

Being open but not naked: Balancing transparency with resilience in science

Stephan Lewandowsky

University of Bristol

and

University of Western Australia



@STWorg



Australian Government
Australian Research Council

stephan.lewandowsky@bristol.ac.uk

THE
ROYAL
SOCIETY

Transparency and Open Data are Essential



Transparency and Open Data are Essential

- But we need a conversation about the full implications
 - political context
 - intentions of people who request data
 - consent and ethics
 - competence of people who request data
 - communication platforms
 - cherry-picking

I. Open Data is Political

- U.S. data access act 1998 (and 2000)
 - all data from federally funded projects available
 - agency must obtain data from grantees
 - data disseminated by government must adhere to act's definition of quality
 - citizens can challenge and agencies must respond
 - “influential data” reanalysis by “open source”

- The acts were drafted by the tobacco industry
- Implementation was overseen by tobacco industry

(Baba et al., 2005, *American Journal of Public Health*)

“A delicate diplomatic situation”: tobacco industry efforts to gain control of the Framingham Study

Janine K. Cataldo^{a,*}, Lisa A. Bero^b, Ruth E. Malone^c

^a*Department of Physiological Nursing—Gerontology, University of California San Francisco, San Francisco, CA, USA*

^b*Department of Clinical Pharmacy and Health Policy, University of California San Francisco, San Francisco, CA, USA*

^c*Department of Social and Behavioral Sciences, University of California San Francisco, San Francisco, CA, USA*

Accepted 30 January 2010

- Tobacco industry used access to raw data for re-analysis by industry consultant
- Tobacco-related morbidity disappeared
 - Lead (paint, gasoline)
 - PVC
 - Many examples ...

Open Data may Never be Open Enough

- Lamar Smith (R-TX), Chair of the House of Representatives Science Committee
- Issued subpoena in 2016 to National Oceanic and Atmospheric Administration (NOAA) ...
- ... for climate data that were publically available at the time

Open Data or Open Season?

- Energy & Environment Legal Institute (“E&E”)
 - mission is “free market environmentalism”
 - in its own words, it “pepper[s] universities around the country” with open records requests
- E&E interested in science?
 - no requests for reports
 - no requests for data
 - no requests for methodologies
 - no requests for funding source
 - no requests for funding amounts
- Only request for e-mails between climate researchers

Secret Science Reform Act

- Now before Congress, rebranded as the Honest and Open New EPA Science Treatment Act (HONEST)
- Would limit EPA to using only data that *can be replicated* and publically available for independent analysis.
- *All* the underlying information must be public

(Un-)intended Consequences of HONEST for EPA

- Replicability requirement rules out data from Deepwater Horizon or Exxon Valdez spill.
- Data ruled out if a business claimed that *any* of the information EPA used to regulate was “confidential business information”.
- Studies about the health impacts of pollution ruled out if they relied *in any part* on confidential patient health data.
- Decisions about safety of food and drugs made if they rely on info from industry “trade secrets.”

EPA estimated to rely on
50,000 articles annually

II. Do Intentions Matter?

- Open Data advocates: No
- Tobacco industry: No, why?
- Public health researchers: Absolutely yes!

Nefarious Intentions and Science

- Dr. Smith publishes a paper suggesting interference causes forgetting from memory.
- Dr. Jones requests data from Dr. Smith and discovers that it is actually leprechauns that caused the forgetting.
- Dr. Jones is Dr. Smith's ex-spouse and recently lost a custody battle over a hamster.
- Ultimately the scientific community ignores the issue.

- Egos are bruised
- Careers hampered
- But no (not much?) lasting harm done

Nefarious Intentions and Public Policy

- Dr. Smith publishes a paper suggesting that Product X causes cancer.
- Dr. Jones, who works for Manufacturer X, requests data and purports to discover that cancer is caused by “dispositional factors” rather than Product X.
- Ultimately the scientific community resolves the issue.

