

**STATISTICAL METHODS FOR  
EVIDENCE SYNTHESIS OF INDIVIDUAL  
PARTICIPANT DATA**

A two-day training course,  
7<sup>th</sup> & 8<sup>th</sup> June 2017,  
Keele University



# **STATISTICAL METHODS FOR EVIDENCE SYNTHESIS OF INDIVIDUAL PARTICIPANT DATA**

**7<sup>th</sup> & 8<sup>th</sup> June, 2017**

## **LOCATION:**

Keele Sustainability Hub, Home Farm, Keele University,  
Keele, Staffordshire, ST5 5AA

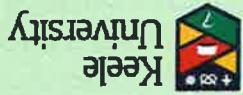


**Keele  
University**

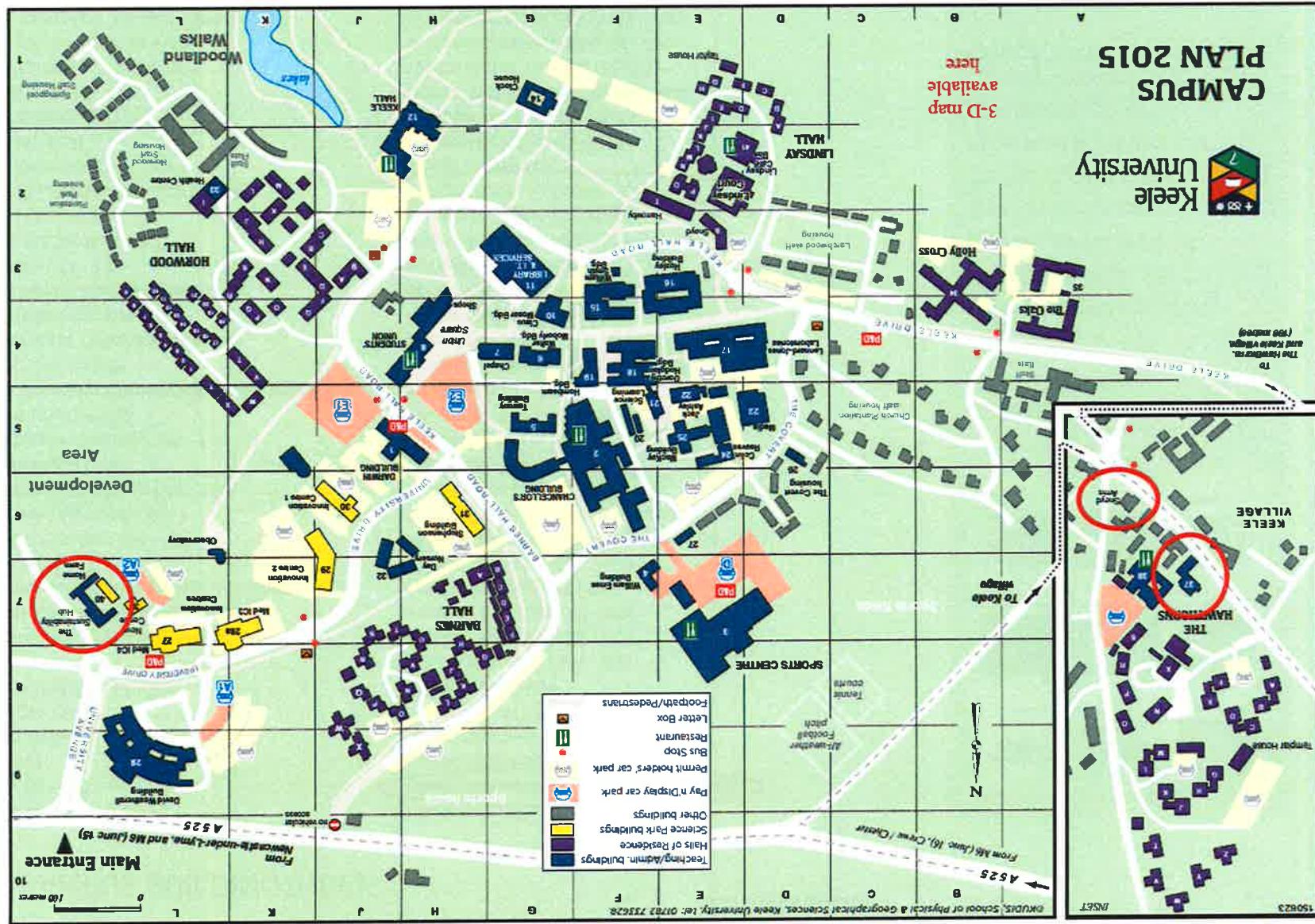
*Research Institute for Primary Care  
& Health Sciences*



# PLAN 2015 CAMPUSES



Here  
available  
3-D map



## KEELE CAMPUS BUILDING INDEX

Buildings & Facilities  
 Art Gallery (2 - F5)  
 Bank (8 - H4)  
 Barnes Hall (H7)  
 Bookshop (8 - H4)  
 Chancellor's Building (2 - F5)  
 Chapel (7 - G4)  
 Church Plantation housing (C5 etc)  
 Claus Moser Research Centre (10 - G4)  
 Clinical Education Centre (Hospital campus)  
 Clock House (14 - G1)  
 Colin Reeves Building (24 - E5)  
 Covert housing (around D6)  
 David Weatherall Building (28 - L9)  
 Darwin Building (1 - J5)  
 Dorothy Hodgkin Building (18 - F4)  
 Exhibition Suite (2 - F2)  
 Guy Hilton Research Centre (Hospital campus)  
 Harrowby (Hall) (in F2)  
 Hartshill Campus (Hospital campus)  
 Hawthorns Hall (see inset)  
 Hawthorns Restaurant (38 - inset)  
 Health Centre (33 - L2)  
 Hexagon (Lindsay Hall) (E2)  
 Holly Cross (34 - B3)  
 Home Farm Buildings (39 & 40 in L7)  
 Hornbeam Building (19 - F4)  
 Horwood Hall (L3)  
 Horwood housing (L2)  
 Huxley Building (16 - E3)  
 Innovation Centre IC1 (30 - J6)  
 Innovation Centre IC2 (29 - J7)  
 Innovation Centres IC3 & 4 (27,28a - K6)  
 Jack Ashley Building (22 - E5)

Keele Hall (12 - H1)  
**Keele Management Centre (37 - inset)**  
 Larchwood housing (D3)  
 Lennard-Jones Labs. (17 - E4)  
 Library and IT Services (11 - G3)  
 Lindsay Hall (D2)  
 Lindsay Studio 2 (41 - D2)  
 Mackay Building (25 - E5)  
 MacKay Institute (22 - E5)  
 Media Building (23 - D5)  
 Med IC 3 & 4 (27,28a - L7)  
 Medical School (Keele Campus) (28 - L9)  
 Medical School (Hospital Campus)  
 Novacentre (39 in J7)  
 Oaks (35 - A3)  
 Observatory (in L6)  
 Plantation Park housing (in L2)  
 Reception (2 - F5)  
 Science Learning Centre (21 - F5)  
 Shops (8 - H4)  
 Sneyd (Hall) (E3)  
 Sports Centre (3 - E7)  
 Springpool housing (L1)  
 Stephenson Building (31 - H6)  
**Students' Union (8 - H4)**  
**Sustainability Hub (40 - L7)**  
 Taylor House (E1)  
 Tawney Building (5 - G5)  
 Templar House (see inset)  
 Walter Moberly Building (6 - G4)  
 Westminster Theatre (2 - F5)  
 William Emes Building (4 - F7)  
 William Smith Building (15 - F4)

Accommodation

Course location



## COURSE TIMETABLE

### DAY 1:

9:00 – 9:30: Registration & tea/coffee

9:30 - 9:45: Welcome & Agenda for the day (includes MAD MINUTES for faculty)

9:45 - 10:15: Lecture 1: Rationale for meta-analysis & IPD (includes objectives for the course)

10:15 - 11:15: Lecture 2: Two-stage IPD meta-analysis

*11:15 - 11:40: Tea & Coffee*

11:40 – 11:55: “MAD MINUTE”: Part A

11:55 – 13:00: Stata Practical 1: Two-stage IPD meta-analysis

*13:00 – 13:45: Lunch*

13:45 - 14:45: Lecture 3: One-stage IPD meta-analysis

14:45 – 15:00: “MAD MINUTE”: Part B

15:00 - 16:00: Stata practical 2: One-stage IPD meta-analysis

*16:00 – 16:30: Tea and coffee*

16:30 – 16:45: “MAD MINUTE”: Part C

16:45 – 17:45: Lecture 4: Special topics (Power, one-stage weights, prognostic factors, DTA)

*19:00: EVENING MEAL AT LOCAL PUB (SNEYD ARMS, KEELE)*

### DAY 2:

09:00 – 10:00: Lecture 5: Estimation of effect modifiers (interactions)

10:00 - 10:45: Stata practical 3: Interactions

*10:45 – 11:05: Tea and Coffee*

11:05 - 11:30: Stata practical 3: Interactions (continued)

11:30 – 12:30: Lecture 6 (Part 1): Multivariate meta-analysis using IPD

*12:30 – 13:15: Lunch*

13:15 – 14:00: Lecture 6 (Part 2): Multivariate meta-analysis using IPD

14:00 – 15:15: Stata practical 4: Multivariate and network meta-analysis

*15:15 – 15:30: Tea and coffee*

15:30 – 16:45: GUEST LECTURE: Dr Thomas Debray (Utrecht) “Multiple imputation in an IPD meta-analysis”

16:45: CLOSE & FEEDBACK

## COURSE FACULTY

### **Professor Richard Riley**

Richard Riley is a Professor of Biostatistics at Keele University, having previously held posts at the Universities of Birmingham, Liverpool and Leicester. He joined Keele in October 2014 and his role focuses on statistical and methodological research for prognosis and meta-analysis, whilst supporting clinical projects in these areas. He is also a Statistics Editor for the BMJ and a co-convenor of the Cochrane Prognosis Methods Group. In meta-analysis, he specialises in methods for dealing with multiple correlated outcomes, and for synthesising individual participant data (IPD). In prognosis, Richard co-leads the PROGRESS initiative (PROGnosis RESearch Strategy) that seeks to improve the standards of prognosis research. This includes statistical methods to identify prognostic factors, to develop and validate prognostic models (risk prediction models), and to identify predictors of treatment response for stratified medicine (predictive markers). Richard will deliver three of the lectures and facilitate the practical sessions.

### **Dr Danielle Burke**

Danielle Burke is a Research Associate in Medical Statistics at Keele University. She previously held a biostatistician post within the Dental School at the University of Birmingham and completed her PhD at the University of Birmingham in December 2014. She joined Keele in June 2015 and is currently undertaking a post-doctoral fellowship from the School for Primary Care Research to apply and develop methods for individual participant data meta-analysis in primary care research. Danielle will deliver two of the lectures and facilitate the practical sessions.

### **Mr Joie Ensor**

Joie Ensor is a Research Associate in Medical Statistics at Keele University. He previously worked at the University of Birmingham as a biostatistician on a Health Technology Assessment project to develop and validate a prognostic model and clinical decision rule using IPD meta-analysis techniques. He has recently submitted his PhD in meta-analysis for prognosis and prediction, which has a specific interest in how to use IPD meta-analysis to explain heterogeneity. He joined Keele in January 2015 and his research interests continue to focus on methodological advances in using IPD meta-analysis for the development and validation of prediction models. Joie will facilitate the practical sessions.

### **Mr David Fisher**

David Fisher is a statistician at the MRC Clinical Trials Unit at UCL in London, where he has worked for nine years. He has a dual role, spending half his time as a trial statistician in the field of colorectal cancer, including the UK-wide stratified medicine platform trial FOCUS4; and half his time as lead statistician for the CTU's Meta-Analysis Group, which has a long history of successfully forming collaborative groups to obtain and analyse IPD, mostly in the field of cancer and with time-to-event outcomes. He is involved in methodological research in areas such as comparing one-stage and two-stage meta-analysis models; investigating whether, by how much, and under what circumstances aggregate data (AD) might be biased compared to IPD; and how best to analyse, present and interpret treatment-covariate interactions. He is also the author and maintainer of the ipdmetan package for Stata which facilitates two-stage IPD meta-analyses of overall effects and interactions. David will assist with the practical sessions and deliver the lecture on treatment-covariate interactions.

## COURSE PARTICIPANTS

Name	Institution / Company
Alessandro Chiarotto	VU Amsterdam, Netherlands
Charlotte Schreurs	VU Amsterdam, Netherlands
Eirini Karyotaki	VU Amsterdam, Netherlands
Elaine Nicholls	Keele University, UK
Ellen Driessen	VU Amsterdam, Netherlands
Emma Hawk	RTI Health Solutions, USA
Gerjon Hannink	Radboudumc Nijmegen, Netherlands
Hannah Lennon	University of Manchester, UK
Jean-Gabriel Le Moine	RTI Health Solutions, USA
Mirjam Reijnders	VU Amsterdam, Netherlands
Rebecca Whittle	Keele University, UK
Reuben Ogollah	Keele University, UK
Verena Hoffman	Klinikum der Universität München, Germany
Xin Hui Chan	University of Oxford, UK
Joshua Vogel	WHO, Switzerland
Andrew Renahan	University of Manchester, UK
Siobhan Bourke	Bangor University, UK
Brent Leininger	University of Minnesota, USA
Feitong Wu	University of Tasmania, Australia
Sandro Gsteiger	Roche, Switzerland



## LECTURE 1: Rationale for meta-analysis and IPD

### Summary:

Meta-analysis methods combine the quantitative evidence across related studies to produce results based on a whole body of research, and as such they are an integral part of evidence-based medicine. Traditional methods for meta-analysis synthesise aggregate data obtained from study publications or study authors, such as a treatment effect estimate (e.g. an odds ratio) and its associated uncertainty (e.g. a standard error or confidence interval). An alternative but increasingly popular approach is the **meta-analysis of individual participant data (IPD)**, or individual patient data, where the raw individual-level data are obtained for each study and used for synthesis. We are in an era where data sharing is becoming expected, and the opportunities to conduct an IPD meta-analysis are increasing. In this first lecture, we discuss the rationale for meta-analysis in general, and then focus broadly on why IPD meta-analysis offers many potential opportunities. We also discuss the aims and learning objectives of this course.

### Learning objectives:

- Understand why meta-analysis is fundamental to evidence-based decision-making
- Understand the differences between aggregate data and IPD
- Begin to recognise the potential advantages of an IPD meta-analysis rather than an aggregate data meta-analysis
- Appreciate that the need for IPD depends on the research question and availability of suitable aggregate data from primary studies
- Appreciate that an IPD meta-analysis also has potential disadvantages, and may be vulnerable to selection bias, publication bias and availability bias
- Understand the objectives and structure of the course

### Key introductory references:

Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: conduct, rationale and reporting. *BMJ* 2010; 340: c221

Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet* 1993;341(8842):418-22.

Oxman AD, Clarke MJ, Stewart LA. From science to practice. Meta-analyses using individual patient data are needed. *JAMA* 1995;274(10):845-6.

Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. *Statistics in Medicine* 1995; 14: 2057-2079.

Stewart LA, Tierney JF. To IPD or Not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Evaluation & the Health Professions* 2002; 25: 76-97.

Simmonds MC, Higgins JPT, Stewart LA, Tierney JF, Clarke MJ, et al. Meta-analysis of individual patient data from randomised trials - a review of methods used in practice. *Clinical Trials* 2005; 2: 209-217.

Simmonds M, Stewart G and Stewart L. A decade of individual participant data meta-analyses: A review of current practice. *Contemp Clin Trials*. 2015; 45: 76-83.

Tierney JF, Vale C, Riley R, Smith CT, Stewart L, Clarke M, et al. Individual Participant Data (IPD) Meta-analyses of Randomised Controlled Trials: Guidance on Their Use. *PLoS Med*. 2015;12:e1001855.



## LECTURE 1: Rationale for meta-analysis & IPD

Mr Joie Ensor

Research Associate  
Keele University

e-mail: j.ensor@keele.ac.uk

### Aims of this lecture

- Discuss the rationale for meta-analysis
- Introduce rationale for collecting individual participant data (IPD)
- Advantages, disadvantages of an IPD meta-analysis
- Learning objectives for this course

### Essential reading

Much of the course is based on the principles within this article

## RESEARCH METHODS & REPORTING

### Meta-analysis of individual participant data: rationale, conduct, and reporting

Richard D Riley,<sup>1</sup> Paul C Lambert,<sup>2</sup> Ghada Abu-Zaid<sup>1</sup>

The use of individual participant data instead of aggregate data in meta-analyses has many potential advantages, both statistically and clinically. Richard D Riley and colleagues describe the rationale for individual participant data meta-analysis and outline how to conduct this type of study.

Meta-analysis involves combining and analysing quantitative evidence from related studies to produce

individual participant data meta-analyses, the starting point in conducting one, how the analysis should be reported, and what challenges this approach may bring.

#### What are individual participant data?

The term individual participant data refers to data collected from participants in a study. In a randomised trial, for example, the individual participant data could be the pre-treatment and post-treatment blood pressure, a treatment group (placebo, active), hypertension (baseline and at each time point), sex, age, and race, for each patient in each study (table). A set of individual participant data from multiple studies often comprises thousands of

### Some essential reading

#### GUIDELINES AND GUIDANCE

### Individual Participant Data (IPD) Meta-analyses of Randomised Controlled Trials: Guidance on Their Use

ayne F. Turner,<sup>1</sup> Claire Vale,<sup>1</sup> Richard Riley,<sup>2</sup> Caitrin Tupper Smith,<sup>3</sup> Lesley Stewart,<sup>4</sup> Michael Currie,<sup>5</sup> Michael Royston<sup>6</sup>

<sup>1</sup> Medical Research Institute of GOSH, London, London, United Kingdom; <sup>2</sup> Research Institute of Primary Care and Health Sciences, Keele University, Keele, United Kingdom; <sup>3</sup> MRC Health Psychology Unit, MRC Methodology Research Group, Department of Biostatistics, University of Liverpool, Liverpool, United Kingdom; <sup>4</sup> Centre for Review and Dissemination, University of York, York, United Kingdom; <sup>5</sup> Centre for Primary Care and Quantitative Health, Bristol, United Kingdom; <sup>6</sup> Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands

[View in full article](#)

#### Summary Points

- Systematic reviews can now incorporate Individual participant data extracted from publications or obtained from trial investigators
- Data (IPD) usually are larger-scale individual-level data that can bring about substantial improvements to the quantity and quality of data, give greater scope in the analysis, and provide more detailed and robust results
- The process of collecting, checking, and analysing IPD is more complex than for aggregate data, and not all IPD meta-analyses are done to the same standard, making it difficult to compare them

### Some essential reading

#### Tutorial

#### Research Synthesis Methods

Received 21 November 2014 Accepted 13 May 2015 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/jrsm.1160

### Get real in individual participant data (IPD) meta-analysis: a review of the methodology

Thomas P. A. Debray,<sup>a,b,\*</sup> Karel G. M. Moons,<sup>a,b</sup> Gert van Valkenhoef,<sup>c</sup> Orestis Eftimou,<sup>d</sup> Noemi Hummel,<sup>e</sup> Rolf H. H. Groenwold,<sup>a</sup> Johannes B. Reitsma,<sup>a,b</sup> and on behalf of the Getreal methods review group

Individual participant data (IPD) meta-analysis is an increasingly used approach for synthesising and investigating treatment effect estimates. Over the past few years, numerous methods for conducting an IPD meta-analysis (IPD-MAs) have been proposed, often making different assumptions and involving choices with regard to the way they are applied. This review aims to provide an overview of the available methods for performing an IPD-MA using evidence from clinical trials or non-randomised studies while investigating treatment effects. With this review, we aim to assist researchers in choosing the appropriate method for their IPD-MA and to highlight the strengths and weaknesses of the different methods for performing an IPD-MA.

\*Correspondence to: Thomas P. A. Debray, Department of General Internal Medicine, Erasmus MC, Rotterdam, The Netherlands.  
<sup>a</sup>Department of General Internal Medicine, Erasmus MC, Rotterdam, The Netherlands; <sup>b</sup>Department of Clinical Epidemiology and Biostatistics, Maastricht University, Maastricht, The Netherlands; <sup>c</sup>Department of Clinical Epidemiology and Biostatistics, Leiden University Medical Center, Leiden, The Netherlands; <sup>d</sup>Department of General Internal Medicine, Academic Medical Center, Amsterdam, The Netherlands; <sup>e</sup>Department of General Internal Medicine, Utrecht University, Utrecht, The Netherlands

Keywords: meta-analysis, IPD, evidence synthesis, randomised trials, meta-analysis of individual participant data, trial design

5

### Some essential reading

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: [www.elsevier.com/locate/conctri](#)

#### Review

### A decade of individual participant data meta-analyses: A review of current practice

Mark Simmonds,<sup>a,\*</sup> Gavin Stewart,<sup>b</sup>, Lesley Stewart,<sup>b</sup>

<sup>a</sup>Centre for Evidence-Based Learning, University of East Anglia, Norwich, UK

<sup>b</sup>School of Agriculture, Food and Rural Development, Newcastle University, UK

#### ARTICLE INFO

Editor's choice

Received 4 April 2015

Reviewed 13 June 2015

Accepted 13 June 2015

Available online 11 July 2015

Keywords: individual participant data, meta-analysis, systematic review, contemporary clinical trials

#### ABSTRACT

Individual participant data (IPD) systematic reviews and meta-analyses are often considered to be the gold standard for meta-analyses. For the past 10 years, the field of IPD meta-analyses has expanded rapidly, and the methodology and reporting of these analyses has changed in the field. This paper investigates current strengths and areas for improvement in IPD systematic reviews.

