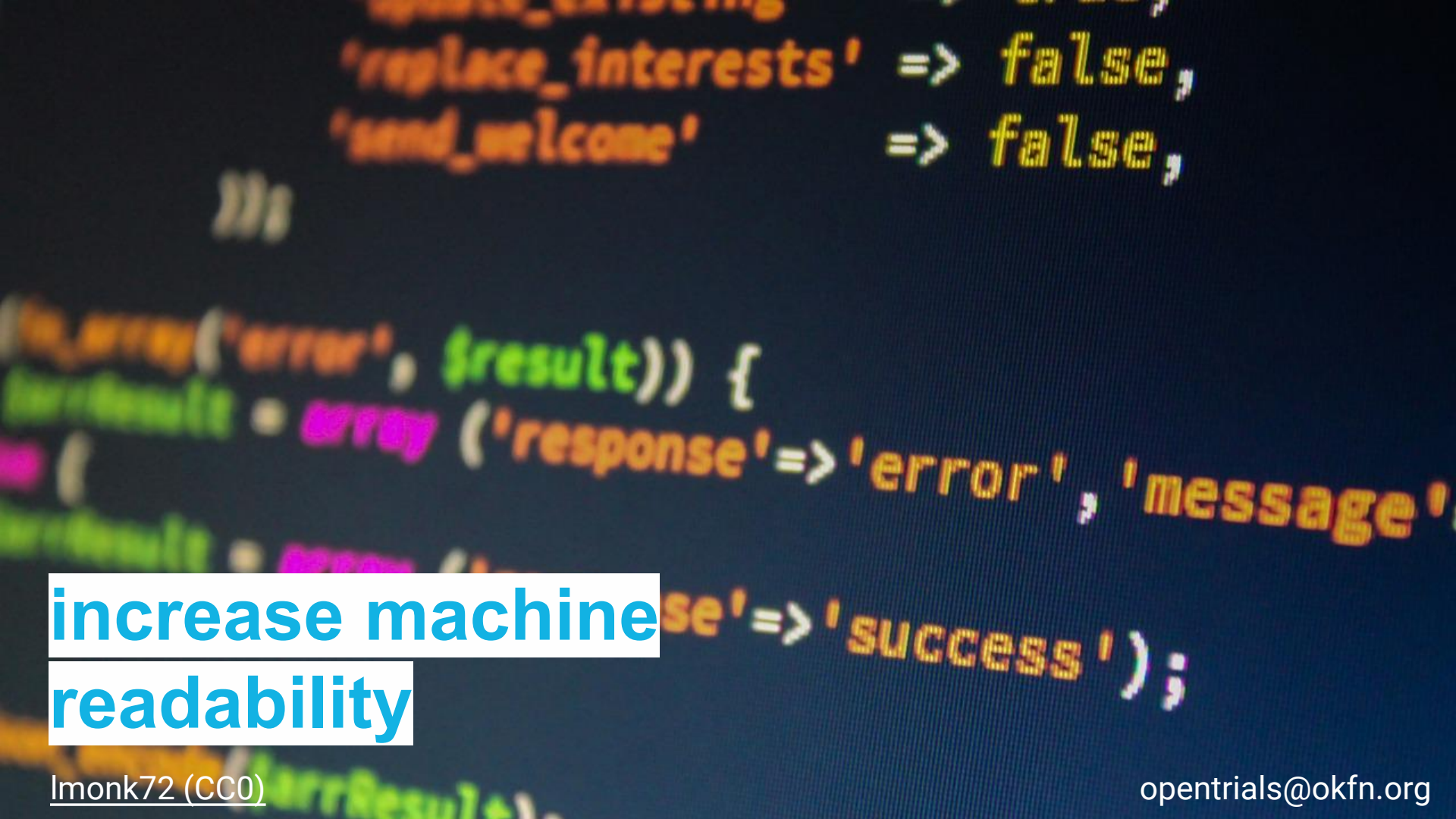
opentrials@okfn.org



investigate discrepancies

Deedster (CC0)

opentrials@okfn.org



increase machine readability

lmonk72 (CC0)

