

PhUSE Annual Conference

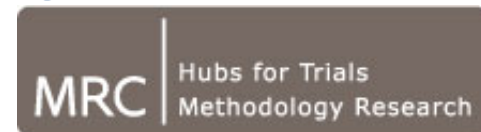
London, 12th-15th October

Data Transparency

Understanding that Data Transparency Matters: The Academic View

Sarah J. Nolan and Dr. Catrin Tudur Smith
Department of Biostatistics
University of Liverpool

(Acknowledgement: Professor Anthony Marson)



DISCLAIMER



views expressed are our own personal views and do not necessarily reflect other academic researchers

Understanding that data transparency matters: An academic view

- Activities of an academic researcher
- Our experiences of data requesting
 - Past and present
- The academic view of ‘transparency’
 - Why transparency is important
- *Data transparency in academic trials*

Activities of an academic researcher

- October 2011: 3 year research project for Cochrane Epilepsy Group funded by National Institute of Health Research (NIHR):

“Clinical and cost effectiveness of interventions for epilepsy in the NHS”

- 7 Cochrane Review Updates
 - Randomised controlled trials of monotherapy
 - Individual Participant Data Meta-Analyses
- Cochrane Network Meta Analysis (10 drugs)

Data requesting: the past

Original Cochrane reviews (approx. 1995-2005)

- 29 data requests made (5881 participants)
 - 12 Academic studies (1383 ppts)
 - 13 Pharmaceutical studies (3320 ppts)
 - 4 Government studies (1178 ppts)
- Data provided from 18 studies (4697 ppts, 80%)
 - 2 Academic studies (286 ppts, 21%)
 - 13 Pharmaceutical studies (3320 ppts, 100%)
 - 3 Government studies (1091 ppts, 93%)

Data for 80% of all eligible ppts (50 - 100% per review)

Data requesting: the past

Original Cochrane reviews (approx. 1995-2005)

Data for 80% of all eligible participants

- Correspondence in '**BE days**' (before e-mail!)
 - Contacts in the field – very important
 - Fax, letter, telephone, face-to-face meetings
 - Data provided on floppy disks, paper copies
- Remaining 20% of data from 11 studies
 - Often studies conducted before 1990
 - Data lost / destroyed / no longer available
 - Data of interest not recorded

Data requesting: the past and present

Ongoing work (SJN): Jan 2012 – present

- 41 new studies identified (10679 ppts)
 - 24 Academic studies (4090 ppts)*
 - 16 Pharmaceutical studies (5996 ppts)
 - 1 Government study (593 ppts)

**2 studies conducted at our site (2437 ppts)*

- First Step: Contact authors / sponsors
- Second Step: Data Requests (8242 ppts)

Data requesting: the past and present

Ongoing work (SJN): Jan 2012 – present

- **First Challenge:** Contact authors / sponsors
 - Who do we contact?
 - Who owns the data?
 - How do we make contact?
 - What do we say?
 - What do we do next?

Data requesting: the past and present

Data request timeline

June 2013 - 1 request successful (academic)

– **168 ppts out of 8242 to request (2%)**

January 2014 – 2 requests successful (academic)

– **299 ppts out of 8242 to request (3%)**

March 2014 – 5 requests successful (4 academic, 1 pharmaceutical)

– **660 ppts out of 8242 to request (8%)**

Data requesting: the past and present

Data request timeline: 2014

Many stalled requests outstanding for > 2 years

- X No response at all

 - X Have we contacted the correct person?

 - X Have we been ignored?