**But massive
harm is done**

Scientific Debate \neq Public Debate

nature
climate change

ARTICLES

PUBLISHED ONLINE: 28 OCTOBER 2012 | DOI:10.1038/NCLIMATE1720

The pivotal role of perceived scientific consensus in acceptance of science

Stephan Lewandowsky^{*}, Gilles E. Gignac and Samuel Vaughan

- The appearance of a scientific debate, whether real or not, prolongs public indecisiveness.
- Tobacco control legislation was delayed by decades due to *appearance* of a scientific debate.

II. Do Intentions Matter?

- Open Data advocates: No
- Tobacco industry: No, why?
- Public health researchers: Absolutely yes!



Now you know why
they say that

III. Consent and Ethics

- Medical or clinical research
 - patient confidentiality
 - anonymization can be difficult
 - even de-linking is insufficient unless the linking key has been destroyed or is held by another institution (U.K. data protection act)
 - consent may have been given for one stated purpose of a study only

Consent: Exploring Implications

- Ms. Jones consents to participate in an experiment that observes the effect of WM training on IQ
 - The experimenter collects the usual covariates and demographics such as gender and ethnicity
 - The Ku-Klux-Klan reads the results and discovers that blacks have higher IQ overall but that training is ineffective
- The *Kluxer's Trumpet* headline: "No matter how much you train them, they cannot get smarter"
 - FOX news picks up the story

Consent: Exploring Implications

- Ms. Jones consents to participate in an experiment that observes the effect of WM training on IQ
- Suppose Ms. Jones was black
 - did she realize what she consented to?
 - would she have given consent if she knew this could happen?
 - given what happened, would she consent again?
 - note that anonymity is not an option

Researchers lose control over their data when they are completely open

Concerns are Being Articulated

Sydes et al. *Trials* (2015) 16:104
DOI 10.1186/s13063-015-0604-6



METHODOLOGY

Open Access

Sharing data from clinical trials: the rationale for a controlled access approach

Matthew R Sydes^{*}, Anthony L Johnson, Sarah K Meredith, Mary Rauchenberger, Annabelle South and Mahesh KB Parmar

Abstract

Background: The move towards increased transparency around clinical trials is welcome. Much focus has been on under-reporting of trials and access to individual patient data to allow independent verification of findings. There are many other good reasons for data sharing from clinical trials. We describe some key issues in data sharing, including the challenges of open access to data. These include issues in consent and disclosure; risks in identification, including

IV. Does Competence Matter?

- Researchers operate in an institutional context
 - ethics boards
 - data management plans
 - preregistration
 - peer-reviewed literature
- Mr. Tom D. Harry from Center for Transparency
 - none of the above
 - but he has a blog!