opentrials@okfn.org

change sharing norms

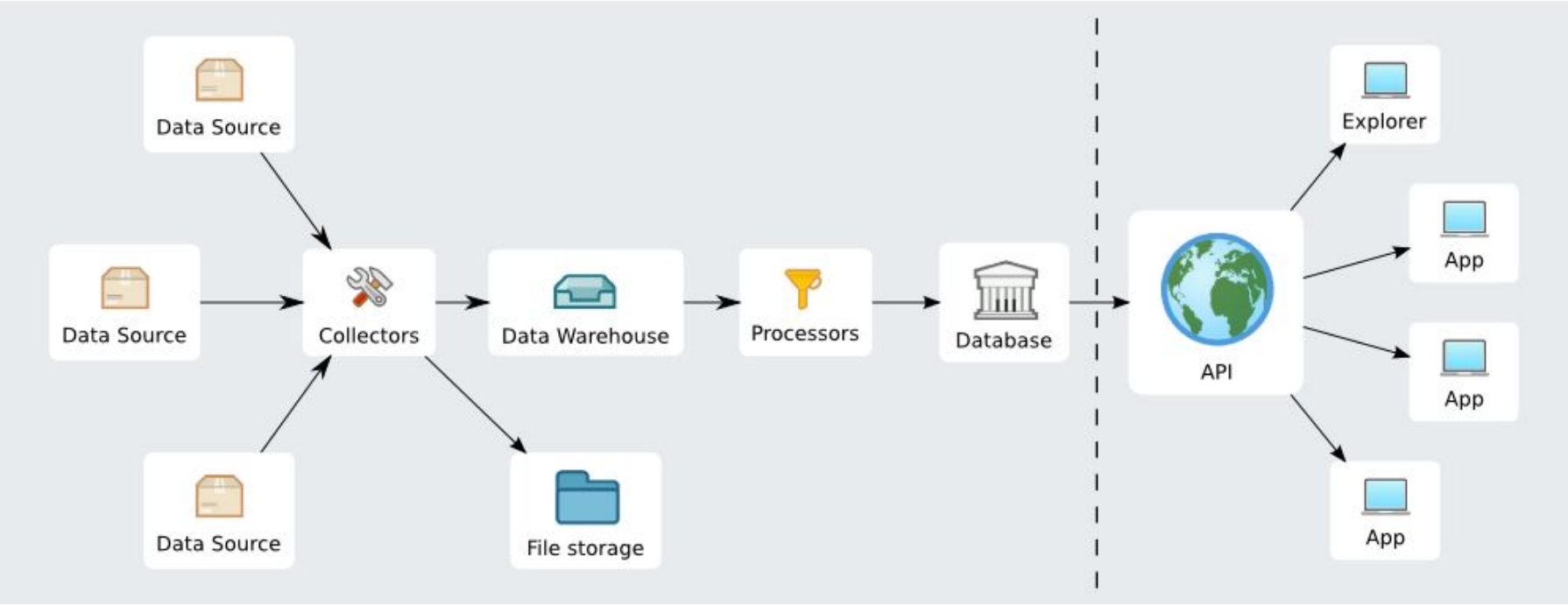
422737 (CC0)

opentrials@okfn.org

1. introduction
2. we (still) have a problem
3. OpenTrials overview
4. techy stuff
5. data sources
6. user examples
7. demo
8. Q&A

opentrials@okfn.org

OpenTrials architecture



opentrials@okfn.org

scraping

data donations

crowdsourcing

opentrials@okfn.org

structured data

HTML

ClinicalTrials.gov Identifier: NCT01013220 First received: November 12, 2009</
Last
December 15, 2014 <
class="info-date">Last verified:
December 2014 <div
class="info-date"> <a
href="/ct2/archive/NCT01013220"
title="Historical versions of study NCT01013220 on ClinicalTrials.gov Archive Site - opens in new window"
onclick="openNewWindow('/ct2/archive/NCT01013220'); return false;">History of Changes

XML

<primary_outcome>
<measure>employer purchase of depression management product</measure>
<time_frame>two years after intervention</time_frame>
<
primary_outc
secondary_ou
<me
management model purchased to evidence-based models
</measure>
<time_frame>two years after intervention</time_frame>
<safety_issue>No</safety_issue>
</secondary_outcome>
<number_of_arms>2</number_of_arms>

API

"id": "a5d82296-8cb8-11e6-988b-0242ac12000c",
"source_id": "nct",
"identifiers
"nct": "NCT02920840"
"p
Synch
Prefro
"gender": "both",
"has_published_results": false,
"status": "ongoing",
"recruitment_status": "recruiting",
"registration_date": "2016-09-26T00:00:00.000Z"