Abstract. A decade of review work has led to identify systematic reviews of collected and analysed IPD. Data were collected from 100 systematic reviews and publication on a variety of topics related to the topic of individual participant data, and the main methodological issues.

Results. There has been considerable growth in the area of IPD-MAs, and this is likely to continue. The main methodological issues at present are the lack of reporting of the systematic reviews, the lack of reporting of subgroup analyses or including covariates in a stage regression models, the lack of effect size analysis, however, are not often used. Another issue is the lack of reporting of the quality of the included studies. The quality of the included studies is often not assessed, and the data are not always well presented, including bias. IPD-MAs were often not well designed and should be well designed and conducted. The main issues studies and findings are IPD-MAs were sought and selected. Conclusion. This paper provides a summary of the current state of IPD systematic reviews and meta-analyses, and identifies a range of topics for improving the quality of reporting the field, the process of IPD systematic reviews, and the statistical methods employed there. In 2014, before the publication of the CONSORT IPD guidelines, specific to IPD systematic reviews, there is a need to improve reporting of this area.

© 2015 Elsevier Inc. All rights reserved.

## **Part 1:**

### **The need for meta-analysis**

7

## **Introduction**

- Decisions in healthcare & medicine need to be based on **all the available evidence**
- Individual studies may be subject to chance, have low power, and/or disagree in their recommendations
- We need formal and explicit methods for identifying, extracting, and combining the evidence across multiple studies, to help decision-makers
- Last 20 years has seen a drive toward **systematic reviews** for this purpose
- See Cochrane library: <http://www.cochrane.org/index.htm>

8

### **Meta-analysis (evidence synthesis)**

The statistical analysis that usually follows a systematic review, to combine the quantitative results from studies identified

#### **Derivation:**

meta: 'after', 'above', 'transcending'

#### **Definition:**

'the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings'  
Glass (1976)

A broader term is **evidence synthesis**

(we use meta-analysis & evidence synthesis interchangeably)

9

### **Why meta-analysis?**

- To quantify effect sizes and their uncertainty to produce **summary (pooled) results** that informs clinical decision making
- To reduce problems of interpretation due to sampling error of study estimates (chance findings)
- Quantify genuine differences in the treatment effect across studies ('between-study heterogeneity')
- Identify factors that modify treatment effect ('**effect modifiers**')
- Identify statistical issues with evidence: e.g. publication bias,
- Identify further research needs

10

## **Part 2:**

### **Traditional meta-analysis framework**

11

### **Meta-analysis using aggregate data**

- Traditional meta-analysis uses **aggregate data**
- Obtainable from publications or study authors
- Meta-analysis of RCTs usually requires from each study:

an estimate of the treatment effect

e.g. odds ratio, relative risk, hazard ratio etc

and the standard error of this estimate

e.g. standard error of log hazard ratio

12

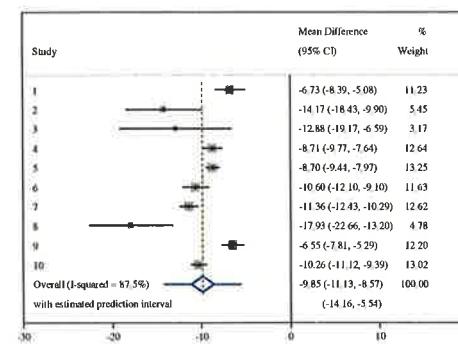
## Meta-analysis using aggregate data

### Advantages:

- Quick (in theory at least, if studies are well reported)
- Cheap
- Meta-analysis methods well established:  
such as inverse-variance, Mantel-Hansel,  
fixed effect, random effects, etc. (more later)
- Software suitable for non-statisticians (e.g. RevMan)
- Presentation includes forest plots, SROC plots etc.

13

### Example: meta-analysis of 10 hypertension trials



14

## Meta-analysis using aggregate data

### Disadvantages:

- Reliant on reporting of published articles
- Often face poor reporting (e.g. p-values rather than estimates)
- Not in control of the statistical analysis method used (see later)
- Inconsistency in choice of effect (hazard ratio, odds ratio, etc.)
- Vulnerable to publication bias: studies with significant results more likely to be published (or reported well) than non-significant studies
- Vulnerable to outcome reporting bias – studies report only those outcomes that were significant or most interesting

15

## Meta-analysis using aggregate data

### Disadvantages:

- Aggregate data collapses participant-level information
  - Observe study-level summaries, such as mean age, proportion male, overall treatment effect
  - Therefore, lose power to explain participant-level variation
  - Thus usually cannot identify subgroup results, treatment-covariate interactions (effect modifiers), etc.  
i.e. which patients do better than others
- Therefore hard to 'explain' causes of differences in effect across studies (more later)

16

## Part 3:

### Rationale for an IPD meta-analysis

17

## Call for IPD meta-analysis

IPD: Individual Patient Data , Individual Participant Data  
(the latter is now being adopted, as more inclusive)

- The original, raw individual-level data from the primary studies identified by the review
- The original source material, from which aggregate data are derived

### IPD meta-analysis:

The synthesis (in a statistical model) of the IPD from multiple studies for the purpose of summarising the evidence

- Increasingly relevant with the advent of 'stratified medicine' – the tailoring of treatment decisions for individual patients

18

## Call for IPD meta-analysis

- The limitations of aggregate data meta-analysis led to articles during the 1990s calling for more IPD meta-analyses

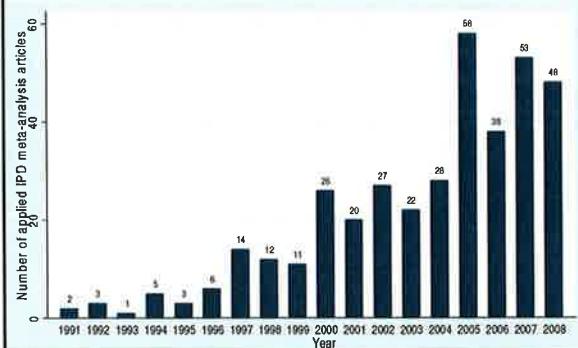
Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet* 1993;341(8842):418-22.

Oxman AD, Clarke MJ, Stewart LA. From science to practice. Meta-analyses using individual patient data are needed. *JAMA* 1995;274(10):845-6.

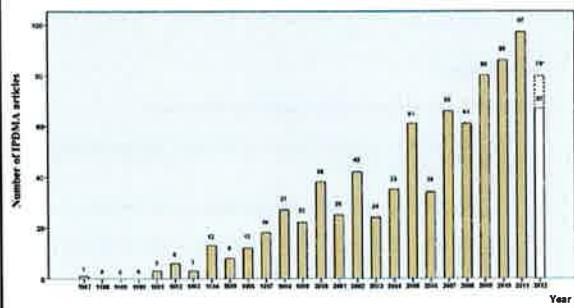
Stewart LA, Clarke MJ, on behalf of the Cochrane Working Party Group on Meta-analysis using Individual Patient Data (1995) Practical methodology of meta-analyses (overviews) using updated individual patient data. *Stat Med* 14: 2057-2079.

19

## Number of IPD meta-analysis articles over time (Riley et al., 2010)

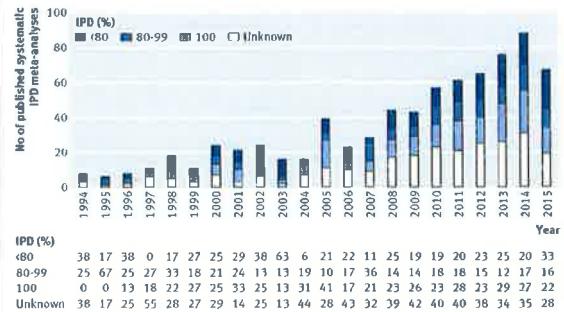


## A more recent review (Huang et al, 2014)



Huang, Y et al (2014) Distribution and Epidemiological Characteristics of Published Individual Patient Data Meta-Analyses. *PLoS ONE* 9(6): e100151

## An even more recent review (Nevitt et al, 2017)



Nevitt, S et al (2017) Exploring changes over time and characteristics associated with data retrieval across Individual participant data meta-analyses: systematic review. *BMJ* 357;j390

## Example: IPD from a meta-analysis of hypertension trials

Study	Patient	SBP Initial	SBP Final	treat	placebo	age	sex
1	1	190	185	1	0	58	1
1	2	175	172	1	0	69	1
1	3	184	185	0	1	39	0
1	4	192	182	0	1	45	1
2	1	201	199	1	0	51	0
2	2	169	154	1	0	42	1
2	3	171	170	0	1	50	1
2	4	179	168	0	1	67	0
3	1	197	167	1	0	83	1
3	2	189	171	1	0	78	0
3	3	184	188	0	1	55	1
3	4	168	161	0	1	61	0

etc ....  
(often contain thousands of patients)

23

## Example: IPD from a meta-analysis of prognosis studies

study	Patient	Marker levels			Adjustment factors		Survival & disease status		
		TH	LDH	MYCN	Age	Stage	Time of recurrence	Final survival time	Final disease status
1	1	Pos	200	5	3 yrs	1	*	150 days	ALIVE
1	2	Neg	350	3	2 yrs	4	330 days	390 days	DEAD
1	3	Neg	120	1	2 yrs	3	230 days	250 days	ALIVE with disease
2	1	Neg	320	1	6 yrs	4	27 days	48 days	DEAD

24

## Data sharing

We are in an era where data sharing is becoming expected...so the IPD meta-analyses approach is here to stay



BMJ 2013;306:e59 doi: 10.1136/bmj.e59 (Published 2 February 2013)

Page 1 of 2

## EDITORIALS

### Why data sharing should be the expected norm

The Institute of Medicine takes a step in the right direction but we should move even faster

Harlan M Krumholz professor of medicine

<sup>1</sup>Sieben of Cardiovascular Medicine and Robert Wood Johnson Foundation Clinical Scholars Program, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT 06510, USA; <sup>2</sup>Department of Health Policy and Management, Yale School of Public Health, New Haven; <sup>3</sup>Center for Outcomes Research and Evaluation, Yale New Haven Hospital, New Haven

25

## Data sharing

We are in an era where data sharing is becoming expected...so the IPD meta-analyses approach is here to stay



BMJ 2013;306:e59 doi: 10.1136/bmj.e59

*norm.” The IOM joins many other organizations, including drug companies,<sup>13</sup> the European Medicines Agency,<sup>14</sup> the National Institutes of Health,<sup>15</sup> and the Bill and Melinda Gates Foundation,<sup>16</sup> in making clear that study reporting and data sharing in medical research are imperative and the questions ahead are how, not whether.*

### Why data sharing should be the expected norm

The Institute of Medicine takes a step in the right direction but we should move even faster

Harlan M Krumholz professor of medicine

<sup>1</sup>Sieben of Cardiovascular Medicine and Robert Wood Johnson Foundation Clinical Scholars Program, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT 06510 USA; <sup>2</sup>Department of Health Policy and Management, Yale School of Public Health, New Haven; <sup>3</sup>Center for Outcomes Research and Evaluation, Yale New Haven Hospital, New Haven

26

## Obtaining IPD

Two common themes:

- (1) Perform **systematic review**, identify relevant studies, and then contact (email, phone) primary study authors to request their data  
and/or
- (2) Form a **collaborating group** with known or leading researchers in the field – the collaboration is mentioned on all publications, with a list of members; may be more likely to obtain IPD
  - Ideally we want all relevant studies to provide their IPD (or at least an unbiased sample of all the relevant studies)
  - Thus ideal approach is (1) [which may ultimately include (2)]
  - Many IPD meta-analyses are just from (2), which may be more susceptible to selection bias

27

## Meta-analysis using IPD

### Potential advantages:

- Use **consistent inclusion and exclusion criteria** across studies, and if appropriate reinstate individuals into the analysis who were originally excluded
- Observe and **account for missing data** at the individual-level
- Verify results presented in the original study publications (assuming IPD provided can be matched to that IPD used in the original analyses)
- **Inform risk of bias assessments:** for example, in regard to whether groups were balanced at baseline

28

## Meta-analysis using IPD

### Potential advantages:

- Use up-to-date follow-up information
  - potentially longer than that used in the original study publications
- Identify those studies which contain the same or overlapping sets of participants
- Calculate and incorporate results for those missing or poorly reported outcomes and summary statistics across published studies
  - may reduce the problem of selective within-study reporting (e.g. of outcomes)
- Calculate and incorporate results for unpublished studies
  - may thus reduce the problem of publication bias

29

## Meta-analysis using IPD

### Potential advantages:

- Standardise the strategy of statistical analysis across studies
  - e.g. the analysis method, how continuous variables are analysed, etc.)
  - use more appropriate/advanced methods than primary studies where necessary
- Assess model assumptions in each study
  - e.g. proportional hazards in Cox regression model
- Produce estimates adjusted for baseline (prognostic) factors
  - may increase power and allow adjustment for confounding factors
- Adjust for a more consistent set of baseline (prognostic) factors across studies

30

## Meta-analysis using IPD

### Potential advantages:

- Obtain meta-analysis results for specific subgroups of participants, and assess differential (treatment) effects across individuals
  - this facilitates individualised or stratified medicine
- Generate and validate prognostic/prediction models (risk scores), and examine multiple individual-level factors in combination
  - e.g. multiple biomarkers and genetic factors, and their interaction
- Account for the correlation between multiple endpoints
  - a meta-analysis of longitudinal data where each participant provides results at multiple time-points

31

## Meta-analysis using IPD

### Possible disadvantages:

- Costly (data managers, advanced statistics, specialised techniques, travel to see collaborators)
- Time-consuming (e.g. to obtain, collate, manage IPD)
- Inconsistent variables and data coded used from study to study
- Does not solve a study being poor quality (high risk of bias)
- More advanced statistical methods required
- Dealing with missing patient-level data
- Biases may remain

32

## Collecting IPD can be painful



BMJ 2012;344:e776 doi: 10.1136/bmj.e7762 (Published 2 December 2012)



Page 1 of 2

### VIEWS & REVIEWS

#### For just one of the trials:

##### OPEN DATA CAMPAIGN

#### Why did it take 19 months to retrieve clinical trial data from a non-profit organisation?

Asbjørn Hróbjartsson *The Nordic Cochrane Centre, Copenhagen, Denmark*

The emails we received during the prolonged exchange were all friendly, and the individuals involved were helpful and understood the need for data sharing, but they were hampered by inflexible, formalistic, and slow bureaucratic procedures. Since our first inquiry we communicated with four people, sent 25 emails, filled in four data use agreement forms, and waited one year and seven months.

33

## Biases in IPD meta-analyses

*Do not automatically view an IPD meta-analysis as 'gold standard' without considering how IPD studies were chosen*

- IPD only sought from published studies?
  - potential **publication bias** (published studies may have larger effects than unpublished studies)
- IPD only sought from a subset of known studies?
  - potential **selection bias** (chosen studies may not be an unbiased sample of all existing studies)
- IPD not provided by all those studies asked?
  - potential **availability bias** (studies providing their IPD may be systematically different from those refusing)

34

## Biases in IPD meta-analyses

*Do not automatically view an IPD meta-analysis as 'gold standard' without considering how IPD studies were chosen*

BMJ 2011;344:d7762 doi: 10.1136/bmj.d7762 (Published 3 January 2012)

Page 1 of 2

### RESEARCH

#### Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey

##### OPEN ACCESS

Ikhlaaq Ahmed *postgraduate student<sup>1</sup>*, Alexander J Sutton *professor of medical statistics<sup>2</sup>*, Richard D Riley *senior lecturer in medical statistics<sup>3</sup>*

<sup>1</sup> MRC Methods for Trials Methodology Research, School of Health and Population Sciences, University of Birmingham, Birmingham B15 2TT, UK; <sup>2</sup> Department of Health Sciences, University of Leicester, Leicester LE1 2PH, UK; <sup>3</sup> School of Health and Population Sciences, University of Birmingham

35

## Evidence synthesis using IPD: it may not be needed

### Decision depends on:

- The research question:
  - interest in overall treatment effect
  - interested in subgroups
  - identifying non-linear trends
  - developing risk prediction models
- The current analysis methods within primary studies (Are they appropriate? Do they give unbiased results?)
- The current reporting standards within primary studies (Is the aggregate data required for meta-analysis available?)

Need for  
IPD  
increases

36

### Evidence synthesis using IPD: it may not be needed

#### *Decision process for IPD approach:*

- What is the research question?
- Has a previous review been done before to answer this question?
- What aggregate data are required to answer the question?
- Are such aggregate data available in the majority of studies?
- If not, will availability of IPD allow them to be calculated?
- How much IPD can I realistically obtain?
- How long will it take to obtain it?
- Do I have the resources for obtaining, collating, checking and managing large sets of IPD?
- Do I have statistical resources to analyse the IPD?

**Aided by:** collaborating groups, different disciplines working together, leaders in the field being involved – & of course funding

### Part 4:

## Structure and objectives for this course

38

### Structure of the course

- Focus on **day one** is primarily on synthesis of IPD from RCTs to answer questions about an intervention effect
- Focus on **day two** is using evidence synthesis for more novel extensions including multiple outcomes and multiple treatments.
- Mixture of **6 lectures & 4 Stata practicals**
- **1 guest lecture** on day two
- 45 min lunch & tea/coffee breaks for **networking**
- Three '**Mad MInute**' sessions
- **Pub meal tonight at 7pm** in Sneyd Arms, Keele Village

39

### Aims of the course

#### *Fundamental statistical methods for evidence synthesis of IPD*

- Be able to conduct one-stage and two-stage IPD meta-analyses
- Understand when and why one-stage and two-stage methods differ
- Recognise how to account for the clustering of participants within studies when conducting an IPD meta-analysis
- Ability to write-down and fit key IPD meta-analysis models for continuous, binary and time-to-event outcomes from RCTs
- Understand how to extend these models to estimate effect modifiers (**treatment-covariate interactions**) for stratified medicine

40

### Aims of the course

#### *Fundamental statistical methods for evidence synthesis of IPD*

- Appreciate how publication and related biases may affect an IPD meta-analysis
- Appreciate how to combine IPD studies and non-IPD studies
- Be able to use IPD to deal with multiple correlated outcomes in a multivariate meta-analysis
- To recognise when a multivariate IPD meta-analysis is needed
- Be able to perform a network IPD meta-analysis of multiple treatment groups for a binary outcome in a multivariate meta-analysis framework

41

### Aims of the course

#### *Fundamental statistical methods for evidence synthesis of IPD*

- Appreciate the extension of standard percentage study weight to the one-stage IPD meta-analysis model
- Recognise novel approaches to deal with missing data & outcomes within an IPD meta-analysis
- Understand the opportunities & challenges of using IPD to identify risk or prognostic factors
- Practise the key methods in Stata

42

## It's only a 2-day course

We focus here on adapting well-known statistical tools for the purposes of IPD meta-analysis, such as:

- linear, logistic and Cox regression
- parametric survival models
- multivariate-normal models

We cannot cover everything, but please ask if you want pointers for non-standard topics. For example, we do not consider:

- How to examine risk of bias of individual IPD studies (should be done before any IPD meta-analysis ... but actually rarely done)
- How to check model assumptions (e.g. normality of residuals in linear regression models, proportional hazards in Cox models)
- Ordinal outcomes, competing risks, Bayesian methods

43

## What we assume

Please recognise the mixture of skills amongst all the participants taking this course (statisticians, epidemiologists, clinicians, etc.)