- ✓ Positive response

 - X No further contact

 - X No progress made on data transfer

Projects delayed due to lack of data

Data requesting: the past and present

Data request Timeline: 2014

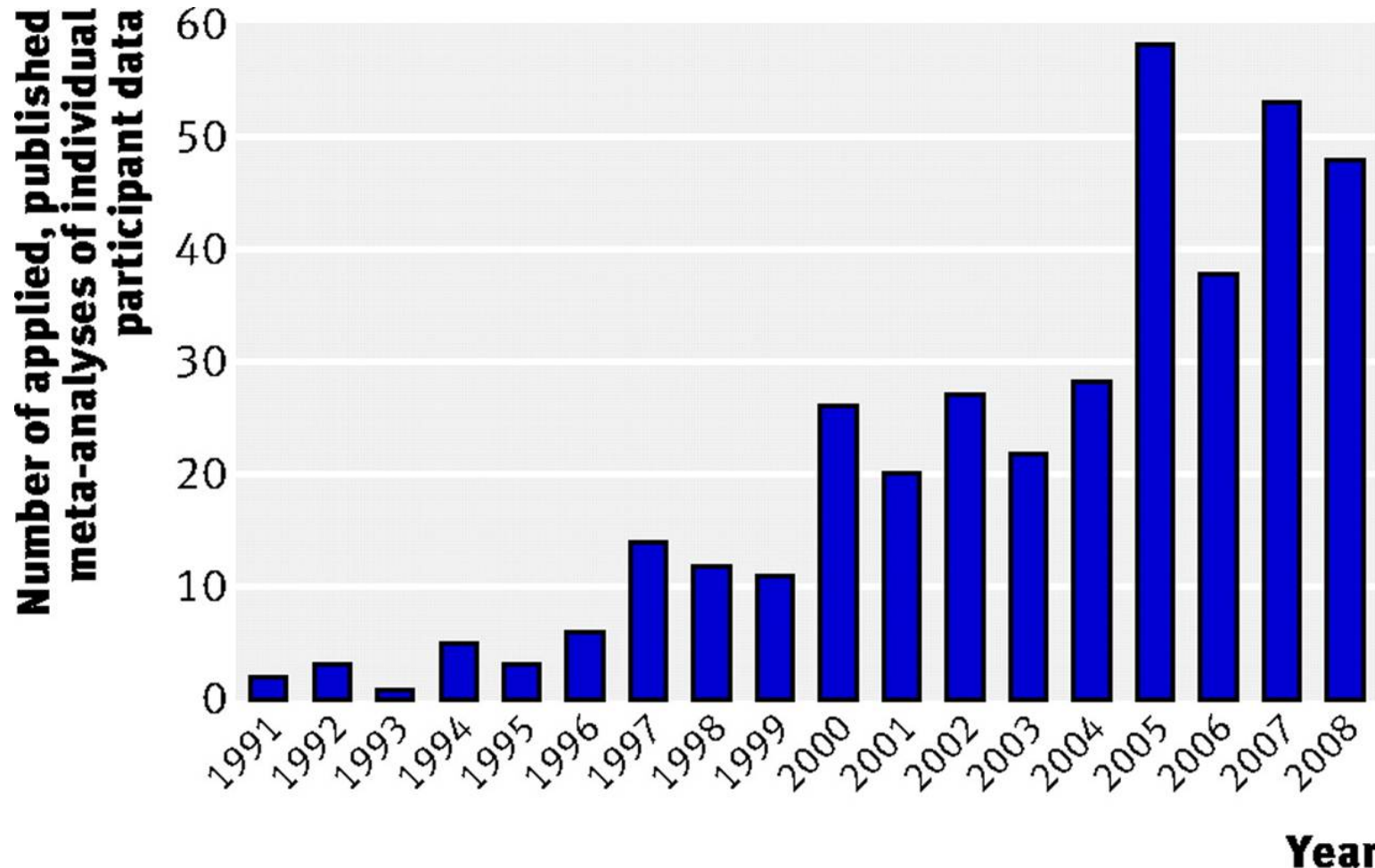
Many stalled requests outstanding for > 2 years

Projects delayed due to lack of data

**PROBLEM: Lack of transparency
in the data requesting process**

Why transparency is important

Number of distinct, applied meta-analyses of individual participant data published up to March 2009



Riley R D et al. **BMJ** 2010;340:bmj.c221 (©2010 by British Medical Journal Publishing Group)

Why transparency is important

Overall completeness and applicability of evidence

“We have gratefully received individual patient data (IPD) for 669 individuals (67% of individuals from all eligible trials) from the authors of five trials”

“However 376 individuals (33%) from four relevant trials could not be included in any analysis as IPD were not available and outcomes of interest were not reported in the published reports.”

“Having to exclude data for a third of eligible participants due to lack of individual patient data and insufficient reporting in study publications is likely to impact on the applicability of the evidence, however it is difficult to quantify exactly how large this impact could be”

Phenytoin versus valproate monotherapy for partial onset seizures and generalised onset tonic-clonic seizures (Review)

Nolan SJ, Marson AG, Pulman J, Tudur Smith C



Data requesting: the present

Data request timeline

03/06/2013: Enquiry submitted to GSK Share (now ClinicalStudyDataRequest.com) – **Request 1**

06/08/2013: Enquiry submitted to GSK Share – **Request 2**

14/11/2013: Research proposal and statistical analysis plan submitted for Requests 1 and 2

10/12/2013: Enquiry submitted to GSK Share – **Request 3**

17/03/2014: Enquiry submitted to Clinicalstudydatarequest.com (GSK study) – **Request 4**

Data requesting: the present

Data request timeline

06/12/2013: Negative response to Request 2

German Law prevents anonymisation of datasets

03/02/2014: Negative response to Request 3

Data and documents could not be located, study 10 years old

19/05/2014: Negative response to Request 4

Costs to provide data are prohibitive, age of study

26/03/2014: Request 1 approved by Independent Review Board

15/05/2014: Data sharing agreement signed for Request 1.

02/06/2014: De-identified data available in the Data Access System

Data requesting: the present

Data request timeline

March 2014 – 5 successful requests, 8% of requested data

September 2014 – 11 requests successful

- 7 academic and 4 pharmaceutical studies
- **2174 participants out of 8242 to request (26%)**
- Positive ongoing requests: 5 studies (3735 ppts)
 - 1 Government and 4 Pharmaceutical studies

Estimated success: 72% of requested data

Estimated success: 78% of eligible data*

*(*plus 2 studies, 2437 participants data owned by us)*

Why transparency is important

What does a researcher want?

- A clear contact for data request enquiries 
- Fully transparent data request process 
 - Who can request data and for what purpose? 
 - What are the stages of the data request process? 
 - Will reasons be given if access is declined? 
 - How long will each stage take? 
- Contact point for subsequent queries 
- Responses to contact in a reasonable time frame 

Why transparency is important

What does a researcher want?

- Data must be accompanied by related documentation 
 - Data labels with descriptions 
 - Annotated CRFs (case report forms) etc. 
- De-identified data should still be **clinically useful** 
 - Sensitive information (age, ethnicity, rare events, death information etc.) may be important to the research question
 - Data Sharing Agreement protects the privacy of research participants

Why transparency is important



What does a researcher want?

Data should be provided in a user friendly, flexible format

- Easily accessible 
- Downloadable? 
- Possible to combine with other data sets 
- But uploading other sponsors data to the research environment may violate data sharing agreements 
- **Restrictions on analyses we can perform**
 - Choice / complexity of software 
 - One stage meta-analysis models? 