http://api.opentrials.net/v1/search?q=depression&page=1&per_page=20

Database

MySQL

PostgreSQL

Microsoft SQL Server

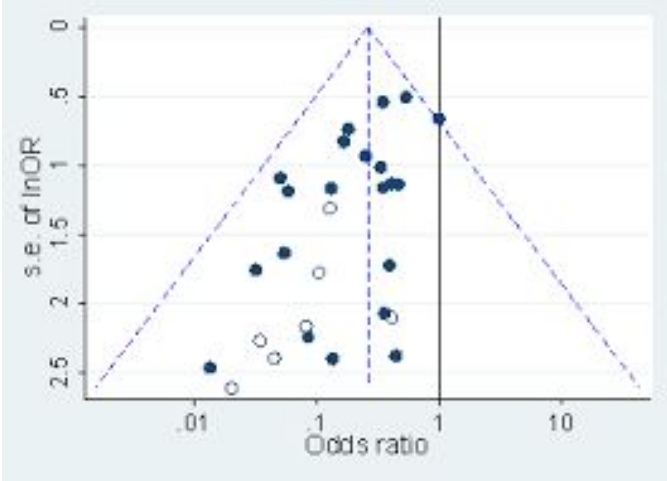
scraping

harder → easier

opentrials@okfn.org

data donations

	A	B	G
1	NCT_ID	OFFICIAL_TITLE	PMID
2	NCT00002762	MENSTRUAL CYCLE	19487378
3	NCT00002879	A PHASE II TRIAL O	18470909
4	NCT00003134	A Phase II Trial of I	19066728
5	NCT00003596	A Phase III Random	18430910
6	NCT00003762	Randomized Phase	19188136
7	NCT00003829	`A Phase II Study o	
8	NCT00003849	A Phase II Trial of O	15657404
9	NCT00003996	A Phase II Trial of P	
10	NCT00005601	A Phase II Trial of O	18569634
11	NCT00005829	Phase II Study of G	
12	NCT00005963	Phase II Trial of Do	16118507
13	NCT00006007	A Phase II Study of	16303865
14	NCT00006010	Phase II Trial of Ge	21555932
15	NCT00006305	Bypass Angioplasty	19502645
16	NCT00009893	Phase II Trial Of Ge	15558814
17	NCT00017186	Phase II Study of G	18224661
18	NCT00022139	A Phase II Trial of P	17921712
19	NCT00022646	A Phase II Clinical T	16135464
20	NCT00025025	Colorectal Cancer S	19026650
21	NCT00026403	A Phase II Study of	17577035
22	NCT00027612	Pilot And Phase II T	20063115
23	NCT00027963	The Efficacy Of Gab	17853395
24	NCT00028925	Phase II Trial of Ora	19935387
25	NCT00032032	Phase I/II Study Of	16730134
26	NCT00040859	A Phase II Study of	16303863
27	NCT00040885	Docetaxel And Infli	19665818
28	NCT00043069	Osteoporosis Preve	19468757
29	NCT00045162	Randomized Phase	16648503
30	NCT00045201	Phase I Trial of OSI	
31	NCT00049673	A Randomized Pha	23297129
32	NCT00053027	Phase II Trial of Rit	18470909
33	NCT00054418	A Phase III Random	19075260
34	NCT00054457	A Phase II Study Of	16497828




```
<QUALITY_ITEM CORE_ITEM="YES" ID=
<NAME>Random sequence generation
<DESCRIPTION>
<P>Selection bias (biased allocat
</DESCRIPTION>
<QUALITY_ITEM_DATA/>
</QUALITY_ITEM>
<QUALITY_ITEM CORE_ITEM="YES" ID=
<NAME>Allocation concealment (sel
<DESCRIPTION>
<P>Selection bias (biased allocat
</DESCRIPTION>
<QUALITY_ITEM_DATA/>
</QUALITY_ITEM>
<QUALITY_ITEM CORE_ITEM="YES" ID=
<NAME>Blinding (performance bias
<DESCRIPTION>
<P>Performance bias or detection
</DESCRIPTION>
<QUALITY_ITEM_DATA/>
</QUALITY_ITEM>
<QUALITY_ITEM CORE_ITEM="NO" ID="
<NAME>Blinding of participants an
<DESCRIPTION>
<P>Performance bias due to knowle
</DESCRIPTION>
<QUALITY_ITEM_DATA/>
</QUALITY_ITEM>
<QUALITY_ITEM CORE_ITEM="NO" ID="
<NAME>Blinding of outcome assessm
<DESCRIPTION>
<P>Detection bias due to knowledg
</DESCRIPTION>
<QUA
</QU
```

opentrials@okfn.org

OpenTrials
All the Data, on All the Trials, Linked

Explore Contribute Data About Statistics Ben Meghreblian

OPEN KNOWLEDGE INTERNATIONAL

Which category best describes your contribution?
Analytics code
Case report forms
Clinical study report
Clinical study report synopsis
Condition
Consent forms
Critical review
European Public Assessment Report (EPAR) document
Intervention
Lay summaries
Other
Patient information sheet
Publication (press release, article, blog post, etc.)
Registry entry
Statistical analysis plan
Trial protocols
Trial webpage
U.S. Food and Drug Administration (FDA) document

, you're agreeing with our **Terms of Use**. The contributions will be moderated before being displayed

crowdsourcing

sting link, like a trial's webpage (e.g. http://www.somewhere.com/data.pdf)

Which category best describes your contribution?