However, this is a statistical course with practical sessions in Stata, so we do assume:

- Attendees are familiar with basic statistical principles such as p-values, CIs, and effect estimates such as ORs, RRs, HRs, & mean differences
- Attendees are familiar with the use and interpretation of regression models, especially linear, logistic and Cox regression
- Have some experience of using Stata, in particular loading data, running do files, and/or interpreting regression results from Stata

44

## Enjoy it!

- We really hope you enjoy this course
- Many opportunities to ask questions:
  - at the end of lectures
  - during coffee and lunch breaks
  - during the practical sessions
- To stick to time, we ask that only essential questions or requests for clarification are made during the lectures
- We can also learn from you and your experiences - we certainly don't know all the answers
- There will be a feedback form for completion at end

45

## References (Introductory texts)

- Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: conduct, rationale and reporting. *BMJ* 2010; 340: c221
- Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: Is there a difference? *Lancet* 1993;341(8842):418-22.
- Oxman AD, Clarke MJ, Stewart LA. From science to practice. Meta-analyses using individual patient data are needed. *JAMA* 1995;274(10):845-6.
- Stewart LA, Clarke MJ, on behalf of the Cochrane Working Party Group on Meta-analysis using Individual Patient Data (1995) Practical methodology of meta-analyses (overviews) using updated individual patient data. *Statistics in Medicine* 14: 2057-2079.
- Stewart LA, Tierney JF (2002) To IPD or Not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Evaluation & the Health Professions* 25: 76-97.
- Stewart L, Tierney J, Burdett S (2005) Do systematic reviews based on individual patient data offer a means of circumventing biases associated with trial publications? In: Rothstein H, Sutton A, Borenstein M, editors. *Publication Bias in Meta-Analyses: Prevention, Assessment and Adjustments*. Chichester: John Wiley & Sons. pp 261-286
- Simmonds MC, Higgins JP, Stewart LA, Tierney JF, Clarke MJ, et al. (2005) Meta-analysis of individual patient data from randomised trials - a review of methods used in practice. *Clinical Trials* 2: 209-217.
- Ahmed I, Sutton AJ, Riley RD (2012) Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. *BMJ* 344: d7762

## LECTURE 2: Two-stage IPD meta-analysis

### Summary:

The statistical approach to IPD meta-analysis involves a two-stage or a one-stage approach. In the two-stage approach, the IPD are first analysed separately in each study using a statistical method appropriate for the type of data being analysed; for example, for continuous outcomes such as blood pressure a linear regression might be fitted, or for time-to-event data such as mortality a Cox regression might be applied. This produces aggregate data for each study, such as a treatment effect estimate and its standard error; these are then synthesised in the second stage using a suitable model for meta-analysis of aggregate data, such as one weighting by the inverse of the variance whilst assuming fixed or random (treatment) effects across studies. In this lecture, the two-stage approach is covered in detail using equations and illustrated examples, and the key estimation methods and summary inferences are discussed.

### Learning objectives:

- Understand the meaning of a ‘two-stage’ approach to IPD meta-analysis
- Know how to use the IPD to derive treatment effect estimates and variances using linear, logistic, and Cox regression
- Understand the difference between ANCOVA, change score and final score approaches to analysis of continuous outcome data
- Know how to perform fixed effect and random effects meta-analysis to combine effect estimates across studies
- Appreciate different estimation methods for implementing the random effects model
- Appreciate the difference in interpretation of summary results from a fixed effect and random effects meta-analysis
- Understand how to derive approximate 95% prediction intervals following a random effects meta-analysis
- Understand how to combine IPD and non-IPD studies using the two-stage approach

### Key references:

1. Higgins JP, Whitehead A, Turner RM, Omar RZ, Thompson SG: Meta-analysis of continuous outcome data from individual patients. *Stat Med* 2001; 20: 2219-4
2. Turner RM, Omar RZ, Yang M, et al: A multilevel model framework for meta-analysis of clinical trials with binary outcomes. *Stat Med* 2000, 19: 3417-3432.
3. Tudur-Smith C, Williamson PR, Marson AG: Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. *Stat Med* 2005, 24: 1307-1319.
4. Riley RD, Kauser I, Bland M, Wang J, Gueyffier F, Thijs L, Deeks JJ. Meta-analysis of continuous outcomes according to baseline imbalance and availability of individual participant data. *Stat Med* 2013; 32(16):2747-66. doi: 10.1002/sim.5726
5. Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society, Series A*. 2009;172:137-59.
6. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ*. 2011;342:d549.
7. Debray TP, Moons KG, van Valkenhoef G, et al. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. *Res Synth Methods*. 2015.



## LECTURE 2: Two-stage IPD meta-analysis

**Prof Richard Riley**

Professor of Biostatistics  
*Keele University*

e-mail: r.rlley@keele.ac.uk

### Aims

- Introduce most common IPD meta-analysis method
- Fixed effect and random effects approaches
- Clarify interpretation of summary results
- Confidence intervals
- Prediction intervals
- Consider how to combine IPD and aggregate data

### Part 1:

#### Traditional meta-analysis of aggregate data

#### Meta-analysis of aggregate data

- What if IPD are not available?
- Traditional meta-analysis uses a two-stage approach
  - STEP 1: Obtain aggregate results from each study
  - STEP 2: Combine results using fixed or random effects meta-analysis
- Consider step 1 in more detail ...  
e.g., Mean age, Proportion male, 2 by 2 table  
Overall treatment effect estimate and its 95% CI
- These are study-level summaries
- Problematic if interested in patient-level interactions (see Lecture 5)
- Theoretically fine if just want overall treatment effect  
... however, reliant on methods & reporting within primary studies<sup>4</sup>

#### How are treatment effects derived and reported for continuous outcomes?

Available treatment effect estimate depends on:  
- statistical approach used by original authors  
- the reporting quality within the publication

For continuous outcomes, reported effect estimate could be:

- mean treatment difference for SBP(final)
- mean treatment difference for SBP(final) – SBP(initial)
- mean treatment difference for SBP(final), adjusted for SBP(initial)
- mean treatment difference for SBP(final), adjusted for SBP(initial) & age etc.

#### How are treatment effects derived and reported for continuous outcomes?

Available treatment effect estimate depends on:  
- statistical approach used by original authors  
- the reporting quality within the publication

For continuous outcomes, reported effect estimate could be:

- mean treatment difference for SBP(final)
- mean treatment difference for SBP(final) – SBP(initial)
- mean treatment difference for SBP(final), adjusted for SBP(initial)
- mean treatment difference for SBP(final), adjusted for SBP(initial) & age etc.

*Increasing statistical power*

**How are treatment effects derived and reported for continuous outcomes?**

Available treatment effect estimate depends on:

- statistical approach used by original authors
- the reporting quality within the publication

For continuous outcomes, reported effect estimate could be:

- FINAL score
- CHANGE score
- FINAL score, ADJUSTED for baseline score ('ANCOVA')
- FINAL SCORE, ADJUSTED for baseline score & prognostic factors

*Increasing statistical power*



**How are treatment effects derived and reported for continuous outcomes?**

Available treatment effect estimate depends on:

- statistical approach used by original authors
- the reporting quality within the publication

For continuous outcomes, reported effect estimate could be:

- FINAL score
- CHANGE score
- FINAL score, ADJUSTED for baseline score ('ANCOVA')
- FINAL SCORE, ADJUSTED for baseline score & prognostic factors

*Increasing statistical power (Riley et al., 2013)*



**With IPD, greater potential to fit the model *you* want, use the same model consistently in each study, and obtain estimates *you* want ...**

**Statistics in Medicine** 2013; 32: 2747–2766

Received 14 November 2011; Accepted 22 December 2012; Published online 10 January 2013 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/sim.5726

**Research Article**

**Meta-analysis of randomised trials with a continuous outcome according to baseline imbalance and availability of individual participant data<sup>a</sup>**

Richard D. Riley,<sup>a,\*</sup> Iram Kauser,<sup>a</sup> Martin Bland,<sup>b</sup> Lutgarde Thijs,<sup>c</sup> Jan A. Staessen,<sup>c,d</sup> Jiguang Wang,<sup>e</sup> François Guéyffier<sup>d</sup> and Jonathan J. Deeks<sup>a</sup>

We describe methods for meta-analysis of randomised trials where a continuous outcome is of interest, such as blood pressure measured at both baseline, pre-treatment and 12-week post-treatment. We use four examples for illustration, comparing studies with and without individual participant data (IPD) and with and without baseline imbalance. We compare two approaches to estimate treatment effect estimates derived using analysis of covariance (ANCOVA), a restriction of just scores, or a regression of the change scores. When there is baseline imbalance in the trials, the treatment effect estimates from the four approaches are very similar, provided that the trials are balanced. However, we show that meta-analytic results for the summary treatment effect are similar regardless of the approach taken. Then, without IPD, if trials are balanced, it is better to fit a simple model treatment effect estimate derived from ANCOVA, as this adjusts for imbalance and accounts for the correlation between baseline and follow-up measurements. If trials are unbalanced, it is better to fit a regression model treatment effect estimate derived from IPD and with stratifiable ANCOVA estimates, researchers should limit meta-analyses to those trials with no baseline imbalance. Treatment's model treatment effect for baseline imbalance without IPD performs poorly in our example and so is not recommended.

Finally, we extend the ANCOVA model to estimate the interaction between treatment effect and baseline values and compare options for estimating this interaction given only aggregate data. Copyright © 2013 John Wiley & Sons, Ltd.

**Similar issue for other outcomes**

Available treatment effect estimate depends on:

- statistical approach used by original authors
- the reporting quality within the publication

e.g. For **binary outcomes**, effect estimate could be given as:

- Odds ratio or Relative risk
- Difference in proportions
- Odds ratio adjusted for baseline imbalances (logistic regression)
- Chi-square test

(may even have the 2x2 table showing event numbers in each group)

**With IPD, greater potential to fit the model *you* want, use the same model consistently in each study, and obtain estimates *you* want ...**

**Similar issue for other outcomes**

Available treatment effect estimate depends on:

- statistical approach used by original authors
- the reporting quality within the publication

e.g. For **time-to-event outcomes**, effect estimate could be given as:

- Hazard ratio
- Difference in median survival
- Kaplan Meier curves, log-rank test
- % survival at n years
- Hazard ratio adjusted for baseline imbalances (Cox model)

**With IPD, greater potential to fit the model *you* want, use the same model consistently in each study, and obtain estimates *you* want ...**

**Part 2:**

---

**Two-stage IPD meta-analysis**

## Now assume we have IPD for all studies, such as for continuous outcomes ...

Study	Patient	SBP initial	SBP final	treat	placebo	age	sex
1	1	190	185	1	0	58	1
1	2	175	172	1	0	69	1
1	3	184	185	0	1	39	0
1	4	192	182	0	1	45	1
2	1	201	199	1	0	51	0
2	2	169	154	1	0	42	1
2	3	171	170	0	1	50	1
2	4	179	168	0	1	67	0
3	1	197	167	1	0	83	1
3	2	189	171	1	0	78	0
3	3	184	188	0	1	55	1
3	4	168	161	0	1	61	0

13

## Two-stage IPD approach

- Consider IPD available for all studies
- Want to assess a treatment effect (e.g. mean difference)
- Two-step approach:**
  - STEP 1: reduce the IPD to aggregate data in each study
  - STEP 2: pool aggregate data using standard meta-analysis methods
- Thus we mirror traditional two-stage approach
- But rather than extract aggregate data from published articles, we now analyse the IPD separately in each study to obtain it ourselves
- Stata module *ipdmetan* (Fisher, 2014) – see practical

## Two-stage IPD approach

- Let us focus on obtaining and pooling effect estimates
- So in general our approach is:

**STEP 1:** Perform a regression analysis to obtain effect estimate in each study separately  
(e.g. mean difference, odds ratio, or hazard ratio)

**STEP 2:** Pool these in a fixed effect or random effects model

NB: Other approaches are possible, for instance for binary outcomes could pool 2x2 tables from the studies using Peto or Mantel-Haenszel method; we do not consider these on this course as we prefer one-stage regression methods when analysing 2 by 2 tables directly

15

STATISTICS IN MEDICINE  
Statist. Med. 2001; 20:2219–2241 (DOI: 10.1002/sim.918)

### Meta-analysis of continuous outcome data from individual patients

Julian P. T. Higgins<sup>1,2\*</sup>, Anne Whitehead<sup>3</sup>, Rebecca M. Turner<sup>3</sup>,  
Runsun Z. Omar<sup>4</sup> and Simon G. Thompson<sup>5</sup>

<sup>1</sup>MRC Biostatistics Unit, Institute of Public Health, University, Fortie St George's, Robinson Way,  
Cambridge CB2 2RE, UK

<sup>2</sup>Medical and Pharmaceutical Statistics, Department of Mathematics, University of Reading, P.O. Box 240,  
Earley Gate, Reading RG6 6FX, UK

<sup>3</sup>MRC Clinical Trials Unit, 222 Euston Road, London NW1 2DA, UK

<sup>4</sup>Department of Epidemiology and Public Health, Imperial College School of Medicine,  
Hammersmith Hospital, Du Cane Road, London W12 0NN, UK

**SUMMARY**  
Meta-analyses using individual patient data are becoming increasingly common and have several advantages over meta-analysis of summary statistics. We explore the use of multilevel or hierarchical models for the meta-analysis of continuous individual patient outcome data from clinical trials. A general model is developed which encompasses traditional meta-analysis, as well as meta-regression and the inclusion of patient-level covariates. The investigation is helped by a comparison to two other different ways of analysing trials to determine their strengths. We focus on models with fixed trial effects, although an extension to a random effect for trial is described. The methods are illustrated on an example in Alzheimer's disease in a classical framework using SAS PROC MIXED and MLwiN, and in a Bayesian framework using BUGS. Relative merits of the three software packages for such meta-analyses are discussed. We also discuss the assumptions made and extensions to incorporate more than two treatments. Copyright © 2001 John Wiley & Sons, Ltd.

16

STATISTICS IN MEDICINE  
Statist. Med. 2002; 21:1843–1853  
Published online in 2002. DOI: 10.1002/sim.3165

### Meta-analysis of continuous outcomes combining individual patient data and aggregate data

Richard D. Riley<sup>1,4</sup>, Paul C. Lambert<sup>2</sup>, Jan A. Staessen<sup>3</sup>, Jigang Wang<sup>4</sup>,  
Francois Guillet<sup>5</sup>, Lodewijk Thijssen<sup>3</sup> and Florent Bouafia<sup>6</sup>

<sup>1</sup>Centre for Medical Statistics and Health Evaluation, Faculty of Medicine, University of Liverpool,  
Shore's Cottage, Broadgreen Campus, Liverpool L19 3LG, UK

<sup>2</sup>Centre for Biostatistics and Clinical Trials, Department of Health Sciences, University of Leicester, 2nd Floor, Adam Building, University Road, Leicester LE1 9RH, UK

<sup>3</sup>Staats Cordwagen Centrum, Division of Hypertension and Cardiovascular Rehabilitation, Department  
of Cardiovascular Disease, University of Antwerp, Antwerp, Belgium

<sup>4</sup>Department of Epidemiology and Clinical Trials, Ruijin Hospital, Shanghai Jiaotong University School  
of Medicine, Ruijin 2nd Road 1973, Shanghai 200032, China

<sup>5</sup>CHU Sainte-Justine, Department of Cardiology, 3175 Côte Sainte-Catherine, Montréal, Québec H3T 1C5, Canada

<sup>6</sup>CHU Sainte-Justine, Research Center, 3175 Côte Sainte-Catherine, Montréal, Québec H3T 1C5, Canada

### SUMMARY

Meta-analysis of individual patient data (IPD) is the gold standard for combining evidence across clinical studies. However, in some studies IPD may not be available and only aggregate data (ADs), such as a treatment effect estimator and its standard error, may be obtained. In this situation, methods for combining IPD and AD are required to utilize all the available evidence. In this paper, we describe and assess a range of statistical methods for combining IPD and AD in such settings. A number of issues from randomized controlled trials,

STATISTICS IN MEDICINE  
Statist. Med. 2000; 19:3417–3432

### A multilevel model framework for meta-analysis of clinical trials with binary outcomes<sup>†</sup>

Rebecca M. Turner<sup>1,2\*</sup>, Runsun Z. Omar<sup>3</sup>, Min Yang<sup>3</sup>, Harvey Goldstein<sup>3</sup>  
and Simon G. Thompson<sup>4</sup>

<sup>1</sup>MRC Clinical Trials Unit, 122 Euston Road, London NW1 2DA, UK

<sup>2</sup>Department of Epidemiology and Public Health, Imperial College School of Medicine, Du Cane Road,  
London NW1 2DA, UK

<sup>3</sup>Mathematical Sciences Group, Institute of Education, 26 Bedford Way, London WC1H 0AE, UK

<sup>4</sup>MRC Biostatistics Unit, Institute of Public Health, Cambridge CB2 2SR, UK

### SUMMARY

In this paper we explore the potential of multilevel models for meta-analysis of trials with binary outcomes for both summary data such as log-odds ratios, and individual patient data. Compared to fixed effect and random effect models, multilevel models allow for correlations between trials and between studies. In contrast to standard maximum likelihood estimation, we use the results from 22 trials to perform Bayesian trials synthesis, as this makes comparisons with a second example dataset comprising 20 trials easier. We compare the multilevel model with a standard random effects model. The multilevel variance may be derived from likelihood based methods or a parametric bootstrap as well as from Wald methods. When modelling individual patient data, a binomial bootstrap may be used to provide unbiased estimates of the parameters. The multilevel model is also compared to a model for the between-trial variance. The trial effects may be modelled as either fixed or random within individual data models, and we discuss the corresponding assumptions and implications. If random trial effects are used, the multilevel model is equivalent to a random effects model. If fixed trial effects are used, the multilevel model is equivalent to a logistic approach to meta-analysis. Having implemented these techniques, the flexibility of multilevel modelling may be explored in facilitating extensions to standard meta-analysis methods. Copyright © 2000 John Wiley & Sons, Ltd.

18

### Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes

Catrin Tudur Smith<sup>1,2,\*</sup>, Paula R. Williamson<sup>1</sup> and Anthony G. Marson<sup>2</sup>

<sup>1</sup>Centre for Medical Statistics and Health Evaluation, University of Liverpool, Liverpool L69 3BX, UK

<sup>2</sup>Department of Neurological Sciences, University of Liverpool, Liverpool, UK

#### SUMMARY

Differences between studies in terms of design features and methodology, clinical procedures, and patient characteristics, i.e. factors that can contribute variation in the treatment effect between studies in a meta-analysis (statistical heterogeneity), regression modeling can be used to examine relationships between treatment effect and covariates with the aim of explaining the variability in terms of clinical, methodological, and other factors. Such models can be fitted at the study level or at the level of the individual patient data. An aggregate data approach can be problematic as sufficient data are rarely available and translating aggregate effects to individual patient can often be misleading. An individual patient data approach, although usually more resource demanding, allows a thorough investigation of potential sources of heterogeneity and enables a full analysis of the data in event time meta-analysis.

Hierarchical Cox regression models are used to identify and explore the evidence for heterogeneity in event analysis and examine the relationship between covariates and survival. Fisher exact data in the context of meta-analysis and alternative formulations of the model are possible and illustrated using individual patient data from a meta-analysis of five randomized controlled trials which compare two drugs for the treatment of stroke. The results show that the treatment effect is significant and the magnitude of treatment effect is small. The behaviour of each model in each situation is explored and compared. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: heterogeneity; time to event; metaanalysis; Cox regression model; individual patient data

19

## Two-step approach

STEP 1: For each IPD study separately, reduce the IPD to AD

This requires a statistical analysis of the IPD in each study

20

## Two-step approach

STEP 1: For each IPD study separately, reduce the IPD to AD

This requires a statistical analysis of the IPD in each study

Q: What is the treatment effect on systolic blood pressure, SBP

- $i = 1$  to  $k$  trials
- The  $j^{th}$  patient provides their SBP after treatment,  $SBP_{ij}$  and their SBP at baseline  $SBP_{0ij}$
- Treatment group ( $x_{ij} = 1$ ) and Control group ( $x_{ij} = 0$ )
- Fit:

21

## Two-step approach

STEP 1: For each IPD study separately, reduce the IPD to AD

This requires a statistical analysis of the IPD in each study

Q: What is the treatment effect on systolic blood pressure, SBP

- $i = 1$  to  $k$  trials: Continuous outcome – ANCOVA model
- The  $j^{th}$  patient provides their SBP after treatment,  $SBP_{ij}$  and their SBP at baseline  $SBP_{0ij}$
- Treatment group ( $x_{ij} = 1$ ) and Control group ( $x_{ij} = 0$ )
- Fit:

$$SBP_{ij} = \varphi_i + \beta_{1i} SBP_{0ij} + \theta_i x_{ij} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N(0, \sigma^2)$$

Final SBP                                  Adjust for baseline SBP  
 Control effect                              Treatment effect                            Residual error

## Two-step approach

STEP 1: For each IPD study separately, reduce the IPD to AD

This requires a statistical analysis of the IPD in each study

Q: What is the treatment effect on systolic blood pressure, SBP

- $i = 1$  to  $k$  trials: Continuous outcome – ANCOVA model
- The  $j^{th}$  patient provides their SBP after treatment,  $SBP_{ij}$  and their SBP at baseline  $SBP_{0ij}$
- Treatment group ( $x_{ij} = 1$ ) and Control group ( $x_{ij} = 0$ )
- Fit:

$$SBP_{ij} = \varphi_i + \beta_{1i} SBP_{0ij} + \theta_i x_{ij} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N(0, \sigma^2)$$

Gives mean difference estimate ( $\hat{\theta}_i$ ) & its variance ( $V(\hat{\theta}_i)$ )

## Two-step approach

STEP 1: For each IPD study separately, reduce the IPD to AD

This requires a statistical analysis of the IPD in each study

Q: What is the treatment effect on a binary outcome (e.g. safe birth at end of pregnancy)?