Tom D Harry shocker:
Vaccinations kill!!!!
Truth revealed by re-analysis

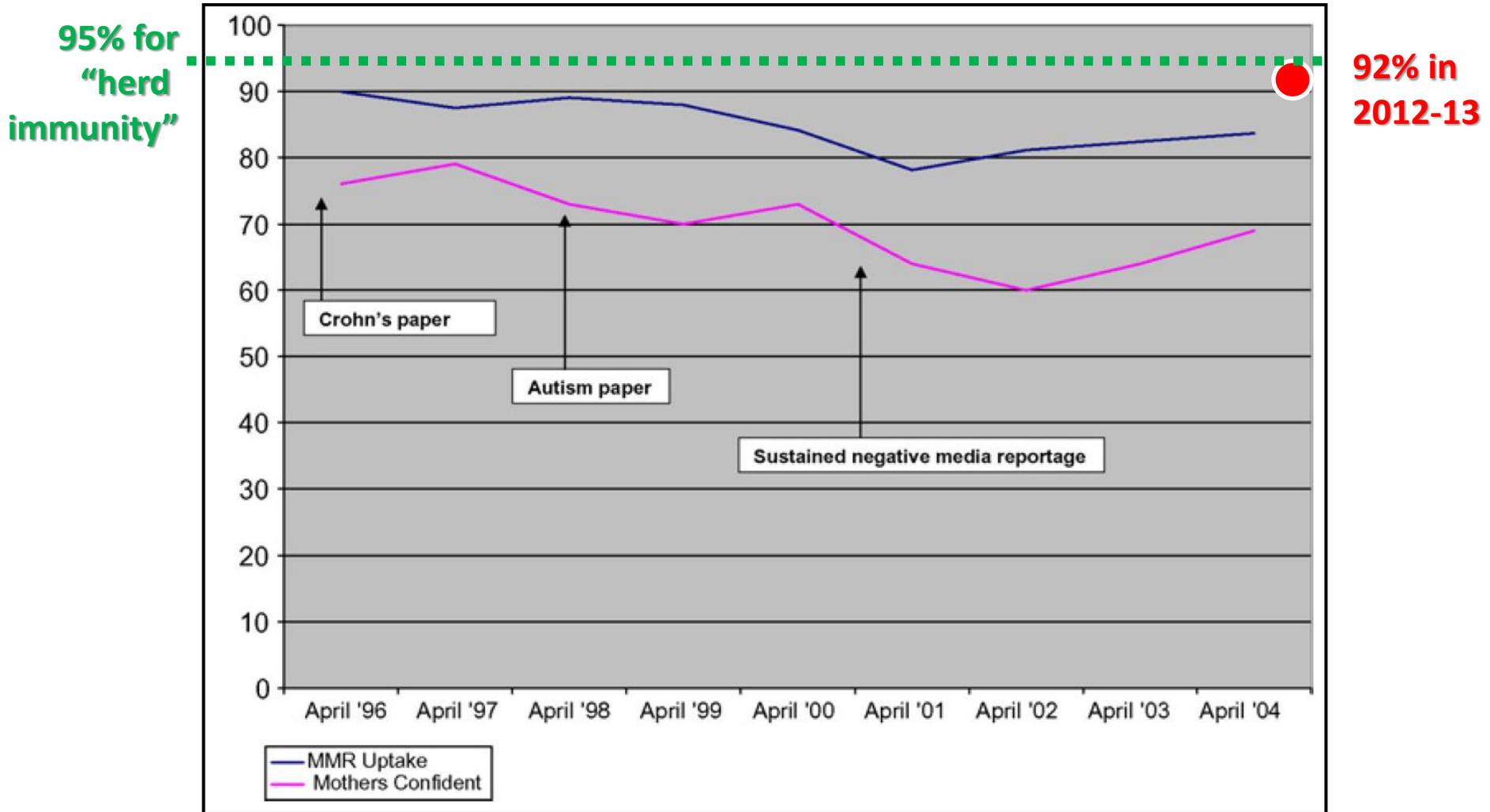
Importance of Competence

- U.K. Medical Research Council's guidelines:
“The custodian [of the data] must ensure that the group [receiving the data] accepts a duty of confidence and protects confidentiality through training procedures, etc, **to the same standards as the custodian.**”

Consideration of the Platforms

- Dr. Smith **publishes a paper** suggesting that Product X causes cancer.
- Dr. Jones, who works for Manufacturer X, requests data and purports to discover that cancer is caused by “dispositional factors” rather than Product X.
- **Dr. Jones and Manufacturer X publish analysis on blogs and Twitter. FOX News picks it up.**
- Ultimately the scientific community resolves the issue.

U.K. MMR Vaccination Rates (Smith et al., 2007)



V. Cherry-Picking

- We urge scientists to preregister their hypotheses and analysis plans to guard against cherry-picking of results or outcome measures.
- We do not keep track of the Ku-Klux-Klan requesting 90 data sets with a racial-identifier variable
- So they can trumpet the one result that yields the “desired” racial differences

What Does this Add up to?

- Science should be open and transparent
- But there is a distinction between science on the one hand, and noise, nonsense, commercial interests, or political propaganda on the other
- Openness and transparency facilitate science, but they disproportionately also aid in the dissemination of noise, nonsense, commercial interests, and political propaganda

Solutions? Symmetry

- People who request data must be competent and must operate in an institutional context of accountability
- People who request data must preregister their intentions (and conform to them)
- Participants' consent must be considered
- Data availability (and limits) should be enshrined in peer-review record at the time of publication

Thank you