Comments

Tell us about what you're sending

Submit

opentrials@okfn.org

attribution and licensing

Open Definition | The Definition | Conformant Licenses | Participate | News

The Open Definition

The **Open Definition** sets out principles that define "openness" in relation to **data and content**.

It makes **precise** the meaning of "open" in the terms **"open data"** and **"open content"** and thereby ensures **quality** and encourages **compatibility** between different pools of open material.

It can be summed up in the statement that:

> *"Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)."*

Put most succinctly:

> *"Open data and content can be freely used, modified, and shared by anyone for any purpose"*

Read the full Open Definition »

THE OPEN DEFINITION IN YOUR LANGUAGE

العربية | Беларуская | Български | Català | Czech | Dansk | Deutsch | Ελληνικά | English | Español | Euskara | Suomi | Français | Galego | עברית | हिन्दी | Croatian | Magyar | Indonesian | Íslenska | Italiano | 日本語 | ಕನ್ನಡ | 한국어 | македонски јазик | Norsk (bokmål) | Polszczyzna | Português Brasileiro | Português | Русский | Shqip | Srpski | Svenska | తెలుగు | Türkçe | Українська | 简化中国 | 中文

Fir0002 (CC BY-SA)

opentrials@okfn.org

1. introduction
2. we (still) have a problem
3. OpenTrials overview
4. techy stuff
5. data sources
6. user examples
7. demo
8. Q&A

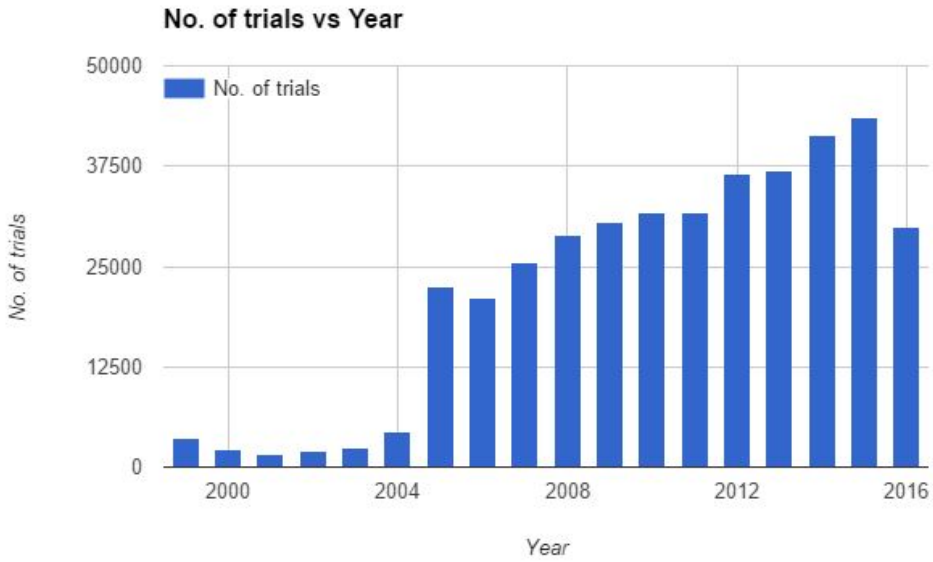
opentrials@okfn.org

imported: **trial registers**

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

European Clinical Trials Database
EudraCT

International Clinical Trials Registry Platform
Search Portal



337,911 deduplicated trials

opentrials@okfn.org

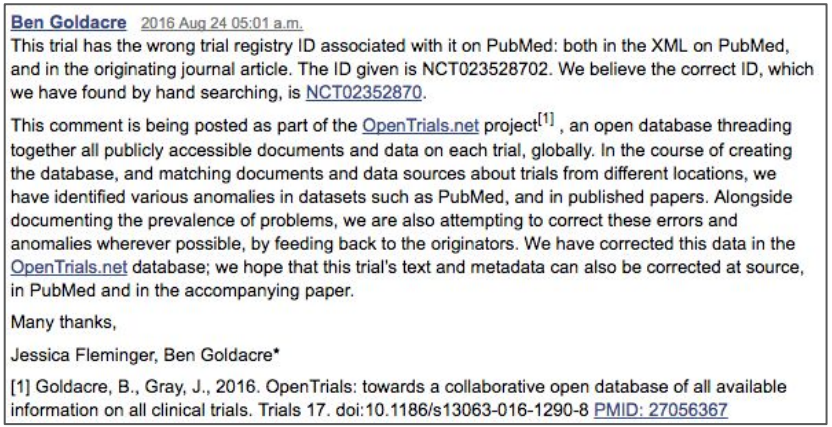
imported: **publications**

~510,000 publications

(~24,000 linked with trials)

errors found in some registry IDs

⇨ feedback via PubMed Commons



Ben Goldacre 2016 Aug 24 05:01 a.m.