- $i = 1$  to  $k$  trials: Binary outcome - logistic regression
- The  $j^{th}$  patient provides their outcome response (0 or 1)
- $p_{ij}$  is the probability of an event
- Treatment group ( $x_{ij} = 1$ ) and Control group ( $x_{ij} = 0$ )
- Fit:

$$\ln(p_{ij}/(1-p_{ij})) = \varphi_i + \beta_{1i} SBP_{0ij} + \theta_i x_{ij}$$

Gives log odds ratio estimate ( $\hat{\theta}_i$ ) & its variance ( $V(\hat{\theta}_i)$ )

24

## Two-step approach

**STEP 1:** For each IPD study separately, reduce the IPD to AD

This requires a statistical analysis of the IPD in each study

**Q: What is the treatment effect on a time-to-event outcome (e.g. mortality rate after diagnosis of lung cancer)?**

- $i = 1$  to  $k$  trials: Time-to-event outcomes - Cox regression
- The  $j^{\text{th}}$  patient provides their follow-up time ( $t_{ij}$ ), and outcome response (0 or 1, censored or event) at that time
- $h_{ij}(t)$  is the hazard rate over time;  $h_{0i}(t)$  is the baseline hazard
- Treatment group ( $x_{ij} = 1$ ) and Control group ( $x_{ij} = 0$ )
- Fit:

$$h_{ij}(t) = h_{0i}(t)\exp(\beta_{0i} + \beta_{1i}x_{ij})$$

Gives log hazard ratio estimate ( $\hat{\theta}_i$ ) & its variance ( $V(\hat{\theta}_i)$ )

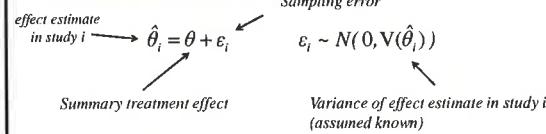
25

## Two-step approach

**STEP 2:** Combine the aggregate data (i.e.  $\hat{\theta}_i$  &  $\text{var}(\hat{\theta}_i)$ ) using standard meta-analysis methods

**OPTION (i):**

- fixed-effect meta-analysis



26

## Two-step approach

**STEP 2:** Combine the aggregate data (i.e.  $\hat{\theta}_i$  &  $\text{var}(\hat{\theta}_i)$ ) using standard meta-analysis methods

**OPTION (i):**

- fixed-effect meta-analysis

Maximum likelihood estimation leads to summary estimate

$$\hat{\theta} = \frac{\sum_{i=1}^k \hat{\theta}_i w_i}{\sum_{i=1}^k w_i} \quad \text{var}(\hat{\theta}) = \frac{1}{\sum_{i=1}^k w_i}$$

$$\text{where } w_i = \frac{1}{V(\hat{\theta}_i)}$$

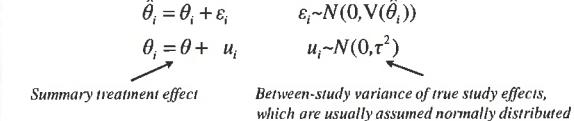
27

## Two-step approach

**STEP 2:** Combine the aggregate data (i.e.  $\hat{\theta}_i$  &  $\text{var}(\hat{\theta}_i)$ ) using standard meta-analysis methods

**OPTION (ii):**

- random-effects meta-analysis



28

## Two-step approach

**STEP 2:** Combine the aggregate data (i.e.  $\hat{\theta}_i$  &  $\text{var}(\hat{\theta}_i)$ ) using standard meta-analysis methods

**OPTION (ii):**

- random-effects meta-analysis

$$\begin{aligned} \hat{\theta}_i &= \theta_i + \varepsilon_i & \varepsilon_i &\sim N(0, V(\hat{\theta}_i)) \\ \theta_i &= \theta + u_i & u_i &\sim N(0, \tau^2) \end{aligned}$$

Summary treatment effect      Between-study variance of true study effects

Calculate  $I^2$ : % total variation due to between-study variance (Higgins et al., 2003)

## Two-step approach

**STEP 2:** Combine the aggregate data (i.e.  $\hat{\theta}_i$  &  $\text{var}(\hat{\theta}_i)$ ) using standard meta-analysis methods

**OPTION (ii):**

- random-effects meta-analysis

Maximum likelihood estimation leads to:

$$\hat{\theta} = \frac{\sum_{i=1}^k \hat{\theta}_i w_i^*}{\sum_{i=1}^k w_i^*} \quad \text{var}(\hat{\theta}) = \frac{1}{\sum_{i=1}^k w_i^*}$$

$$\text{where } w_i^* = \frac{1}{V(\hat{\theta}_i) + \tau^2}$$

&  $\tau^2$  typically estimated by either Dersimionian & Laird (methods of moments) or restricted maximum likelihood (REML)

## Two-step approach

**STEP 2:** Combine the aggregate data (i.e.  $\hat{\theta}_i$  &  $\text{var}(\hat{\theta}_i)$ ) using standard meta-analysis methods

### OPTION (ii)

- random-effects meta-regression

$$\begin{aligned}\hat{\theta}_i &= \theta_i + \varepsilon_i & \varepsilon_i &\sim N(0, V(\hat{\theta}_i)) \\ \theta_i &= a + b \times \text{dose}_i + u_i & u_i &\sim N(0, \tau^2)\end{aligned}$$

*A study's true treatment effect depends on dose used*      *Remaining between-study variance*

- Easily extend to meta-regression: include study-level covariates (e.g. year, mean age, dose) to explain/reduce  $\tau^2$

## Aside: Fixed-effect & random-effects model specification

- The models in the second stage have been presented in a linear model framework
- A fixed-effect meta-analysis can also be written:

$$\hat{\theta}_i \sim N(\theta, V(\hat{\theta}_i))$$

- A random-effects meta-analysis can also be written:

$$\begin{aligned}\hat{\theta}_i &\sim N(\theta_i, V(\hat{\theta}_i)) \\ \theta_i &\sim N(\theta, \tau^2)\end{aligned}$$

## Applied example: hypertension

- Wang et al. (2005) performed a quantitative overview of trials in hypertension to investigate hypertension treatments and their lowering of systolic blood pressure (SBP).
- They selected randomised controlled trials that tested active anti-hypertensive drugs against placebo or no treatment
- IPD was sought from trials in the INdividual Data ANalysis of Anti-hypertensive intervention trials (INDANA) data set or at the Studies Coordinating Centre in Leuven (Belgium)
- Ten trials were ultimately included, and these provided IPD for a total of 28581 patients.

## Applied example: hypertension

- An IPD meta-analysis of the 10 trials is important to summarise the effect of anti-hypertensive drugs on SBP
- Specifically
  - (i) to examine the distribution of treatment effects across the trials in order to estimate the average effect (that is, how much anti-hypertensive drugs reduce SBP compared to control on average across the trial populations)
  - (ii) to quantify the amount of between-trial variation in the effect of anti-hypertension drugs
  - (iii) identify effect modifiers: patient-level factors that modify (interact with) treatment effect

## Sample of the data ...

Study	Patient	SBP Initial	SBP final	treat	placebo	age	sex
1	1	190	185	1	0	58	1
1	2	175	172	1	0	69	1
1	3	184	185	0	1	39	0
1	4	192	182	0	1	45	1
2	1	201	199	1	0	51	0
2	2	169	154	1	0	42	1
2	3	171	170	0	1	50	1
2	4	179	168	0	1	67	0
3	1	197	167	1	0	83	1
3	2	189	171	1	0	78	0
3	3	184	188	0	1	55	1
3	4	168	161	0	1	61	0

35

## Two-stage analysis

- The treatment and control groups are well balanced in SBP in each trial at baseline, and this was true for other patient characteristics.

**STEP 1:** ANCOVA in each trial, to obtain mean difference in SBP at follow-up after adjusting for baseline

Stata code example:

```
. bysort trialdummy: reg sbpl treat sbpi
```

## Stage 1 aggregate data

- Aggregate data for the 10 hypertension trials

ID	Trial name*	Control	Treatment	SBP baseline (mmHg)		SBP final (mmHg)		ANOVA: Treatment effect estimate (variance)
				Mean (SD)	Mean (SD)	Control	Treatment	
1	ATMH	750	789	153.05 (15.73)	152.28 (15.25)	139.75 (17.81)	132.85 (16.68)	-6.66 (0.72)
2	HEP	199	150	191.55 (17.64)	189.94 (16.15)	179.89 (22.15)	165.06 (20.03)	-14.17 (4.73)
3	EWPH	82	90	178.23 (15.06)	177.33 (15.85)	170.45 (26.91)	156.88 (21.26)	-12.88 (10.31)
4	HDFP	2371	2427	151.00 (19.53)	151.68 (19.83)	138.54 (21.26)	130.09 (19.25)	-8.71 (0.30)
5	MRC-1	3445	3548	156.65 (15.96)	156.60 (16.09)	144.25 (17.38)	135.49 (16.32)	-8.70 (0.14)
6	MRC-2	1337	1314	182.13 (17.27)	182.13 (12.63)	168.08 (19.71)	153.99 (20.13)	-10.60 (0.59)
7	SHEP	2371	2365	170.12 (0.24)	170.12 (0.50)	156.24 (20.12)	155.10 (19.05)	-11.36 (0.30)
8	STOP	131	137	191.15 (11.16)	191.63 (12.21)	189.11 (21.9)	171.46 (19.29)	-17.93 (5.82)
9	Sy-Chi	1139	1252	170.25 (11.41)	170.73 (10.90)	156.55 (16.86)	150.20 (15.84)	-6.55 (0.41)
10	Sy-Eur	2297	2308	173.94 (10.09)	173.75 (9.86)	165.24 (16.33)	154.87 (16.31)	-10.26 (0.20)

## How do ANCOVA results compare to other methods?

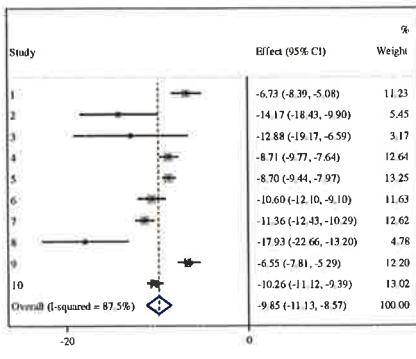
ID	Trial name*	FINAL SCORE: Treatment effect estimate (variance)		ANCOVA: Treatment effect estimate (variance)	CHANGE SCORE: Treatment effect estimate (variance)
		Treatment effect estimate (variance)	Treatment effect estimate (variance)		
1	ATMH	-6.90 (0.78)	-6.66 (0.72)	-6.66 (1.02)	-6.13
2	HEP	-14.83 (5.29)	-14.17 (4.73)	-14.17 (5.86)	-13.23
3	EWPH	-13.57 (13.56)	-12.88 (10.31)	-12.88 (10.53)	-12.68
4	HDFP	-8.44 (0.34)	-8.71 (0.30)	-8.71 (0.42)	-9.13
5	MRC-1	-8.76 (0.16)	-8.70 (0.14)	-8.70 (0.19)	-8.62
6	MRC-2	-10.59 (0.60)	-10.60 (0.58)	-10.60 (0.72)	-10.65
7	SHEP	-11.14 (0.32)	-11.36 (0.30)	-11.36 (0.31)	-11.51
8	STOP	-17.66 (6.34)	-17.93 (5.82)	-17.93 (6.28)	-18.18
9	Sy-Chi	-6.36 (0.45)	-6.55 (0.41)	-6.55 (0.49)	-6.84
10	Sy-Eur	-10.37 (0.23)	-10.26 (0.20)	-10.26 (0.21)	-10.18

## Two-stage analysis

- The treatment and control groups are well balanced in SBP in each trial at baseline, and this was true for other patient characteristics.

- STEP 1:** ANCOVA in each trial, to obtain mean difference in SBP at follow-up after adjusting for baseline
- STEP 2:** Random effects meta-analysis of the treatment effect estimates  
(with  $\tau^2$  estimated using Dersimonian & Laird's methods of moments approach)  
(though note: REML gives slightly larger  $\tau^2$   
– see lecture 3)

## Results



## Stata code and results

```
admetan treat se, re
Meta-analysis pooling of aggregate data
using the random-effects inverse-variance model
with DerSimonian-Laird estimate of tau2
```

Study	Effect	[95% Conf. Interval]	% Weight
1	-6.732	-8.389	11.13
2	-14.166	-18.426	5.48
3	-12.881	-19.174	3.17
4	-8.709	-9.774	12.64
5	-8.702	-9.436	7.967
6	-10.602	-12.109	11.63
7	-11.357	-12.428	12.62
8	-17.926	-22.655	4.78
9	-6.548	-7.810	12.20
10	-10.256	-11.122	13.02
Overall effect	-9.048	-11.126	100.00
Heterogeneity Measures			
	Value	df	p-value
Cochran's Q	72.02	9	0.000
I <sup>2</sup> (%)	87.5%		
Modified H <sup>2</sup>	7.003		
tau <sup>2</sup>	3.0653		

40

## Stata code and results

Alternatively there is a wonderful package called 'ipdmetan' that automates the first and second stages:

```
. ipdmetan, study(trial) re: reg sbpl treat sbpi
```

Studies included: 10  
Participants included: 28592

Meta-analysis pooling of main (treatment) effect estimate treat using the random-effects inverse-variance model with DerSimonian-Laird estimate of tau<sup>2</sup>

The pooled results are identical to those on the previous slides.

42

## Reminder:

### IPD may not be needed...

- Perhaps all trials:
  - used ANCOVA
  - reported the treatment effect estimate
  - reported its 95% confidence interval
- If we are only interested in summarising the overall treatment effect, the IPD is giving us nothing new (other things being equal, like length of follow-up, number of included patients, etc)
- Advantages of having IPD begin to arise when studies do not report the results, outcomes, subgroups of interest; use inconsistent analysis methods; etc .... (see lecture 1)

## Part 3:

### Interpreting summary results from fixed-effect and random-effects meta-analysis

44

## Meta-analysis: fixed vs random

- Meta-analysts usually choose a fixed-effect or random-effects approach ...

#### Fixed-effect model:

- The true treatment effect is the same (**fixed**) in each study
- Differences across studies are only due to chance

#### Random-effects model:

- Allows the true treatment effect to be different in each study
- Allows for **heterogeneity** caused by, e.g., the types of patients, follow-up length, dose, method of measurement, ...

## Meta-analysis: fixed vs random

- Meta-analysts usually choose a fixed-effect or random-effects approach ...

#### Fixed-effect model:

- The true treatment effect is the same (**fixed**) in each study
- Differences across studies are only due to chance

#### Random-effects model:

- Allows the true treatment effect to be different in each study
- Allows for **heterogeneity** caused by, e.g., the types of patients, follow-up length, dose, method of measurement, ...

**But BOTH approaches provide a pooled treatment effect .... Should their interpretation actually be very different?**

Journal of the  
Royal Statistical Society  
*A* Statistics  
Society

J.R. Statist. Soc. A (2009)

172, Part 1, pp. 137–159

## A re-evaluation of random-effects meta-analysis

Julian P. T. Higgins, Simon G. Thompson and David J. Spiegelhalter  
Medical Research Council Biostatistics Unit, Cambridge, UK

[Received March 2007; Revised March 2008]

OnlineOpen This article is available from online at [www.biomedcentral.com](http://www.biomedcentral.com)

**Summary.** Meta-analysis in the presence of unexplained heterogeneity is frequently undertaken by using a random-effects model, in which the effects underlying different studies are assumed to be drawn from a normal distribution. Here we discuss the justification and interpretation of such models, by addressing in turn the aims of estimation, prediction and hypothesis testing. We argue that the random-effects model is appropriate for estimation, but that the assumptions of the random-effects distribution and inference on the whole distribution. We suggest that random-effects meta-analyses as currently conducted often fail to provide the key results, and we investigate the extent to which distribution-free, classical and Bayesian approaches can provide satisfactory methods. We conclude that the Bayesian approach has the advantage of naturally addressing the problem of estimation in the presence of heterogeneity. It also has no problems, including computational infeasibility and sensitivity to a priori judgements. We propose a simple prediction interval for classical meta-analysis and offer extensions to standard practice of Bayesian meta-analysis, making use of an example of studies of 'set shifting' ability in people with eating disorders.

**Keywords:** Meta-analysis; Prediction; Random-effects models; Systematic reviews

## RESEARCH METHODS & REPORTING

### Interpretation of random effects meta-analyses

Richard D Riley,<sup>1</sup> Julian P T Higgins,<sup>2</sup> Jonathan J Deeks<sup>3</sup>

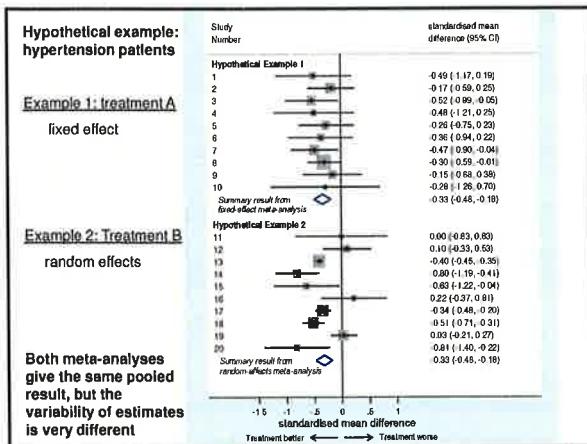
<sup>1</sup>Department of Primary Care and Population Health, University College London, London, UK  
<sup>2</sup>MRC Biostatistics Unit, University of Cambridge, Cambridge, UK  
<sup>3</sup>Department of Primary Care and Population Health, University College London, London, UK  
Accepted: 1 November 2010  
Published online: 29 January 2011  
Cite this as: BMJ 2011; 342:d549  
doi: 10.1136/bmj.d549

Summary estimates of treatment effect from random effects meta-analyses give only the average effect across all studies. Inclusion of prediction intervals, which estimate the likely effect in an individual setting, could make it easier to apply the results to clinical practice.

Meta-analysis is used to synthesise quantitative information from related studies and can form a basis for clinical practice.

Riley RD, Higgins JP, Deeks JJ.  
Interpretation of random-effects meta-analyses.  
*BMJ* 2011; 342:d549

48



## Interpretation of fixed and random-effects pooled estimate

### Fixed-effect approach

- assumes a single (fixed) treatment effect across studies
- thus pooled estimate gives the best estimate of this single treatment effect

## Interpretation of fixed and random-effects pooled estimate

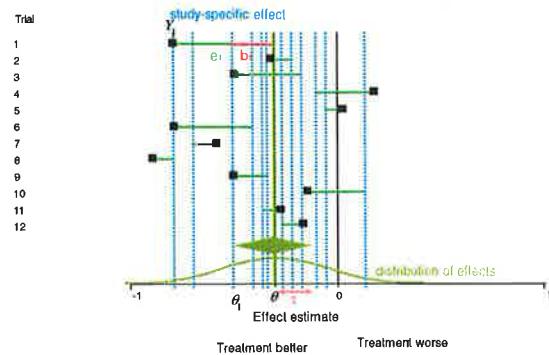
### Fixed-effect approach

- assumes a single (fixed) treatment effect across studies
- thus pooled estimate gives the best estimate of this single treatment effect

### Random-effects approach

- assumes a distribution of treatment effects across studies
- each study can have a different treatment effect
- thus the pooled estimate gives the estimate of the average treatment effect across studies
- Individual studies may have a treatment effect that varies considerably away from this average value

## Random-effects meta-analysis



### Example

- Bachmann et al. (BMJ, 2010) perform a random-effects meta-analysis of 12 randomised trials to summarise the effect of inpatient rehabilitation, compared with usual care, on improving functional outcome in geriatric patients

Pooled odds ratio = 1.36 (95% CI: 1.07 to 1.71)

- indicates that the average effect of the intervention is to make the odds of functional improvement 1.36 times higher than usual care.
- As the confidence interval is above one, it provides strong evidence that the **average** intervention effect is beneficial.

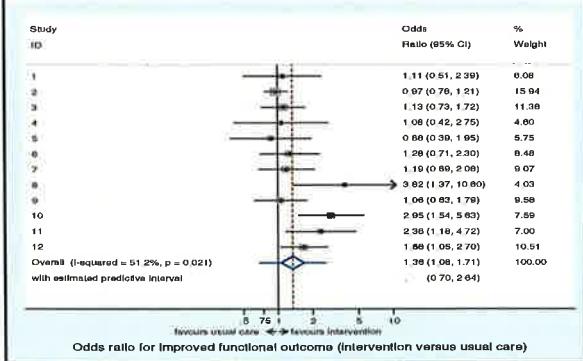
### Example

- However, there is large between-study heterogeneity ( $I^2 = 51\%$ ;  $\tau^2 = 0.27$ ), potentially due to study differences in the intervention type (e.g. general or orthopaedic rehabilitation) and length of follow-up, amongst other factors.