This trial has the wrong trial registry ID associated with it on PubMed: both in the XML on PubMed, and in the originating journal article. The ID given is NCT023528702. We believe the correct ID, which we have found by hand searching, is NCT02352870.

This comment is being posted as part of the OpenTrials.net project[1] , an open database threading together all publicly accessible documents and data on each trial, globally. In the course of creating the database, and matching documents and data sources about trials from different locations, we have identified various anomalies in datasets such as PubMed, and in published papers. Alongside documenting the prevalence of problems, we are also attempting to correct these errors and anomalies wherever possible, by feeding back to the originators. We have corrected this data in the OpenTrials.net database; we hope that this trial's text and metadata can also be corrected at source, in PubMed and in the accompanying paper.

Many thanks,

Jessica Fleminger, Ben Goldacre*

[1] Goldacre, B., Gray, J., 2016. OpenTrials: towards a collaborative open database of all available information on all clinical trials. Trials 17. doi:10.1186/s13063-016-1290-8 PMID: 27056367

imported: **risk of bias scores**

```
<QUALITY_ITEM CORE_ITEM="YES" ID="QIT-01" LEVEL="STUDY" NO="1">
<NAME>Random sequence generation (selection bias)</NAME>
<DESCRIPTION>
<P>Selection bias (biased allocation to interventions) due to inad
</DESCRIPTION>
<QUALITY_ITEM_DATA/>
</QUALITY_ITEM>
<QUALITY_ITEM CORE_ITEM="YES" ID="QIT-02" LEVEL="STUDY" NO="2">
<NAME>Allocation concealment (selection bias)</NAME>
<DESCRIPTION>
<P>Selection bias (biased allocation to interventions) due to inad
</DESCRIPTION>
<QUALITY_ITEM_DATA/>
</QUALITY_ITEM>
<QUALITY_ITEM CORE_ITEM="YES" ID="QIT-03" LEVEL="GROUP" NO="3">
<NAME>Blinding (performance bias and detection bias)</NAME>
<DESCRIPTION>
<P>Performance bias or detection bias due to knowledge of the allo
</DESCRIPTION>
<QUALITY_ITEM_DATA/>
</QUALITY_ITEM>
<QUALITY_ITEM CORE_ITEM="NO" ID="QIT-07" LEVEL="GROUP" NO="7">
<NAME>Blinding of participants and personnel (performance bias)</N
<DESCRIPTION>
<P>Performance bias due to knowledge of the allocated intervention
</DESCRIPTION>
<QUALITY_ITEM_DATA/>
</QUALITY_ITEM>
<QUALITY_ITEM CORE_ITEM="NO" ID="QIT-08" LEVEL="GROUP" NO="8">
<NAME>Blinding of outcome assessment (detection bias)</NAME>
<DESCRIPTION>
<P>Detection bias due to knowledge of the allocated interventions
</DESCRIPTION>
<QUALITY_ITEM_DATA/>
</QUALITY_ITEM>
```

Methodological rigour

The methodological rigour of this trial was scored in a **systematic review**, where it was scored as being at "low risk" of bias for "sequence generation", "unclear" for "allocation concealment", "unclear" for "attrition", "high risk" of bias for "reporting", "low risk" of bias for "other biases", "unclear" for "blinding (performance)", and "low risk" of bias for "blinding (detection)".

opentrials@okfn.org

NHS
Health Research Authority

imported: **research summaries**

~22,000 research summaries

provides ‘lay summary’ explanation of the trial for a non-medical audience

created for ethics committee

FOIA request ⇨ HRA for Patient Information Sheets

This study is being carried out to see if the drug MPDL3280A can reduce the size of tumours in patients with bladder cancer before surgery. MPDL3280A is currently being investigated in a number of tumour types and has been shown to have activity in bladder cancer which has spread beyond the bladder.

MPDL3280A is designed to stop a protein called PD-L1 (programmed death-ligand 1) being expressed on the cancer. PD-L1 helps to camouflage the cancer, preventing the body’s immune system from identifying the cancer and fighting it. MDPL3280A works against PD-L1, allowing the immune system to recognise the tumour cells as foreign bodies and attack them.

There are strict inclusion and exclusion criteria for this trial. Broadly speaking, patients with histologically confirmed transitional cell carcinoma of the bladder (T2-T4a – this indicates how far into the bladder the cancer cells have grown) are eligible.