- Responding to the heterogeneity, the authors state:

**'pooled effects should be interpreted with caution because the true differences in effects between studies might be due to uncharacterised or unexplained underlying factors or the variability of outcome measures on functional status'**

### Bachmann: pooled odds ratio: 1.36 (95% CI: 1.07 to 1.71)



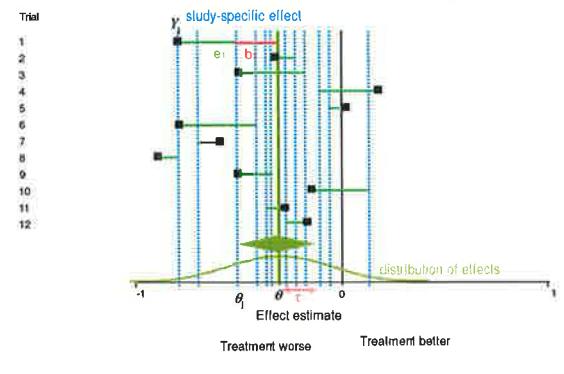
### Prediction interval

- We can quantify the variability in the intervention effect using a prediction interval (option 'rfdist' in admetan or ipdmetan)
- An approximate 95% prediction interval gives the potential intervention effect when it is applied in a single study population similar to one of those included in the meta-analysis
- Shows potential impact of treatment in actual practice
- Ideally use a Bayesian approach, but Higgins et al (JRSS-A, 2009) & Riley et al (BMJ, 2011) suggest an approximate prediction interval;

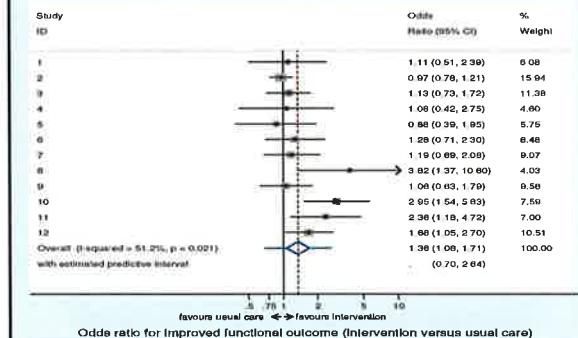
$$\hat{\theta} \pm t_{0.975,k-2} \sqrt{var(\hat{\theta}) + \hat{\tau}^2}$$

where  $k$  is the number of studies, and  $t_{0.975,k-2}$  refers to the 97.5% percentile value of the t-distribution with  $k-2$  degrees of freedom

### Random-effects meta-analysis



### Bachmann: pooled odds ratio: 1.36 (95% CI: 1.07 to 1.71) 95% prediction interval: 0.70 to 2.64



### Bachmann: pooled odds ratio: 1.36 (95% CI: 1.07 to 1.71)

95% prediction interval: 0.70 to 2.64

- On average the intervention appears effective (95% CI does not contain 1)
- But in an individual setting the intervention may not always be beneficial (prediction interval contains 1)
- Further research is needed to identify causes of the heterogeneity, e.g. the subtypes of geriatric rehabilitation programmes that work best

### Return to hypertension example:

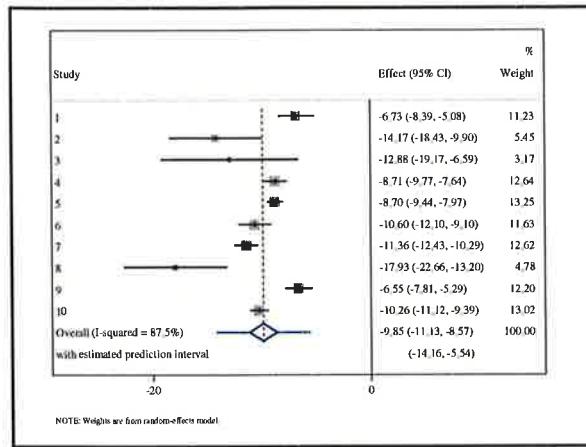
#### Random effects results (methods of moments)

Summary estimate = -9.85

95% CI for average effect = -11.13 to -8.57

95% prediction interval for effect in new study: -14.16 to -5.54

- On average anti-hypertensive drugs appear effective (95% CI does not contain 0)
- In a new population anti-hypertensive drugs are highly likely to always be beneficial (prediction interval does not contain 0  
... a Bayesian could give a probability of over 95% that treatment will be effective in a new population)



## Limitations of the Prediction Interval

- With few studies, width of interval is unhelpfully wide
- Is a normal distribution suitable (Lee and Thompson, 2008)?
- Assumes the available studies are a good representation of the settings of interest
- Width of interval susceptible to publication bias & inclusion of studies with high risk of bias
- Frequentist PI equation is only approximate (Partlett & Riley)
  - improvements are needed (work ongoing)
- Hang on! Do we even know how to estimate  $\tau^2$  & CIs yet?

62

**Research Article** Statistics in Medicine  
Received 2 February 2016 Accepted 9 September 2016 Published online in Wiley Online Library  
[wileyonlinelibrary.com/doi/10.1002/sim.7140](http://onlinelibrary.wiley.com/doi/10.1002/sim.7140)

### Random effects meta-analysis: Coverage performance of 95% confidence and prediction intervals following REML estimation

Christopher Partlett<sup>a,b,\*</sup> and Richard D. Riley<sup>c</sup>

<sup>a</sup> A random effects meta-analysis combines the results of several independent studies to compute the estimates about a particular measure of interest, such as a treatment effect. The approach allows for unexplained between-study heterogeneity in the true treatment effect by incorporating random study effects about the overall mean. The variance of the mean effect estimate is conveniently calculated by assuming that the between-study variance is known. This is appropriate when there are many studies, but becomes problematic when there are few studies. Alternative methods that also account for this uncertainty, such as Hartung-Knapp, Sidik-Jonkman and Knapp-Roger, have been proposed and shown to improve upon the traditional approach in some situations. In this paper, we compare the coverage of the 95% confidence and prediction intervals in terms of the coverage of the 95% confidence and prediction intervals derived from a random effects meta-analysis estimated using restricted maximum likelihood. We show that, in terms of the combined results, the Hartung-Knapp method performs best, followed by the Sidik-Jonkman and Knapp-Roger methods. When the between-study heterogeneity was large and/or study sizes were small, the coverage of the Hartung-Knapp method was slightly less than the heterogeneity in the  $I^2$  ( $I^2 < 30\%$ ) and study sizes were small. In contrast, the prediction interval coverage was slightly higher than the heterogeneity in the  $I^2$  ( $I^2 > 30\%$ ) and study sizes were similar. In other situations, especially where heterogeneity is small and the study sizes are quite varied, the coverage is far from one and could not be easily improved by increasing the number of studies. These findings suggest that the Hartung-Knapp method is the best choice for random effects meta-analysis. However, researchers should continue to develop 95% prediction intervals following a frequentist random effects meta-analysis until a more reliable solution is identified. © 2016 The Authors. Statistics in Medicine Published by John Wiley & Sons Ltd.

\*Keywords: random effects; meta analysis; coverage; REML; simulation

63

## Part 4:

### A note on estimation of random-effects meta-analysis models & derivation of CIs

64

### Estimation of the random-effects meta-analysis model

- Traditionally, researchers either use methods of moments (known as DerSimonian and Laird, DL) or REML
- REML assumes normality of the random effects
- DL makes no assumption about the random effects distribution
- 95% confidence intervals for the pooled (mean) effect typically derived using

$$\hat{\theta} \pm 1.96 \times s.e.(\hat{\theta})$$

**BUT !**

- This approach ignores the uncertainty of  $\hat{\tau}^2$  & the  $V(\hat{\theta}_i)$
- Concern that the traditional confidence interval is too narrow
- Increasing calls for improvement

65

**Annals of Internal Medicine | RESEARCH AND REPORTING METHODS**

### Random-Effects Meta-analysis of Inconsistent Effects: A Time for Change

John E. Cornfield, PhD; Cynthia D. Malone, MD, MSc; Russell Localio, PhD; Catherine B. Stark, PhD, MS; Anne R. Meibohm, PhD; Elise Guller, MD, DrPH; and Steven N. Goodman, MD, PhD

A primary goal of meta-analysis is to improve the estimation of treatment effects by pooling results of similar studies. This article explains how the most widely used method for pooling heterogeneous evidence, the DerSimonian and Laird (DL) estimator, can produce biased estimates with likely high precision. A classic example is presented to show that use of the DL estimator can lead to erroneous conclusions. Particular problems with the DL estimator are discussed, and several alternative methods for summarizing heterogeneous evidence are presented. The authors support replicating universal use of the DL estimator with analyses based on a critical synthesis that recognizes the uncertainty in the evidence, focuses on describing and explaining the probable sources of variation in the evidence, and uses random-effects models that provide more accurate confidence limits than the DL estimator.

Annals Intern Med 2016;164(3):249-259  
For author affiliations, see end of article.  
[www.annals.org](http://www.annals.org)

66

**Random-Effects Meta-analysis of Inconsistent Effects: A Time for Change**

John E. Cornell, PhD; Cynthia D. Mulrow, MD, MSc; Russell Localio, PhD; Catharine B. Stark, PhD, MS; Anne R. Metzbohm, PhD; Bruce Goldfarb, MD, DrPH; and Steven N. Goodman, MD, PhD

A primary goal of meta-analysis is to improve the estimation of treatment effects by pooling results of similar studies. This article explores how the most widely used method for pooling heterogeneous evidence—the DerSimonian Laird (DL) estimator—can produce biased estimates with falsely high precision. A classic example is presented to show that use of the DL estimator can lead to erroneous conclusions. Particular problems with the DL estimator are discussed, and several alternative methods for summarizing heterogeneous evidence are presented. The authors support replacing

ing universal use of the DL estimator with analyses based on a critical synthesis that recognizes the uncertainty in the evidence, focuses on describing and explaining the probable sources of variation in the evidence, and uses random-effects estimates that provide more accurate confidence limits than the DL estimator.

*Ann Intern Med.* 2014;160:267-270  
For author affiliations, see end of first page.  
[www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)

67

**Random-Effects Meta-analysis of Inconsistent Effects: A Time for Change**

John E. Cornell, PhD; Cynthia D. Mulrow, MD, MSc; Russell Localio, PhD; Catharine B. Stark, PhD, MS; Anne R. Metzbohm, PhD; Bruce Goldfarb, MD, DrPH; and Steven N. Goodman, MD, PhD

A primary goal of meta-analysis is to improve the estimation of treatment effects by pooling results of similar studies. This article explores how the most widely used method for pooling heterogeneous evidence—the DerSimonian Laird (DL) estimator—can produce biased estimates with falsely high precision. A classic example is presented to show that use of the DL estimator can lead to erroneous conclusions. Particular problems with the DL estimator are discussed, and several alternative methods for summarizing heterogeneous evidence are presented. The authors support replacing

ing universal use of the DL estimator with analyses based on a critical synthesis that recognizes the uncertainty in the evidence, focuses on describing and explaining the probable sources of variation in the evidence, and uses random effects estimates that provide more accurate confidence limits than the DL estimator.

*Ann Intern Med.* 2014;160:267-270  
For author affiliations, see end of first page.  
[www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)

68

**Which estimation method should I use for a random effects meta-analysis?****Main options in software:**

- DL
- ML
- REML

**But there are many other competitors**

- Paule and Mandel
- Sidik and Jonkman
- DerSimonian and Kacker 2-step methods
- Bayesian methods, ...

69

**Options for deriving CIs following a random-effects model****To account for uncertainty in  $\hat{\tau}^2$ :**

- Bayesian approach
- ML with profile likelihood (Hardy and Thompson, 1996)
- DL or REML followed by:
  - a (modified) Hartung-Knapp-Sidik-Jonkman correction of the s.e. ( $\hat{\sigma}$ )
  - use of value from  $t$ -distribution rather than 1.96

$$\hat{\mu} \pm t_{k-1,0.975} \times s.e._{HKSJ}(\hat{\mu})$$

- Kenward-Roger, **plus others...**

71

**Research Synthesis Methods**Invited Review  
Published online in Wiley Online Library  
Received 20 July 2011; Accepted 20 June 2012  
DOI: 10.1002/sim.4646**Methods to estimate the between-study variance and its uncertainty in meta-analysis**

Areti Angeliki Veroniki,<sup>a,\*</sup> Dan Jackson,<sup>b</sup>  
Wolfgang Viechtbauer<sup>c</sup>  
"We identified 16 estimators ...  
many approaches have not  
been compared under the same simulation  
settings, and hence making any clear  
recommendations about these  
methods is difficult."

Keywords: heterogeneity, mean square error, confidence interval

70

**A refined method for the meta-analysis of controlled clinical trials with binary outcome<sup>†</sup>**Joachim Hartung<sup>a,1§</sup> and Guido Knapp<sup>b</sup>

Department of Statistics, University of Dortmund, D-44221 Dortmund, Germany

**SUMMARY**

For the meta-analysis of controlled clinical trials with binary outcome a test statistic for testing an overall treatment effect is proposed, which is based on a refined estimator for the variance of the treatment effect estimator usually used in the random-effects model of meta-analysis. In simulation studies it is shown that the proposed test keeps the prescribed significance level much better than the commonly used tests in the fixed-effects and random-effects model, respectively. Moreover, when using the test it is not necessary to choose between fixed effects and random effects approaches in advance. The proposed method applies in the same way to the analysis of a controlled multi-centre study with binary outcome, including a possible interaction between drugs and centres. Copyright © 2001 John Wiley & Sons, Ltd.

72

STATISTICS IN MEDICINE  
*Statist. Med.* 2001; **20**:1791–1792 (DOI: 10.1002/sim.791)

## On tests of the overall treatment effect in meta-analysis with normally distributed responses

Joachim Hartung<sup>\*†‡</sup> and Guido Knapp<sup>†</sup>

*Department of Statistics, University of Dortmund, D-44221 Dortmund, Germany*

### SUMMARY

For the meta-analysis of controlled clinical trials or epidemiological studies, in which the responses are at least approximately normally distributed, a refined test for the hypothesis of no overall treatment effect is proposed. The test statistic is based on a direct estimation function for the variance of the overall treatment effect estimator. As outcome measures, the absolute and the standardized difference between means are considered. In simulation studies it is shown that the proposed test keeps the prescribed significance level very well in contrast to the commonly used tests in the fixed effects and random effects model, respectively, which can become very liberal. Furthermore, just for using the proposed test it is not necessary to choose between the fixed effects and the random effects approach in advance.

Copyright © 2001 John Wiley & Sons, Ltd.

73

Inoue et al. *BMC Medical Research Methodology* 2014, **14**:25  
http://dx.doi.org/10.1186/1471-2963-14-25

**RESEARCH ARTICLE** **Open Access**

## The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method

John Hartung<sup>1</sup>, John P. Kleinmuntz<sup>2,3\*</sup> and George V. Kerec<sup>1</sup>

**Abstract**  
Background: The DerSimonian and Laird (DSL) method is widely used for the random effects meta-analysis of continuous data type in meta-analyses. The method described by Hartung, Knapp and Sidik and Jonkman (HKSJ) is known to perform better when trial variances are considered. However, evidence in realistic scenarios, where true and null effects are mixed together, has been scarce, so far. The aim here is to clarify the relative performance of the two methods in terms of Type I error rate and power in a range of realistic scenarios. We also compare the HKSJ method to the Hartung-Knapp-Sidik-Jonkman (HKJS) method, which is a modification of the HKSJ method that uses a different way to estimate the variance of the overall mean.

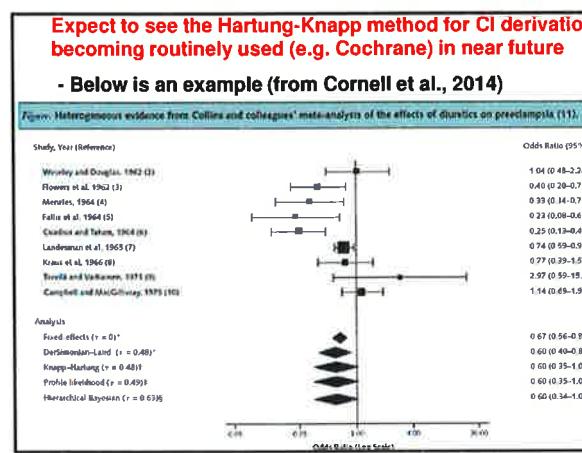
Methods: We evaluate the HKSJ and HKJS methods in terms of the Type I error rate for sample sizes of  $n = 20$  and  $n = 100$  and for various numbers of studies and between study heterogeneity, and also for known values of  $\sigma^2_{\text{true}}$  of the total (true) error, larger than the sample size. We also compare the power of the HKSJ and HKJS methods against the DSL method, using a prior distribution of trial variances and  $n = 3$  studies of intervention sizes  $(n_i)$  drawn from a uniform distribution.

Results: The results show that the HKSJ method considerably reduces the mean adaptive error rate when the true total error is larger than the sample size, if the HKSJ error rates are more stable, whereas the DSL method is less stable. The HKJS method shows similar performance to the HKSJ method when the total error is smaller than the sample size. The HKJS method is slightly more powerful than the HKSJ method when the total error is larger than the sample size, but the HKSJ method is more powerful when the total error is smaller than the sample size.

Conclusion: Our results show that the HKSJ method is considerably more robust than the DSL method, especially when the total error is larger than the sample size, and can easily be applied to real-life meta-analyses. Even with the HKJS method, however, care needs to be taken when there are a lot of very unequal sizes of studies.

**Keywords:** Meta-analysis, Clinical trials, Heterogeneity, Type I error, Random effects, Guller's distribution of the sample mean

74



## Part 5:

### Combining IPD and non-IPD studies

75

## Combining IPD and non-IPD studies

- Sometimes IPD is not available from all studies
- Review of Nevitt et al. (2017, BMJ):  

"Only 188 (25%) of these IPD meta-analyses retrieved 100% of the eligible IPD for analysis, with 324 (43%) of these IPD meta-analyses retrieving 80% or more of relevant IPD"
- Stewart and Tierney (2002) advise that:  

"... an IPD-only meta-analysis may be biased if unavailability of IPD is related to the study results"
- Ahmed et al. (2013, BMJ) also demonstrate 'availability bias' concern
- However, non-IPD studies may also provide AD of  $\hat{\theta}_i$  &  $V(\hat{\theta}_i)$
- Easy to combine IPD and AD studies in step 2 of two-step approach (though issue of how AD studies estimated may often remain)

77

**BMJ**  
doi:10.1136/bmjjournals.0.970365. First published online 3 March 2017. Page 1 of 14

**RESEARCH**

## Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey

**OPEN ACCESS**

*Authors' affiliations:* Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, USA; Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, USA; Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, USA

**Abstract**  
OBJECTIVE To assess the availability of individual participant data (IPD) and associated data in meta-analyses using IPD. DESIGN A systematic review of databases of meta-analyses using IPD. SETTING PubMed, Google Scholar, and the Cochrane Library. POPULATION All meta-analyses using IPD. METHODS We identified 100 meta-analyses using IPD. We assessed the availability of IPD and associated data in each study. MAIN OUTCOME MEASURES The percentage of studies that had IPD available, the percentage of studies that had associated data available, and the percentage of studies that had both IPD and associated data available.

**RESULTS** Of the 100 meta-analyses, 100% had IPD available, 100% had associated data available, and 100% had both IPD and associated data available. The percentage of studies that had IPD available ranged from 0% to 100% across the 100 meta-analyses. The percentage of studies that had associated data available ranged from 0% to 100% across the 100 meta-analyses. The percentage of studies that had both IPD and associated data available ranged from 0% to 100% across the 100 meta-analyses.

**CONCLUSION** Most meta-analyses using IPD have IPD and associated data available. However, a significant number of studies have IPD available but do not have associated data available.

**DATA SOURCE** PubMed, Google Scholar, and the Cochrane Library.

78

### Applied Example (Riley et al., 2008)

- 10 RCTs; each assess treatment effect on SBP
- IPD available for each trial
- Two-step approach using previous models & REML gives:

Parameter estimate	100% IPD trials only
Pooled treatment effect, $\hat{\theta}$	-10.08
Standard error of $\hat{\theta}$	0.91
Between-study variance, $\tau^2$	6.53

### Applied Example (Riley et al., 2008)

- 10 RCTs; each assess treatment effect on SBP
- What if IPD were not available from all studies?
- Two-step approach now gives:

Parameter estimate	100% IPD trials only	90% IPD trials only	50% IPD trials only	10% IPD trials only	0% AD trials only
Pooled treatment effect, $\hat{\theta}$	-10.08	-10.46	-10.96	-10.18	
Standard error of $\hat{\theta}$	0.91	0.86	1.51	0.46	
Between-study variance, $\tau^2$	6.53	5.10	10.20		

### Applied Example (Riley et al., 2008)

- 10 RCTs; each assess treatment effect on SBP
- What if IPD were not available from all studies?
- Two-step approach including non-IPD studies now gives:

Parameter estimate	100% IPD trials only	90% IPD trials only	IPD & AD trials only	50% IPD trials only	10% IPD trials only	0% AD trials only
Pooled treatment effect, $\hat{\theta}$	-10.08	-10.46	-10.08	-10.96	-10.18	
Standard error of $\hat{\theta}$	0.91	0.86	0.91	1.51	0.46	
Between-study variance, $\tau^2$	6.53	5.10	6.53	10.20		

### Applied Example (Riley et al., 2008)

- 10 RCTs; each assess treatment effect on SBP
- What if IPD were not available from all studies?
- Two-step approach including non-IPD studies now gives:

NB results averaged across all permutations of which studies are missing IPD

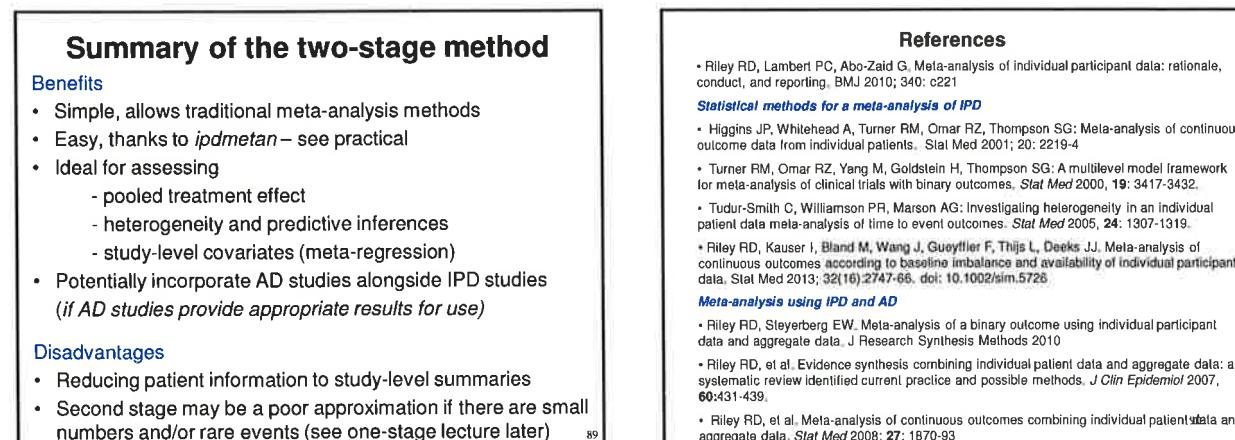
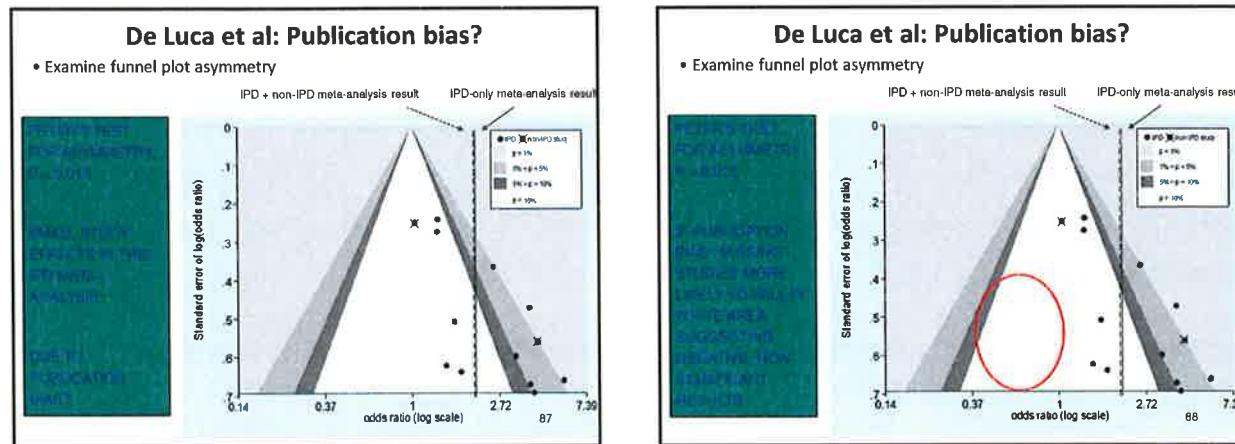
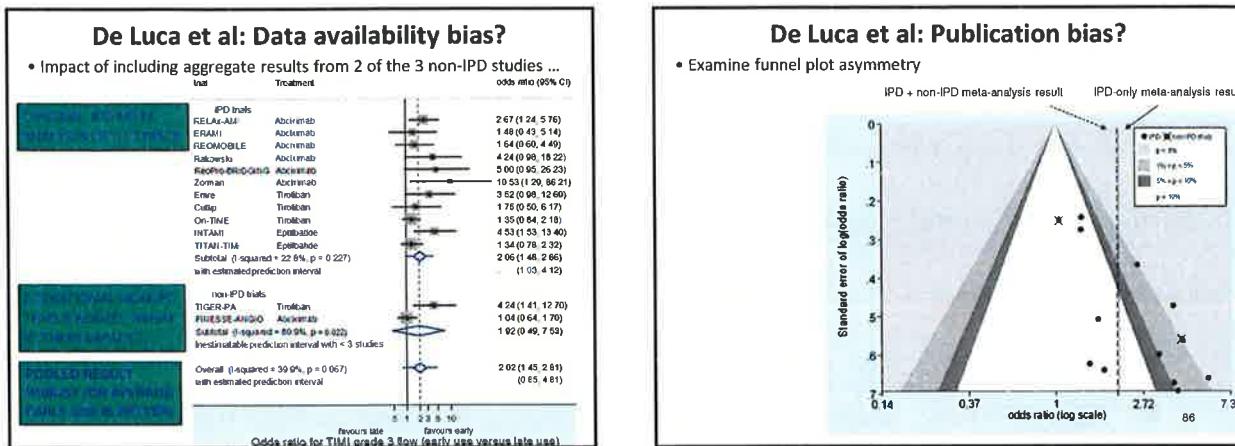
Parameter estimate	100% IPD trials only	IPD & AD trials only	90% IPD & AD trials only	IPD trials only	50% IPD & AD trials only	IPD trials only	10% IPD & AD trials only	IPD trials only	0% AD trials only
Pooled treatment effect, $\hat{\theta}$	-10.08	-10.46	-10.08	-10.96	-10.08	-10.96	-10.08	-10.18	-10.08
Standard error of $\hat{\theta}$	0.91	0.86	0.91	1.51	0.91	0.91	0.46	0.91	0.91
Between-study variance, $\tau^2$	6.53	5.10	6.53	10.20	6.53	10.20	6.53	6.53	6.53

### Issues for combining IPD and AD

- Combining IPD & aggregate data is appealing to use all evidence
- Sensitivity analyses including/excluding non-IPD studies thus recommended
- But may just reinforce why IPD was sought! Non-IPD studies may:
  - not provide information required;
    - e.g. missing standard error of effect estimate
  - not have used correct statistical model
    - e.g. may not have adjusted for baseline imbalances
  - not be of comparable quality
    - e.g. maybe only the 'best' quality studies provide IPD
  - increase the observed heterogeneity
    - e.g. IPD studies adjust for baseline SBP; AD studies may not

### Example: De Luca et al. review

- Review the benefits of early versus late use of Gp IIb-IIIa inhibitors in patients undergoing primary angioplasty for ST-segment elevation myocardial infarction.
- A primary angiographic endpoint was whether patients achieved a preprocedural Thrombolysis in Myocardial Infarction Study (TIMI) grade 3 flow distal embolisation.
- A systematic review identified 14 relevant trials and IPD was sought from them all, so selection bias is not a concern.
- However, availability and publication biases are a threat, as
  - IPD was unavailable for 3 trials (21%)
  - all 11 trials providing IPD were fully published.
- De Luca did not investigate these biases, but Ahmed et al did...  
<sup>84</sup>



## **PRACTICAL 1: Two-stage IPD meta-analysis**

Please follow the instructions in the Stata do files labelled Practical 1a, 1b and 1c

Practical 1(a): Two-stage binary outcomes

Practical 1(b): Two-stage continuous outcomes

Practical 1(c): Two-stage survival outcomes

### **Learning objectives:**

- Gain experience of fitting two-stage IPD meta-analysis models.
- Analyse binary, continuous and survival data using IPD from multiple studies.
- Interpret Stata output after fitting IPD meta-analysis models.
- Gain experience of using key Stata modules for meta-analysis including ‘metan’, ‘metareg’, ‘metaan’, and ‘ipdmetan’.

## **LECTURE 3: One-stage IPD meta-analysis**

### **Summary:**

The statistical approach to IPD meta-analysis involves a two-stage or a one-stage approach. In the one-stage approach, the IPD from all studies are modelled simultaneously whilst accounting for the clustering of participants within studies. This requires a model specific to the type of data being synthesised, alongside appropriate specification of the meta-analysis assumptions (e.g. fixed or random effects across studies). The one-stage approach is more flexible and potentially more exact than the two-stage approach, but may face computational difficulties. In this lecture, we outline how to perform a one-stage IPD meta-analysis for continuous, binary and survival outcomes, and describe the similarities and differences between the one-stage and two-stage approaches.

### **Learning objectives:**

- Understand the meaning of a ‘one-stage’ approach to IPD meta-analysis
- Know why it is inappropriate to lump all the IPD together and analyses ignoring the clustering of patients within studies
- Understand how to perform a one-stage IPD meta-analysis for continuous, binary and survival outcomes
- Understand the advantages and disadvantages of a one-stage approach compared to a two-stage approach
- Understand how to combine IPD and non-IPD studies within the one-stage approach

### **Key references:**

1. Higgins JP, Whitehead A, Turner RM, Omar RZ, Thompson SG: Meta-analysis of continuous outcome data from individual patients. *Stat Med* 2001; 20: 2219-4
2. Turner RM, Omar RZ, Yang M, et al: A multilevel model framework for meta-analysis of clinical trials with binary outcomes. *Stat Med* 2000, 19: 3417-3432.
3. Tudur-Smith C, Williamson PR, Marson AG: Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. *Stat Med* 2005, 24: 1307-1319.
4. Riley RD, Kauser I, Bland M, Wang J, Gueyffier F, Thijs L, Deeks JJ. Meta-analysis of continuous outcomes according to baseline imbalance and availability of individual participant data. *Stat Med* 2013; 32(16):2747-66. doi: 10.1002/sim.5726
5. Mathew T, Nordstrom K. On the equivalence of meta-analysis using literature and using individual patient data. *Biometrics* 1999;55(4):1221-3
6. Debray TPA, Moons KGM, Abo-Zaid GMA, Koffijberg H, Riley RD (2013) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? *PLoS ONE* 8(4): e60650
7. Debray TP, Moons KG, van Valkenhoef G, et al. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. *Res Synth Methods*. 2015.



## LECTURE 3: One-stage IPD meta-analysis

**Dr Danielle Burke**  
Research Associate  
Keele University  
email: [d.burke@keele.ac.uk](mailto:d.burke@keele.ac.uk)

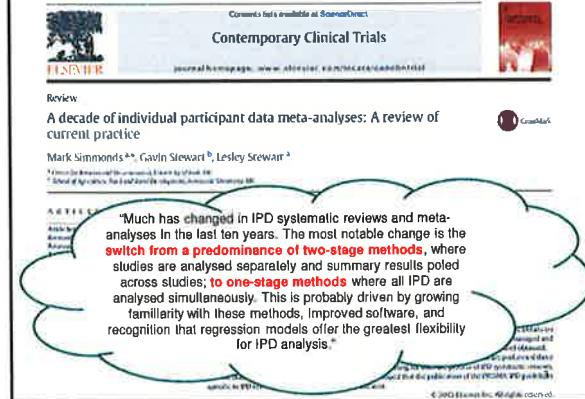
**Prof Richard Riley**  
Professor of Biostatistics  
Keele University  
e-mail: [r.riley@keele.ac.uk](mailto:r.riley@keele.ac.uk)

## Aims

- Introduce one-stage approach
- Must not ignore the clustering of participants by trial
- Demonstrate possible fixed effect and random effects models
- Comparison of one-stage and two-stage approaches
- Reporting considerations

2

### One-stage approach becoming more common ...



## Part 1:

### One-stage analysis ignoring clustering (how not to do it!)

4

## One-stage IPD meta-analysis

- The one-stage approach meta-analyses the IPD from all studies simultaneously, to produce summary results
- Requires the IPD from all studies to be in one dataset
- **But** crucial not to lose the clustering of patients within studies

Researchers may:

- Ignore clustering when merging datasets
- Or, ignore clustering in their analysis

5

## Example – merging datasets

Study	Patient	SBP Initial	SBP final	treat	placebo	age	sex
1	1	190	185	1	0	58	1
1	2	175	172	1	0	69	1
1	3	184	185	0	1	39	0
1	4	192	182	0	1	45	1
2	1	201	199	1	0	51	0
2	2	169	154	1	0	42	1
2	3	171	170	0	1	50	1
2	4	179	168	0	1	67	0
3	1	197	167	1	0	83	1
3	2	189	171	1	0	78	0
3	3	184	188	0	1	55	1
3	4	168	161	0	1	61	0



6

## Example – merging datasets

Patient	SBP initial	SBP final	treat	placebo	age	sex
1	190	185	1	0	58	1
2	175	172	1	0	69	1
3	184	185	0	1	39	0
4	192	182	0	1	45	1
1	201	199	1	0	51	0
2	169	154	1	0	42	1
3	171	170	0	1	50	1
4	179	168	0	1	67	0
1	197	167	1	0	83	1
2	189	171	1	0	78	0
3	184	188	0	1	55	1
4	168	161	0	1	61	0



7

## One-stage IPD meta-analysis ignoring clustering

- Simmonds et al. (2005) examined IPD meta-analyses of randomised trials
  - 3 out of 14 ignored clustering.
- Abo-Zaid et al. (2012) examined IPD meta-analyses of prognostic factor studies (see lecture 6)
  - 5 out of 11 ignored clustering

8

## Example: ANCOVA ignoring clustering

### Continuous outcome – ANCOVA model

- Treatment effect on systolic blood pressure (SBP)
- $i = 1$  to  $k$  trials

9

## Example: ANCOVA ignoring clustering

### Continuous outcome – ANCOVA model

- Treatment effect on systolic blood pressure (SBP)
- $i = 1$  to  $k$  trials

$$SBP_j = \phi + \beta_1 SBP_{0j} + \theta x_j + \varepsilon_j \quad \varepsilon_j \sim N(0, \sigma^2)$$

Final SBP      Adjust for baseline SBP  
      Control effect      Treatment effect      Residual error

- Ignoring clustering ignores the trial strata, and so no  $\beta_i$  term in the regression model

10

## Ignoring clustering – why not?

- Naively assumes all patients came from one big study.
- Combining loses the original randomisation in each study
- Patients in the same study are likely to be similar:
  - From same population
  - Subject to same treatment strategies (dosing, timing, methods, etc)
  - Subject to same doctors, health professionals, support, etc
- Thus correlated responses for those in same study
- Similar to repeated measures models, the correlation should be accounted for.

11

## Ignoring clustering – why not?

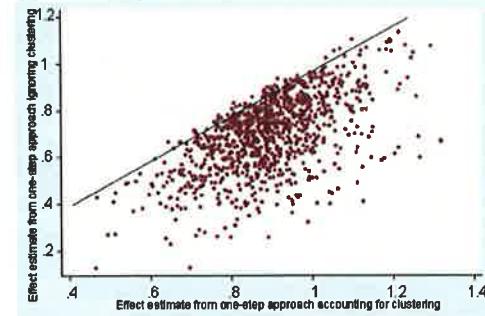
- Assumes baseline risk in all trials was the same – may cause misleading results
- However, each trial should be allowed to have their own baseline risk
- Basically, the naïve approach makes very strong assumptions, that - if wrong - lead to low coverage and potentially biased summary results
- Abo-Zaid et al. (2013) investigated this.

12



## Downward bias in effect estimate by ignoring clustering

### (a) summary log odds ratio

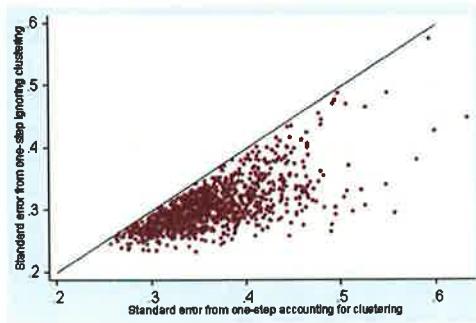


(Abo-Zaid et al., 2013)

14

## Poor coverage by ignoring clustering

### (a) s.e. of summary log odds ratio



(Abo-Zaid et al., 2013)

15

## Applied example

**Table 1:** Results for the effect of a family history of thrombophilia on the odds of truly having deep vein thrombosis, for each of the three IPD models.

Method	Pooled logOR (s.e.)	Odds ratio	95% CI for odds ratio	p-value
Two-step	0.280 (0.135)	1.323	1.015 to 1.725	0.038
One-step ignoring clustering	0.060 (0.128)	1.062	0.825 to 1.365	0.642
One-step accounting for clustering	0.269 (0.138)	1.301	0.996 to 1.697	0.053

16

## Well known statistically already...

- Omission of an important covariate (i.e. study effect) can bias results:

Robinson LD, Jewell NP. Some surprising results about covariate adjustment in logistic regression models. *Int Stat Rev* 1991; 58: 227-240.

Greenland S, Robins MR, Pearl J. Confounding and collapsibility in causal inference. *Statistical Science* 1999; 14: 29-46

Gail MH, Wieand S, Piantadosi S. Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. *Biometrika* 1984; 71: 431-444.

Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959; 22: 719-748.

Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Prog Cardiovasc Dis* 1985; 17: 335-371.

## Part 2:

## One-stage analysis accounting for clustering (how to do it!)

16

## Accounting for clustering

- Allow each trial to have their own baseline risk. Either:
    - stratify by study (separate intercept term for each study)
  - or
    - place a random effect on the intercept term, to allow for heterogeneity in baseline risk across studies
  - The latter makes strong & perhaps unnecessary assumptions:
    - ... trials are drawn from a population of all possible trials
    - ... assumes a (normal) distribution of baseline risks
  - But it may substantially reduce the number of parameters to estimate (potentially important if ML estimation used)
  - The latter is also preferable when developing prognostic models, as then need an average intercept term.

19

## One-stage approach

#### **Continuous outcomes: ANCOVA (fixed effect model):**

- Treatment effect on systolic blood pressure (SBP)
  - $i = 1$  to  $k$  trials:
  - The  $j^{\text{th}}$  patient provides their SBP after treatment,  $\text{SBP}_{ij}$  and their SBP at baseline  $\text{SBP}_{0ij}$
  - Treatment group ( $x_{ij} = 1$ ) and Control group ( $x_{ij} = 0$ )
  - Fit:

$$SBP_{ij} = \varphi_i + \beta_{1i} SBP_{0ij} + \theta x_{ij} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N(0, \sigma_i^2)$$

Gives summary mean difference ( $\hat{\theta}$ ) & its variance ( $V(\hat{\theta})$ )

**Continuous outcomes: ANCOVA (fixed effect model)**  
**Stata code and output:**

— 3 —

**reg sbspi i.trialdummy c.sbspi#trialdummy treat , nocons**

	step	Coef.	Std. Err.	t	t	[95% Conf. Interval]
<b>trialdummy</b>						
3	96.9176	30.5325	3.11	9.01	0.000	76.296811 <b>117.3484</b>
4	30.63773	19.4063	1.59	0.047	0.0438970 <b>60.03157</b>	
5	80.1196	1.935649	41.36	0.000	76.98272 <b>84.44112</b>	
6	181.4567	2.580562	39.71	0.000	77.39462 <b>100.00000</b>	
7	114.0456	2.580562	44.30	0.000	77.39462 <b>100.00000</b>	
8	95.8486	4.610202	21.11	0.001	45.83 <b>43.96052</b>	
9	86.1218	17.7682	4.85	0.000	51.29431 <b>120.9485</b>	
10	89.3662	5.464507	16.35	0.000	76.63554 <b>100.05659</b>	
	80.66918	4.47952	13.05	0.000	49.3169 <b>66.76667</b>	
<b>trialdummy*age</b>						
1	.9170028	.000972	308.56	0.000	.9132045 <b>.922855</b>	
2	.4221751	.0149284	7.69	0.000	.3147144 <b>.5100935</b>	
3	.774647	.063325	8.57	0.000	.04608313 <b>.9349265</b>	
4	.384427	.0127975	30.06	0.000	.3503520 <b>.4079487</b>	
5	.0832283	.0130084	31.00	0.000	.0272283 <b>.0300000</b>	
6	.2410294	.0130084	18.46	0.000	.0109351 <b>.0269072</b>	
7	.5804637	.0210268	21.81	0.000	.3362696 <b>.6245198</b>	
8	.5075668	.0123263	5.57	0.000	.2390577 <b>.6697579</b>	
9	.403303	.0121975	12.61	0.000	.0306416 <b>.0660046</b>	
10	.6135511	.0253522	24.03	0.000	.5634866 <b>.6357395</b>	
CONST	-9.33591	.2061454	-45.28	0.000	-9.73936 <b>-8.931930</b>	

2

**Continuous outcomes: ANCOVA (fixed effect model)**  
**Stata code and output:**

BUT!! The previous model assu

Using the mixed command, we can relax this assumption

```
. mixed sbpl i.trialdummy c.sbp1#trialdummy treat, nocons  
      reml residuals(ind, by(trialdummy))
```

2

## One-stage approach

### **Continuous outcomes: ANCOVA (random effects model):**

- Treatment effect on systolic blood pressure (SBP)
  - $i = 1$  to  $k$  trials:
  - The  $j^{\text{th}}$  patient provides their SBP after treatment,  $\text{SBP}_{ij}$  and their SBP at baseline  $\text{SBP}_{0ij}$
  - Treatment group ( $x_{ij} = 1$ ) and Control group ( $x_{ij} = 0$ )

$$SBP_{ij} = \phi_i + \beta_{1i} SBP_{0ij} + \theta_i x_{ij} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N(0, \sigma^2) \\ \theta_i = \theta + u \quad u \sim N(0, \tau^2)$$

Gives summary mean difference ( $\hat{\theta}$ ) & its variance ( $V(\hat{\theta})$ )