If a patient is eligible for the study and decides to take part, they will be enrolled into the study and will receive up to two 3-weekly cycles of MPDL3280A. 4-8 weeks after being enrolled, the patient will have an operation to remove the bladder (cystectomy) as per normal practise. Treatment with MPDL3280A in the window between enrolment and surgery will not delay this surgery. Following the operation, they will attend three hospital visits (4, 12 and 24 weeks after cystectomy) and their disease progress/survival will be followed over the next 2 years. The clinical team will compare the patient’s tumour tissue samples, scan results and blood results from before and after treatment with MPDL3280A in order to see how well the drug works and if it is safe. Many of the procedures involved in this study are offered as standard care and participation in this trial will not delay surgery.

The study is being carried out in England and will also take place in 5 other European countries (France, Germany, Italy, the Netherlands, Spain).

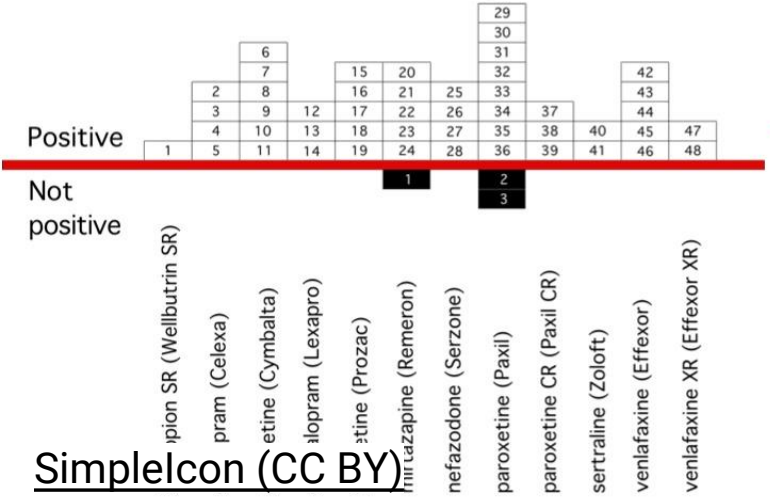
OpenTrialsFDA

live prototype: **FDA Drug Approval Packages**

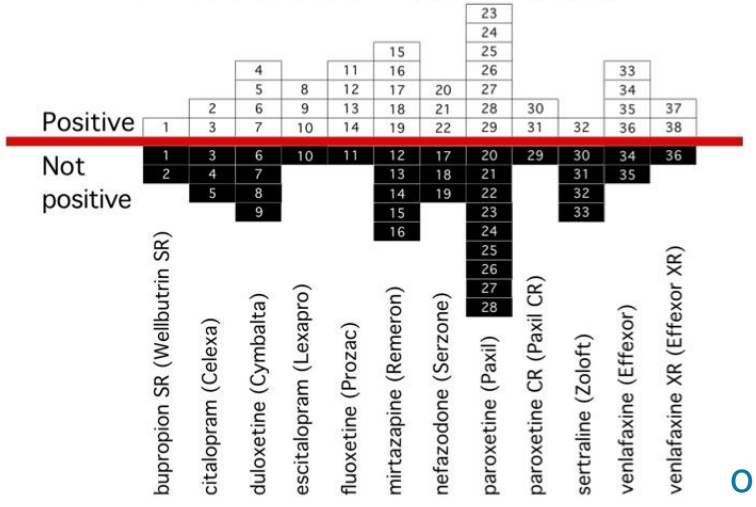
46,332 documents imported to Document Cloud and OCRed

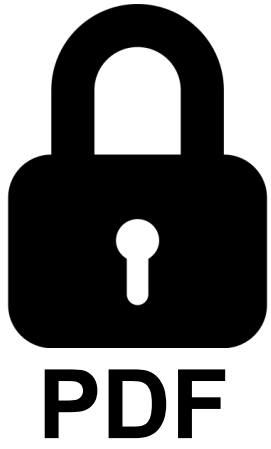
16,329 FDA applications

Journal version of antidepressant trials



FDA version – same trials





SimpleIcon (CC BY)

opentrials@okfn.org

ongoing: **clinical study reports**

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

the YODA PROJECT

abbvie

SYNOPSIS

Name of Sponsor/Company	Janssen Research & Development*
Name of Finished Product	To be determined
Name of Active Ingredients	TMC435 (simeprevir)

* Janssen Research & Development is a global organization that operates through different legal entities in various countries. Therefore, the legal entity acting as the sponsor for Janssen Research & Development studies may vary, such as, but not limited to Janssen Biotech, Inc.; Janssen Products, LP; Janssen Biologics, BV; Janssen-Cilag International NV; Janssen, Inc; Janssen Infectious Diseases BVBA (formerly known as Tibotec BVBA); Janssen R&D Ireland (formerly known as Tibotec Pharmaceuticals); or Janssen Research & Development, LLC (including the former Tibotec Inc. entity). The term "sponsor" is used to represent these various legal entities as identified on the sponsor list.