23

## **Continuous outcomes: ANCOVA (random effects model) Stata code and output:**

```
* mixed sbpl sbp1 treat triall trial2 trial3 trial4 trial5 trial6 trial7 ///
trial8 trial9 trial10 , nocons || trialdummy: treat, noconstant reml
```

tbl1	Chdf	Std. Err.	n	P> t	(95% Conf. Interval)
sbp1	-41.87276	0.073698	56.42	0.000	-40.24829 -43.31703
sbp2	-10.31857	0.193239	30.12	0.000	-12.16423 -8.30916
trial1	5.82685	1.263071	58.95	0.000	73.30573 76.34796
trial2	9.67339	1.263071	54.45	0.000	70.44243 70.76443
trial3	9.43535	1.256204	50.76	0.000	70.96379 70.18379
trial4	5.93368	1.176272	44.55	0.000	73.05937 77.62639
trial5	.16-.67993	1.190873	44.07	0.000	76.34586 81.014
trial6	30.7065	1.214248	62.13	0.000	65.52161 91.09345
trial7	8.41161	1.214248	58.95	0.000	62.46081 75.11458
trial8	10.65264	1.199551	53.44	0.000	62.46081 75.11458
trial9	15.36896	1.352837	63.10	0.000	62.71345 88.02947
trial10	92.40403	1.320945	69.43	0.000	89.79572 95.02598

Random-Effects Parameters	Estimate	Std. Err.	(95% Conf. Interval)
Intercept	17.27844	0.031725	[17.13561, 17.62245]
trialday	2.4001825	.0209193	[1.641756, 5.473969]
adjtrial			

2

### Continuous outcomes: ANCOVA (random effects model)

#### Stata code and output:

```
* mixed sbpl sbpi treat trial1 trial2 trial3 trial4 trial5 trial6 trial7 ///
trial8 trial9 trial10 , nocons || trialdummy: treat, noconstant reml
```

sbpl	Coeff.	Std. Err.	t	P> t	[95% Conf. Interval]
sbpi	-4187276	.0073685	56.62	0.000	-4042849 -4331703
treat	-10.31857	1.039233	-10.12	0.000	-12.31623 -8.320916
trial1	38.82649	1.094203	58.95	0.000	73.30573 70.34796
trial2	99.26730	1.622966	51.45	0.000	95.69443 104.0403
trial3	94.99955	2.142404	44.34	0.000	99.7895 99.18757
trial4	81.10000	1.343122	64.55	0.000	73.05097 77.62639
trial5	78.67903	1.181200	66.01	0.000	76.10000 76.10000
trial6	88.30765	1.212346	65.13	0.000	85.52186 91.03715
trial7	84.86977	1.391150	65.26	0.000	82.43713 87.54226
trial8	104.5368	1.939528	53.44	0.000	102.6294 110.4436
trial9	65.36696	1.352837	63.10	0.000	62.71745 66.02047
trial10	92.40432	1.330845	69.43	0.000	89.79572 95.01293

Random-effects Parameters		Estimate	Std. Err.	[95% Conf. Interval]
trialdummy: Identity	sd(treat)	2.957825	.9209619	1.641756 5.473989
sd(individual)		17.27844	.0731735	17.12951 17.42245

LR test vs. linear model: chibar2(01) = 47.60 Prob > chibar2 = 0.0000

25

### Continuous outcomes: ANCOVA (random effects model)

#### Stata code and output:

```
* mixed sbpl sbpi treat trial1 trial2 trial3 trial4 trial5 trial6 trial7 ///
trial8 trial9 trial10 , nocons || trialdummy: treat, noconstant reml
```

sbpl	Coeff.	Std. Err.	t	P> t	[95% Conf. Interval]
sbpi	-4187276	.0073685	56.62	0.000	-4042849 4331703
treat	-10.31857	1.039233	-10.12	0.000	-12.31623 -8.320916
trial1	75.82685	1.268507	58.95	0.000	73.30573 70.34796
trial2	99.26730	1.622966	51.45	0.000	95.69443 104.0403
trial3	94.99955	2.142404	44.34	0.000	99.7895 99.18757
trial4	81.10000	1.343122	64.55	0.000	73.05097 77.62639
trial5	78.67903	1.181200	66.01	0.000	76.10000 76.10000
trial6	88.30765	1.212346	65.13	0.000	85.52186 91.03715
trial7	84.86977	1.391150	65.26	0.000	82.43713 87.54226
trial8	104.5368	1.939528	53.44	0.000	102.6294 110.4436
trial9	65.36696	1.352837	63.10	0.000	62.71745 66.02047
trial10	92.40432	1.330845	69.43	0.000	89.79572 95.01293

What is wrong with this model? trial10

26

### Continuous outcomes: ANCOVA (random effects model)

#### Stata code and output:

```
* mixed sbpl sbpi treat trial1 trial2 trial3 trial4 trial5 trial6 trial7 ///
trial8 trial9 trial10 , nocons || trialdummy: treat, noconstant reml
```

sbpl	Coeff.	Std. Err.	t	P> t	[95% Conf. Interval]
sbpi	-4187276	.0073685	56.62	0.000	-4042849 -4331703
treat	-10.31857	1.039233	-10.12	0.000	-12.31623 -8.320916
trial1	75.82685	1.268507	58.95	0.000	73.30573 70.34796
trial2	99.26730	1.622966	54.45	0.000	95.69443 102.9403
trial3	94.99955	2.142404	44.34	0.000	99.7895 99.18757
trial4	81.10000	1.343122	64.55	0.000	73.05097 77.62639
trial5	78.67903	1.181200	66.01	0.000	76.10000 76.10000
trial6	88.30765	1.212346	65.13	0.000	85.52186 91.03715
trial7	84.86977	1.391150	65.26	0.000	82.43713 87.54226
trial8	106.5368	1.939529	53.44	0.000	102.6294 110.4436
trial9	65.36696	1.352837	63.10	0.000	62.71745 66.02047
trial10	92.40432	1.330945	69.43	0.000	89.79572 95.01293

Caution:  
These are strong assumptions

Random-effects Parameters		Estimate	Std. Err.	[95% Conf. Interval]
trialdummy: Identity	sd(treat)	2.957825	.9209619	1.641756 5.473989
sd(individual)		17.27844	.0731735	17.12951 17.42245

LR test vs. linear model: chibar2(01) = 47.60 Prob > chibar2 = 0.0000

27

### Continuous outcomes: ANCOVA (random effects model)

#### Stata code and output:

```
* mixed sbpl i.trialdummy c.sbpi#trialdummy treat ,  
|| trialdummy: treat, nocons reml residuals(ind,  
by(trialdummy))
```

This will give us a separate adjustment for baseline SBP and a separate residual variance per study

But will it fit in Stata? See practical ...

Could alternatively put a random effect on SBPi  
But again, will it fit?

```
* mixed sbpl i.trialdummy c.sbpi#trialdummy treat ,  
|| trialdummy: treat, nocons reml residuals(ind,  
by(trialdummy))
```

28

### One-stage approach

#### Binary outcomes: logistic regression (fixed effect model)

- $i = 1 \text{ to } k$  trials
- The  $j^{\text{th}}$  patient provides their outcome response (0 or 1)
- $p_{ij}$  is the probability of an event
- Treatment group ( $x_{ij} = 1$ ) and Control group ( $x_{ij} = 0$ )
- Fit:

$$\ln(p_{ij}/(1-p_{ij})) = \phi_i + \beta_{1i}SBP_{0ij} + \theta x_{ij}$$

Gives summary log odds ratio ( $\hat{\theta}$ ) & its variance ( $V(\hat{\theta})$ )

29

#### Binary outcomes: logistic regression (fixed effect model, stratified intercept)

##### Example code for a treatment called 'eryt'

```
* logistic dvt eryt study1 study2 study3, nocons
```

Logistic regression		Number of obs	Wald chi2(4)	Prob > chi2
		= 1,722	= 620.82	= 0.0000
Log likelihood = -747.08689				

dvt	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
eryt	1.351515	.1864085	2.18	0.029	1.03126 1.771225
study1	+1295397	.0143114	-18.50	0.000	.104319 .160858
study2	+18393	.0407957	-7.63	0.000	.+1190847 .2840857
study3	.2646055	.029437	-11.95	0.000	.+2127668 .3290741

30

### Binary outcomes: logistic regression (fixed effect model, random intercept)

```
+ meqrlogit dvt eryt, nocons || studyid: , var or
Mixed-effects logistic regression
Group variable: studyid
Number of obs = 1,722
Number of groups = 3
Obs per group:
min = 153
avg = 574.0
max = 1,028
Integration points = 7
Log likelihood = -756.37964
Wald chi2(1) = 4.40
Prob > chi2 = 0.036

```

dvt	Odds Ratio	Std. Err.	Z	P> z	[95% Conf. Interval]
eryt	1.335757	.184434	2.10	0.036	1.019057 1.750881

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]
studyid: Identity	2.925121	.2.412149	.581053 14.72557
var(_cons)			

### One-stage approach

#### Binary outcomes: logistic regression (random effects model, stratified intercept)

- $i = 1 \text{ to } k$  trials
- The  $j^{\text{th}}$  patient provides their outcome response (0 or 1)
- $p_{ij}$  is the probability of an event
- Treatment group ( $x_{ij} = 1$ ) and Control group ( $x_{ij} = 0$ )

#### Fit:

$$\ln(p_{ij}/(1-p_{ij})) = \phi_i + \beta_{1i} SBP_{0ij} + \theta_i x_{ij}$$

$$\theta_i = \theta + u_i \quad u_i \sim N(0, \tau^2)$$

Gives summary log odds ratio ( $\hat{\theta}$ ) & its variance ( $V(\hat{\theta})$ ) <sup>32</sup>

### Binary outcomes: logistic regression (random effects model, stratified intercept)

```
+ meqrlogit dvt eryt study1 study2 study3, nocons || studyid: eryt, nocons or
Integration points = 7
Log likelihood = -747.08689
Wald chi2(4) = 620.82
Prob > chi2 = 0.0000
```

dvt	Odds Ratio	Std. Err.	Z	P> z	[95% Conf. Interval]
eryt	1.351516	.1864885	2.18	0.029	1.03126 1.771225
study1	.1295397	.0143114	-18.50	0.000	.104319 .160858
study2	.1839299	.0407957	-7.63	0.000	.1190846 .2840856
study3	.2646055	.029437	-11.95	0.000	.2127660 .3290741

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]
studyid: Identity	3.67e-15	2.00e-08	0
var(eryt)			

LR test vs. logistic model: chibar2(01) = 0.00 Prob >= chibar2 = 1.0000

### Binary outcomes: logistic regression (random effects model, random intercept)

meqrlogit dvt eryt || studyid: eryt, or

dvt	Odds Ratio	Std. Err.	Z	P> z	[95% Conf. Interval]
eryt	1.391412	.2961019	1.55	0.121	.9168036 2.11153
var(eryt)	.1822133	.0345892	-8.97	0.000	.1256025 -.2643393

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]
studyid: Independent			
var(eryt)	.0207127	.0915477	3.59e-06 119.797
var(_cons)	.0760706	.075772	-.0111359 .5306352

We now see that the OR is slightly larger and there is heterogeneity on the effect of erythema; so choice of model specification important (ongoing work to resolve this)

### One-stage approach

#### Time-to-event outcomes: Cox regression (fixed effect model)

- $i = 1 \text{ to } k$  trials
  - The  $j^{\text{th}}$  patient provides their follow-up time ( $t_{ij}$ ), and outcome response (0 or 1, censored or event) at that time
  - $h_j(t)$  is the hazard rate over time;  $h_0(t)$  is the baseline hazard
  - Treatment group ( $x_{ij} = 1$ ) and Control group ( $x_{ij} = 0$ )
- Fit:
- $$h_{ij}(t) = h_0(t) \exp(\beta_{1i} SBP_{0ij} + \theta_i x_{ij})$$

NB Above allows a different baseline hazard shape per study

#### Time-to-event outcomes: Cox regression (fixed effect model)

```
+ stcox treat, strata(trial)
Stratified Cox regression -- no ties
No. of subjects = 3,000
No. of failures = 1,332
Time at risk = 11551,21312
Number of obs = 3,000
LR chi2(1) = 22.12
Prob > chi2 = 0.0000

```

_t	Haz. Ratio	Std. Err.	Z	P> z	[95% Conf. Interval]
treat	.771277	.042684	-4.69	0.000	.6919953 .8596418

Stratified by trial

## One-stage approach

### Time-to-event outcomes: Cox regression (fixed effect model)

- $i = 1$  to  $k$  trials
- The  $j^{\text{th}}$  patient provides their follow-up time ( $t_{ij}$ ), and outcome response (0 or 1, censored or event) at that time
- $h_{ij}(t)$  is the hazard rate over time;  $h_0(t)$  is the baseline hazard
- Treatment group ( $x_{ij} = 1$ ) and Control group ( $x_{ij} = 0$ )
- Fit: 
$$h_{ij}(t) = h_0(t) \exp(\alpha_i + \beta_{1i} SBP_{0ij} + \theta x_{ij})$$

Now each study has same baseline hazard shape (but not same magnitude);  $h_{01}$  is baseline hazard for study 1 and  $\alpha_i$  is omitted.

37

### Time-to-event outcomes: Cox regression (fixed effect model)

		stcox treat trialvar1-trialvar15		Number of obs	LR chi2(15)	Prob > chi2
No. of subjects	No. of failures	t	P>1-t			
3,000	1,332			3,000	530.19	0.0000
Time at risk:	11551.21312					
Log likelihood	-10496.801					
	Haz. Ratio	Std. Err.	Z	P>1-t	[95% Conf. Interval]	
trialvar1	.7714499	.0426183	-4.70	.000	.6922829	.8596702
trialvar1	3.283469	.6191397	6.31	.000	2.268966	4.751577
trialvar2	1.176325	.2560620	0.75	.456	.7677014	1.802259
trialvar3	1.154018	.2538013	0.65	.515	.7499104	1.77589
trialvar4	1.437724	.3032284	1.72	.085	.9508108	2.173957
trialvar5	3.596639	.6720835	6.85	.000	2.493663	5.187476
trialvar6	5.988959	.0808485	9.92	.000	4.204675	8.530414
trialvar7	7.610893	.1360441	11.35	.000	5.361472	10.80406
trialvar8	1.633447	.3351486	2.40	.017	1.093545	2.402556
trialvar9	2.02497	.5400701	5.44	.000	1.944614	4.123237
trialvar10	5.64405	.1406352	9.52	.000	3.9387	6.060937
trialvar11	9.21449	.4504129	8.13	.000	3.11240	6.297337
trialvar12	2.622135	.5044562	5.01	.000	1.798449	3.823072
trialvar13	1.257394	.2710852	1.06	.288	.0240588	1.918595
trialvar14	2.619511	.5066734	4.98	.000	1.792292	3.327032 .38
trialvar15	1	(omitted)				

## One-stage approach

### Time-to-event outcomes: Cox regression (random effects model), e.g. use stmix in Stata (see practical)

- $i = 1$  to  $k$  trials
- The  $j^{\text{th}}$  patient provides their follow-up time ( $t_{ij}$ ), and outcome response (0 or 1, censored or event) at that time
- $h_{ij}(t)$  is the hazard rate over time;  $h_0(t)$  is the baseline hazard
- Treatment group ( $x_{ij} = 1$ ) and Control group ( $x_{ij} = 0$ )
- Fit: 
$$h_{ij}(t) = h_0(t) \exp(\beta_{1i} SBP_{0ij} + \theta_i x_{ij})$$
- $\theta_i = \theta + u_i$        $u_i \sim N(0, \tau^2)$

Gives summary log hazard ratio ( $\hat{\theta}$ ) & its variance ( $V(\hat{\theta})$ )<sup>39</sup>

## Estimation of one-stage models

- Most statistical software packages offer estimation of mixed models for continuous and binary outcomes.
- Mixed linear regression models commonly use restricted maximum likelihood (REML), e.g. Stata 'mixed', SAS 'proc mixed'.
  - Methods for small sample inference also known as denominator-degrees-of-freedom (DDF) adjustments including Satterthwaite and Kenward-Roger.
- Mixed logistic regression typically estimated using ordinary maximum likelihood (via a Gaussian quadrature procedure).
- REML can be implemented through a pseudo-likelihood approach
- Random effects extensions to the Cox model ('Frailty models') possible in some statistical packages, for example 'coxme' in R.
- Can take a long time to estimate one-stage models (see practical 1).
- Alternatively can use a Bayesian approach.

40

STATISTICS IN MEDICINE  
Statist. Med. 2001; 20:2210-2241 (DOI: 10.1002/sim.915)

### Meta-analysis of continuous outcome data from individual patients

Julian P. T. Higgins<sup>1,\*†</sup>, Anne Whitehead<sup>2</sup>, Rebecca M. Turner<sup>3</sup>,  
Rumana Z. Omar<sup>2</sup> and Simon G. Thompson<sup>1</sup>

<sup>1</sup>MRC Biostatistics Unit, Institute of Public Health, University of Cambridge, Robinson Way,  
Cambridge CB2 2SR, UK

<sup>2</sup>Maths and Pharmaceutical Statistics, School of Maths and Physics, University of Reading, P.O. Box 240,  
Earley Gate, Reading, RG6 6FX, UK

<sup>3</sup>MRC Clinical Trials Unit, 22, Weston Road, London NW1 2DA, UK

\*Department of Epidemiology and Public Health, Imperial College School of Medicine,  
Hammersmith Hospital, Du Cane Road, London W12 0NN, UK

#### SUMMARY

Meta-analyses using individual patient data are becoming increasingly common and have several advantages over meta-analyses of summary statistics. We explore the use of multilevel or hierarchical models for the meta-analysis of continuous individual patient outcome data from clinical trials. A general framework is developed which encompasses traditional meta-analysis, as well as multilevel regression and the inclusion of patient-level covariates for investigation of heterogeneity. Unexplained variation in the outcome variable can be accommodated by allowing for random effects at both trial and patient level, although an extension to a random effect for trial is described. The methods are illustrated on an example in Alzheimer's disease in a classical framework using SAS PROC MIXED and MLwiN, and in a Bayesian framework using BUGS. Relative merits of the three software packages for such meta-analyses are discussed in terms of ease of use, computational efficiency and extension to incorporate more than two treatments. Copyright © 2001 John Wiley & Sons, Ltd.

41

STATISTICS IN MEDICINE  
Statist. Med. 2001; 20:2210-2241 (DOI: 10.1002/sim.915)  
Published online 11 December 2001 in Wiley InterScience  
(www.interscience.wiley.com) DOI: 10.1002/sim.915

### Meta-analysis of continuous outcomes combining individual patient data and aggregate data

Richard D. Riley<sup>1,2,\*†</sup>, Paul C. Lambert<sup>2</sup>, Ian A. Stratton<sup>3</sup>, Yiqiang Wang<sup>4</sup>,  
Frances Greyson<sup>1</sup>, Langdon Tugwell<sup>2</sup> and Horst Riedel<sup>5</sup>

<sup>1</sup>Centre for Medical Statistics and Health Evaluation, Faculty of Medicine, University of Liverpool,  
Lancaster Campus, Liverpool, UK; <sup>2</sup>School of Health and Life Sciences, Aston University, Aston Triangle, Birmingham B4 7ET, UK;

<sup>3</sup>Centre for Biomarker Research, Keele University, Keele, UK; <sup>4</sup>Department of Mathematics, University of Bristol, Bristol, UK; <sup>5</sup>Department of Cardiovascular Disease, University of Western Ontario, London, Ontario, Canada N6A 3K7

\*Correspondence to: Richard D. Riley, Department of Primary Care and Population Sciences, University of

London, 187 Colegate, York YO1 5WY, UK; E-mail: richard.riley@liverpool.ac.uk

†See authorship section for details of contributions.

#### SUMMARY

Meta-analysis of individual patient data (IPD) is the gold standard for synthesising evidence across clinical studies. However, for some studies IPD may not be available and only aggregate data (AD) such as a treatment effect estimate and its standard error, may be obtained. In this situation, methods for combining IPD and AD are important to explore all the available evidence. In this paper, we develop and assess a range of methods for combining IPD and AD in meta-analysis of continuous outcomes from randomised controlled trials.

42

STATISTICS IN MEDICINE  
Statist. Med. 2000, 19, 3417–3422

## A multilevel model framework for meta-analysis of clinical trials with binary outcomes<sup>a</sup>

Rebecca M. Turner<sup>a,1</sup>, Runana Z. Omar<sup>a</sup>, Min Yang<sup>b</sup>, Harvey Goldstein<sup>b</sup>  
and Simon G. Thompson<sup>a</sup>

<sup>a</sup> MRC Clinical Trials Unit, 222 Queen's Road, London NW1 2DA, UK  
<sup>b</sup> Department of Epidemiology and Public Health, Imperial College School of Medicine, Du Cane Road, London NW1 4NN, UK

<sup>a</sup> Medical Statistics Group, Institute of Education, 26 Bedford Way, London WC1E 6HT, UK  
<sup>b</sup> MRC Biostatistics Unit, Institute of Public Health, Cambridge CB2 2PS, UK

**SUMMARY**  
In this paper we explore the potential of multilevel models for meta-analysis of trials with binary outcomes. In both summary data such as log-odds ratios and individual patient data, conventional fixed effect and random effects meta-analyses are based on model-free methods, which do not allow for heterogeneity or restricted maximum likelihood estimation. To exemplify the methods, we use the results from 22 trials to prevent respiratory tract infections. We also make comparisons with a second example, data comprising three trials of different types of oral contraceptives. We compare the multilevel model approach with the traditional random effects model and the log-likelihood ratio test. We show how the multilevel model approach can estimate between-trial variance more precisely than the standard random effects model. The multilevel model approach is more robust to heterogeneity than the random effects model. The multilevel model approach can also estimate the between-trial variance. The trial effects may be modelled as either fixed or random within individual data models, and we discuss the corresponding assumptions and implications. If random trial effects are used, the multilevel model approach is equivalent to the random effects model. If fixed trial effects are used, the multilevel model approach is equivalent to a hierarchical approach to meta-analysis. Having implemented these techniques, the flexibility of multilevel modelling may be explored by facilitating extensions to standard meta-analysis methods. Copyright © 2000 John Wiley & Sons, Ltd.