Status: Approved
Date: 4 October 2013
Prepared by: Janssen Infectious Diseases - Diagnostics BVBA

Protocol No.: TMC435-TiDP16-C216

Title of Study: A Phase 3, randomized, double-blind, placebo-controlled study to investigate the efficacy, safety and tolerability of TMC435 versus placebo as part of a treatment regimen including peginterferon α-2a (Pegasys®) and ribavirin (Copegus®) or peginterferon α-2b (PegIntron®) and ribavirin (Rebetol®) in treatment-naïve, genotype 1, hepatitis C-infected subjects

Study Name: TMC435-TiDP16-C216 (QUEST-2)

EudraCT Number: 2010-021174-11

NCT No.: NCT01290679

Clinical Registry No.: CR017380

Coordinating Investigator: [redacted] MD, [redacted] Germany

Study Centers: The study was conducted at 76 sites in 14 countries.

Publication (Reference):

Manns M, Marcellin P, Poordad Fred FP, et al. Simeprevir (TMC435) with peginterferon/ribavirin for treatment of chronic HCV genotype-1 infection in treatment-naïve patients: Results from QUEST-2, a Phase 3 trial; Poster at The International Liver Congress 2013, April 24 - 28 2013, Amsterdam, The Netherlands; Journal of Hepatology 2013 Suppl 1(58) S568.

Study Period: 18 January 2011 to 5 February 2013

Phase of Development: Phase 3

Objectives: The primary objective was to demonstrate the superiority of TMC435 versus placebo as part of a treatment regimen including pegylated interferon alpha-2a (PegIFNα-2a)/ribavirin (RBV) or PegIFNα-2b/RBV, with respect to the proportion of treatment-naïve hepatitis C virus (HCV) genotype 1 infected subjects with sustained virologic response (SVR) 12 weeks after the planned end of treatment (SVR12).

opentrials@okfn.org

ongoing: **more...**

Epistemonikos

even **asking** has been interesting...

opentrials@okfn.org

1. introduction
2. we have a problem
3. OpenTrials overview
4. techy stuff
5. data sources
6. user examples
7. demo
8. Q&A

opentrials@okfn.org

1. researcher
2. doctor
3. patient
4. data journalist
5. developer
6. policy maker
7. regulator

Rhoda Baer (CC0)

opentrials@okfn.org

1. researcher
2. doctor
3. patient
4. data journalist
5. developer
6. policy maker
7. regulator

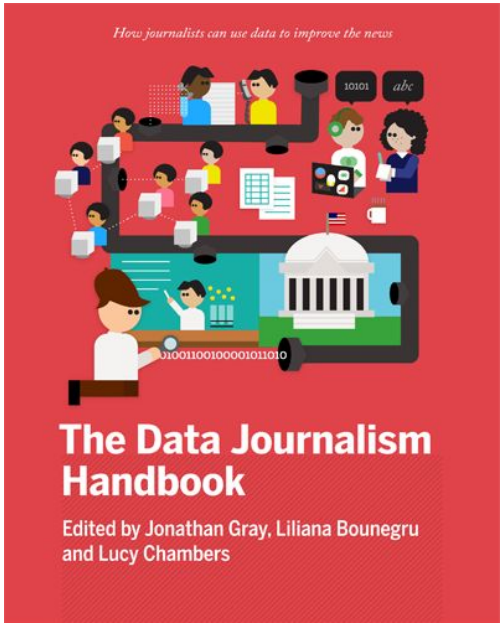
Ilmicrofono Oggiono (CC BY)

opentrials@okfn.org

1. researcher
2. doctor
3. patient
4. data journalist
5. developer
6. policy maker
7. regulator

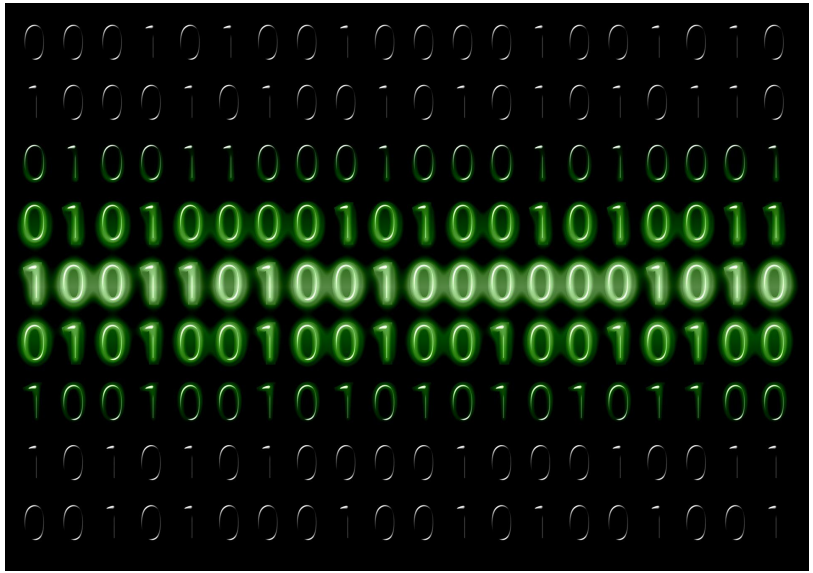
Michael Coghlan (CC BY-SA)

1. researcher
2. doctor
3. patient
4. data journalist
5. developer
6. policy maker
7. regulator



The Data Journalism Handbook (CC BY-SA)

1. researcher
2. doctor
3. patient
4. data journalist
5. developer
6. policy maker
7. regulator



geralt (CC0)

opentrials@okfn.org

1. researcher
2. doctor
3. patient
4. data journalist
5. developer
6. policy maker
7. regulator

olibac (CC BY)

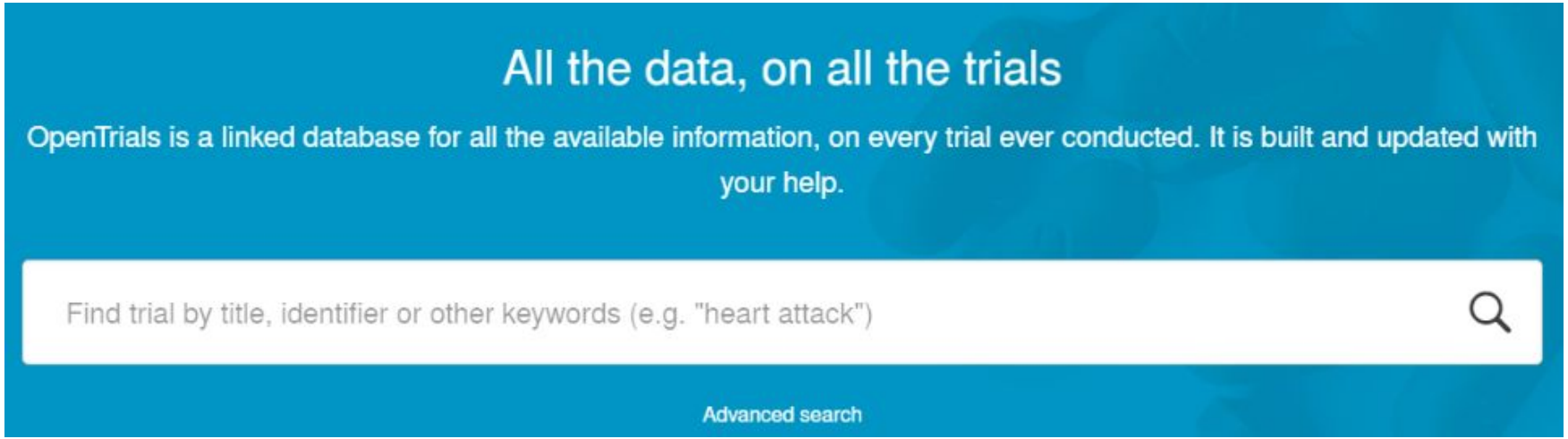
1. researcher
2. doctor
3. patient
4. data journalist
5. developer
6. policy maker
7. regulator

opentrials@okfn.org

1. introduction
2. we have a problem
3. OpenTrials overview
4. techy stuff
5. data sources
6. user examples
7. demo
8. Q&A

opentrials@okfn.org

explorer.opentrials.net



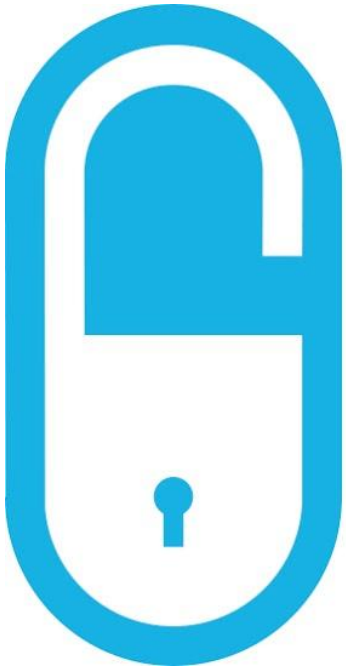
beta version!

opentrials@okfn.org

please

contribute!

opentrials@okfn.org



thank you!
any questions?

opentrials@okfn.org / opentrials.net / @opentrials

Team OpenTrials: Ben Goldacre, Emma Beer, Stephen Abbott Pugh, Paul Walsh, Vitor Baptista, Evgeny Karev, Victor Nițu, Georgiana Bere, Sierra Williams, Lieke Ploeger, Sam Smith, Ben Meghreblian, Sarah McNeill, Jessica Fleminger