43

STATISTICS IN MEDICINE  
Statist. Med. 2005, 24(10):1219

Published online 31 January 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sim.2050

## Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes

Catrin Tudor Smith<sup>1,2</sup>, Paula R. Williamson<sup>3</sup> and Anthony G. Marson<sup>2</sup>

<sup>1</sup> Centre for Medical Statistics and Health Evaluation, University of Liverpool, Liverpool L69 3BV, UK  
<sup>2</sup> Department of Neurology, University of Liverpool, Liverpool L69 3BV, UK

**SUMMARY**  
Different aspects underlie the heterogeneity in meta-analysis: methodology, clinical practice, and patient characteristics, are factors that can contribute to variability in the treatment effect between studies in a meta-analysis (statistical heterogeneity). Regression modeling can be used to examine relationships between treatment effects and covariates with the aim of explaining the variability in terms of clinical and methodological factors. In this paper, we propose a regression model for the analysis of individual patient data. An aggregate data approach can be problematic as sufficient data are rarely available and modelling aggregate effects in individual patients can often be misleading. An individual patient data approach overcomes these problems by allowing the inclusion of all data and the identification of potential sources of heterogeneity, and enables a full analysis of time to event outcomes in meta-analyses.

**Methodology** Cox regression models are used to identify and explore the evidence for heterogeneous effects and examine the relationship between covariates and censored failure time data in this context. Alternative formulations of the model are possible and illustrated using individual patient data from a meta-analysis of five randomized controlled trials, which compare two drugs for the treatment of atrial fibrillation. The model is extended to include time-varying covariates to explore the effect of heterogeneity and magnitude of treatment effect as varied. The behaviour of each model in each situation is explored and compared. Copyright © 2005 John Wiley & Sons, Ltd.

**KEY WORDS:** heterogeneity; time to event; meta-analysis; Cox regression model; individual patient data

44

Coverline et al. BMC Medical Research Methodology 2005, 5:124  
http://www.biomedcentral.com/1471-2298/5/124

**BMC**  
Medical Research Methodology

**RESEARCH ARTICLE** Open Access

## Individual patient data meta-analysis of survival data using Poisson regression models

Michael J Crowther<sup>1</sup>, Richard D Riley<sup>2</sup>, Jan A Stoenen<sup>1,3</sup>, Jiguang Wang<sup>1</sup>, François Guayffier<sup>4</sup> and Paul C Lambert<sup>1,2</sup>

**Abstract**  
Background: An Individual Patient Data (IPD) meta-analysis is often considered the gold standard for synthesising survival data from clinical trials. An IPD meta-analysis can be achieved by either a two-stage or a one-stage approach, depending on whether the trials are analysed separately or simultaneously. A range of one-stage approaches have been developed but have only recently appeared, but these are known to be computationally intensive and are not currently available in all standard statistical software. We describe an alternative approach using Poisson Based Generalised Linear Models (GLMs).

**Methods:** We illustrate, through application and simulation, the Poisson approach both classically and in a Bayesian framework, in two-stage and one-stage approaches. We outline the benefits of our one-stage approach through extension to modelling treatment towards interactions and non proportional hazard. Ten trials of hypertension treatment, with all cause death the outcome of interest, are used to apply and assess the approach.

**Results:** We show that the Poisson approach obtains almost identical estimates to the Cox model. It is additionally computationally efficient and directly estimates the baseline hazard. Some downward bias is observed in classical estimates of the heterogeneity in the treatment effect, with improved performance from the Bayesian approach.

**Conclusion:** Our approach provides a highly flexible and computationally efficient framework available in all standard statistical software, to the investigation of not only heterogeneity, but the presence of non proportional hazards and treatment effect modifiers.

45

## Part 3:

---

### Combining IPD and aggregate data in a one-stage meta-analysis

**Part 3:**

---

### Combining IPD and non-IPD

**Research Article**

Received 21 May 2004 Accepted 21 December 2005  
Published online 21 March 2006

**Binary outcomes – for further details see:**

**Research Synthesis Methods**

**Research Article**

Received 21 May 2004 Accepted 21 December 2005  
Published online 21 March 2006

**Meta-analysis of a binary outcome using individual participant data and aggregate data**

Richard D. Riley<sup>a,1,2</sup> and Ewout W. Steyerberg<sup>b,3</sup>

In this paper, we develop meta-analysis models that synthesize a binary outcome from health-care studies while accounting for participant-level covariates. In particular, we show how to synthesize the observed event-risk across studies while accounting for the study-specific event probability and the study-specific participant-level covariate probability. The models are adjusted for situations where studies provide individual participant data (IPD), or a mixture of IPD and aggregate data. We show that the availability of IPD is crucial in at least some studies; this allows one to estimate the study-specific event probability and the study-specific participant-level covariate probability to account for potential ecological bias and study level confounding. The models can produce pertinent population-level and individual-level results, such as the pooled overall and the covariate-specific event probability for an individual. Furthermore, the models can produce population-level and individual-level results for the association between age and death mortality risk as synthesized in relation to individual age. The results show that as individual age increases the probability of six-month mortality also increases; further, the results reveal clear evidence of ecological bias, with the probability to die in each study additionally influencing an individual's mortality probability. Copyright © 2010 John Wiley & Sons, Ltd.

**Keywords:** meta-analysis; evidence synthesis; individual participant data; IPD; binary data; participant-level covariate

46

### Combining IPD and non-IPD

- Recall that IPD and non-IPD results can be combined in the second stage of the two-step IPD meta-analysis approach
- Also possible in the one-stage approach
- In brief, we need to specify two related regression models (one for the IPD studies, and one for the non-IPD studies)
- Estimate them together (joint likelihood)

47

**Binary outcomes – for further details see:**

**Research Synthesis Methods**

**Research Article**

Received 21 May 2004 Accepted 21 December 2005  
Published online 21 March 2006

**Meta-analysis of a binary outcome using individual participant data and aggregate data**

Richard D. Riley<sup>a,1,2</sup> and Ewout W. Steyerberg<sup>b,3</sup>

In this paper, we develop meta-analysis models that synthesize a binary outcome from health-care studies while accounting for participant-level covariates. In particular, we show how to synthesize the observed event-risk across studies while accounting for the study-specific event probability and the study-specific participant-level covariate probability. The models are adjusted for situations where studies provide individual participant data (IPD), or a mixture of IPD and aggregate data. We show that the availability of IPD is crucial in at least some studies; this allows one to estimate the study-specific event probability and the study-specific participant-level covariate probability to account for potential ecological bias and study level confounding. The models can produce pertinent population-level and individual-level results, such as the pooled overall and the covariate-specific event probability for an individual. Furthermore, the models can produce population-level and individual-level results for the association between age and death mortality risk as synthesized in relation to individual age. The results show that as individual age increases the probability of six-month mortality also increases; further, the results reveal clear evidence of ecological bias, with the probability to die in each study additionally influencing an individual's mortality probability. Copyright © 2010 John Wiley & Sons, Ltd.

**Keywords:** meta-analysis; evidence synthesis; individual participant data; IPD; binary data; participant-level covariate

48

## Continuous outcomes – for further details see:

STATISTICS IN MEDICINE  
Statistics Med 2005; 27:1870-1893  
Published online 11 December 2005 in Wiley InterScience  
(www.interscience.wiley.com) DOI: 10.1002/sim.3165

Meta-analysis of continuous outcomes combining individual patient data and aggregate data

Richard D. Riley<sup>1,\*†</sup>, Paul C. Lambert<sup>2</sup>, Jan A. Staessen<sup>3</sup>, Jiguang Wang<sup>4</sup>, Francois Gueyffier<sup>5</sup>, Lutgarde Thijs<sup>3</sup> and Florent Boutron<sup>6</sup>

Which builds on ...

Goldstein H, Yang M, Omar RZ, Turner RM, Thompson SG. Meta-analysis using multilevel models with an application to the study of class size effects. JRSS Series C. 2000;49:399-412

49

## Part 4:

### One-stage versus two-stage

50

## One-stage vs. two-stage: Does it make a difference?

- Short answer: usually not, but occasionally it might.
- Many theoretical and empirical papers say the estimates are usually similar

Okin I, Sampson A. Comparison of meta-analysis versus analysis of variance of individual patient data. *Biometrics*, 1998;54(1):317-22.

Bowden J, et al. Individual patient data meta-analysis of time-to-event outcomes: one-stage versus two-stage approaches for estimating the hazard ratio under a random-effects model. *Research Synthesis Methods*, 2011;2(3):150-162.

Stewart G, et al. Statistical analysis of individual participant data meta-analyses: a comparison of methods and recommendations for practice. *PLoS One*, 2012; 7(10):e46042

51

## Return to the hypertension example:

One-stage and two-stage give very similar results in this example

Approach	$\beta$ ( $s.e(\beta)$ )	$t^*$
One-stage (REML)	-10.16 (1.08)	7.13
Two-stage (REML)	-10.17 (1.07)	7.10

52

## Empirical evaluation

- Generally we agree that approaches give similar results in most situations.
- But there are situations where (willingly or naively), the assumptions will be different depending on the choice of the one-stage or two-stage approach
  - This can lead to differences in summary estimates.
- We have identified ten key reasons in a recent paper – here I discuss five.

53

## Tutorial in Biostatistics

Received: 11 March 2016 Accepted: 13 September 2016  
(wileyonlinelibrary.com) DOI: 10.1002/sim.7141

Statistics  
in Medicine

### Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ

Danielle L. Burke<sup>\*†</sup>, Joie Ensor and Richard D. Riley

Meta-analysis using individual participant data (IPD) obtains and synthesizes the raw, participant-level data from a set of relevant studies. The IPD approach is becoming an increasingly popular tool as an alternative to traditional aggregate data meta-analysis, especially as it avoids reliance on published results and provides an opportunity to include individual patient data. In this tutorial, we introduce the IPD approach and describe the statistical procedures for conducting an IPD meta-analysis, one-stage and two-stage. The one-stage approach analyses the IPD from all studies simultaneously, for example, in a hierarchical regression model with random effects. The two-stage approach derives aggregate data effect sizes and effect estimates in each study separately and then combines them via a fixed effect model. The authors compare the two-stage and one-stage approaches via theoretical consideration, simulation and empirical examples, yet there remains confusion regarding when each approach should be adopted, and indeed why they may differ.

In this tutorial paper, we outline the key statistical methods for one-stage and two-stage IPD meta-analyses, and highlight the strengths and weaknesses of each approach. We explain the key differences arise because of different modeling assumptions, rather than the choice of one-stage or two-stage itself. We illustrate the concepts with recently published IPD meta-analyses, summarize key statistical software and provide a guide to further IPD meta-analyses. © 2016 The Authors. Statistics in Medicine published by John Wiley & Sons Ltd.

Keywords: individual patient data; individual participant data; meta-analysis; IPD; one-stage; two-stage

54

## Key Reason 1: Exact one-stage versus approximate two-stage likelihoods

### Two-stage

- 2<sup>nd</sup> step, study effect estimates, e.g. log(OR), assumed approximately normally distributed with known variances
- Inappropriate for studies with small sample sizes and/or small numbers of events for binary and time-to-event outcomes
- Arbitrary continuity correction for zero cells with binary outcomes

### One-stage

- Better to model the exact likelihood, e.g.  $r_{ij} \sim \text{Bin}(1, p_{ij})$  with logistic regression model and avoid the normality assumption.

55

**Journal of Clinical Epidemiology**

Journal of Clinical Epidemiology 61 (2008) 11–51

The binomial distribution of meta-analysis was preferred to model within-study variability

Taye H. Hamza<sup>a,b</sup>, Hans C. van Houwelingen<sup>b</sup>, Theo Stijnen<sup>a</sup>

<sup>a</sup>Department of Epidemiology and Biostatistics, Erasmus MC—Erasmus University Medical Center Rotterdam, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands  
<sup>b</sup>Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands

Accepted 21 March 2008

**Abstract**

**Objective:** When meta-analytic methods are used, it is important to know which method should be used. This paper compares the exact likelihood approach with the approximate binomial likelihood approach.

**Study Design:** Coverage probability analysis.

**Setting:** Two-stage likelihood analysis.

**Participants:** The exact likelihood approach gives unbiased estimates, while the approximate approach gives biased estimates.

**Interventions:** The exact likelihood approach gives unbiased estimates, while the approximate approach gives biased estimates.

**Results:** The exact likelihood approach gives unbiased estimates, while the approximate approach gives biased estimates.

**Conclusion:** The exact likelihood approach is the method of preference and should be used whenever feasible. © 2008 Elsevier Inc. All rights reserved.

"The exact likelihood approach is the method of preference and should be used whenever feasible"

using the DerSimonian and Laird's weighted distribution to study distribution of the standard test statistic

the mean squared error and variance of the within-study sample

coverage probability and gives unbiased estimates. The approximate approach gives biased estimates.

Conclusion: The exact likelihood approach is the method of preference and should be used whenever feasible. © 2008 Elsevier Inc. All rights reserved.

OPEN ACCESS freely available online

PLOS ONE

## Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage?

Thomas P. A. Debray<sup>1</sup>, Karol G. M. Moons<sup>1</sup>, Ghada Mohammed Abdalla Abo-Zaid<sup>2</sup>, Hendrik Koffijberg<sup>3</sup>, Richard David Riley<sup>4</sup>

<sup>1</sup> MRC Centre for Health Services and Delivery Research, Medical School of Cancer Research Institute, The University of Nottingham, Nottingham Clinical Research Centre, Nottingham, United Kingdom, <sup>2</sup> School of Health and Population Sciences, and School of Mathematics, Public Health, Epidemiology and Biostatistics, College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom

<sup>3</sup> Department of Epidemiology and Biostatistics, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands, <sup>4</sup> School of Mathematics, Public Health, Epidemiology and Biostatistics, College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom

E-mail: r.d.riley@nottingham.ac.uk

† Present address: Department of Primary Care and Population Sciences, University of California Los Angeles, Los Angeles, California, United States of America

Competing interests: The authors have declared that no competing interests exist.

Received: January 10, 2008; Accepted: April 1, 2008; Published: May 1, 2008

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Copyright: © 2008 Debray et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

URL: <http://dx.doi.org/10.1371/journal.pone.0002121>

Citation: Debray TP, Moons KG, Abo-Zaid GA, Koffijberg H, Riley RD (2008) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? PLoS ONE 3(5): e2121. doi:10.1371/journal.pone.0002121

Editor: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Review by: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Published: May 1, 2008

Copyright: © 2008 Debray et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

URL: <http://dx.doi.org/10.1371/journal.pone.0002121>

Citation: Debray TP, Moons KG, Abo-Zaid GA, Koffijberg H, Riley RD (2008) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? PLoS ONE 3(5): e2121. doi:10.1371/journal.pone.0002121

Editor: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Review by: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Published: May 1, 2008

Copyright: © 2008 Debray et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

URL: <http://dx.doi.org/10.1371/journal.pone.0002121>

Citation: Debray TP, Moons KG, Abo-Zaid GA, Koffijberg H, Riley RD (2008) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? PLoS ONE 3(5): e2121. doi:10.1371/journal.pone.0002121

Editor: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Review by: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Published: May 1, 2008

Copyright: © 2008 Debray et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

URL: <http://dx.doi.org/10.1371/journal.pone.0002121>

Citation: Debray TP, Moons KG, Abo-Zaid GA, Koffijberg H, Riley RD (2008) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? PLoS ONE 3(5): e2121. doi:10.1371/journal.pone.0002121

Editor: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Review by: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Published: May 1, 2008

Copyright: © 2008 Debray et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

URL: <http://dx.doi.org/10.1371/journal.pone.0002121>

Citation: Debray TP, Moons KG, Abo-Zaid GA, Koffijberg H, Riley RD (2008) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? PLoS ONE 3(5): e2121. doi:10.1371/journal.pone.0002121

Editor: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Review by: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Published: May 1, 2008

Copyright: © 2008 Debray et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

URL: <http://dx.doi.org/10.1371/journal.pone.0002121>

Citation: Debray TP, Moons KG, Abo-Zaid GA, Koffijberg H, Riley RD (2008) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? PLoS ONE 3(5): e2121. doi:10.1371/journal.pone.0002121

Editor: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Review by: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Published: May 1, 2008

Copyright: © 2008 Debray et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

URL: <http://dx.doi.org/10.1371/journal.pone.0002121>

Citation: Debray TP, Moons KG, Abo-Zaid GA, Koffijberg H, Riley RD (2008) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? PLoS ONE 3(5): e2121. doi:10.1371/journal.pone.0002121

Editor: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Review by: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Published: May 1, 2008

Copyright: © 2008 Debray et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

URL: <http://dx.doi.org/10.1371/journal.pone.0002121>

Citation: Debray TP, Moons KG, Abo-Zaid GA, Koffijberg H, Riley RD (2008) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? PLoS ONE 3(5): e2121. doi:10.1371/journal.pone.0002121

Editor: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Review by: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Published: May 1, 2008

Copyright: © 2008 Debray et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

URL: <http://dx.doi.org/10.1371/journal.pone.0002121>

Citation: Debray TP, Moons KG, Abo-Zaid GA, Koffijberg H, Riley RD (2008) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? PLoS ONE 3(5): e2121. doi:10.1371/journal.pone.0002121

Editor: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Review by: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Published: May 1, 2008

Copyright: © 2008 Debray et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

URL: <http://dx.doi.org/10.1371/journal.pone.0002121>

Citation: Debray TP, Moons KG, Abo-Zaid GA, Koffijberg H, Riley RD (2008) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? PLoS ONE 3(5): e2121. doi:10.1371/journal.pone.0002121

Editor: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Review by: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Published: May 1, 2008

Copyright: © 2008 Debray et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

URL: <http://dx.doi.org/10.1371/journal.pone.0002121>

Citation: Debray TP, Moons KG, Abo-Zaid GA, Koffijberg H, Riley RD (2008) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? PLoS ONE 3(5): e2121. doi:10.1371/journal.pone.0002121

Editor: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

Review by: Michael J. Pocock, London School of Economics and Political Science, United Kingdom

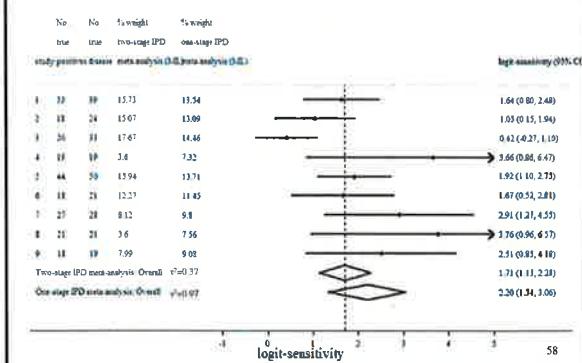
Published: May 1, 2008

Copyright: © 2008 Debray et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

URL: <http://dx.doi.org/10.1371/journal.pone.0002121>

Citation: Debray TP, Moons KG, Abo-Zaid GA, Koffijberg H, Riley RD (2008) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? PLoS ONE 3(5): e2121. doi:10.1371/journal.pone.0002121

## Example (Hamza et al., 2008)



## Key Reason 2: Choice of specification for any adjustment terms

- Can adopt different specifications for any adjustment terms
- Two-stage automatically assumes a different effect of each adjustment factor in each trial
  - First stage

$$SBP_{ij} = \varphi_i + \beta_i SBP_{0ij} + \theta_i x_{ij} + \varepsilon_{ij}; \quad \varepsilon_{ij} \sim N(0, \sigma_i^2)$$

- One-stage can replicate this by stratifying the effect by study

$$SBP_{ij} = \varphi_i + \beta_{1i} SBP_{0ij} + \theta_{1i} x_{ij} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N(0, \sigma_i^2)$$

Stratifying the effect by trial here

## Key Reason 2: Choice of specification for any adjustment terms

- Fixed adjustment effect is problematic if heterogeneity in adjustment factor

$$SBP_{ij} = \varphi_i + \beta_i SBP_{0ij} + \theta_i x_{ij} + \varepsilon_{ij}; \quad \varepsilon_{ij} \sim N(0, \sigma_i^2)$$

- Could assume adjustment effects are drawn randomly from a distribution

$$SBP_{ij} = \varphi_i + \beta_i SBP_{0ij} + \theta_i x_{ij} + \varepsilon_{ij}$$

$$\beta_i \sim N(\beta, \delta^2) \quad \varepsilon_{ij} \sim N(0, \sigma_i^2)$$

- Leads to multiple random effects - covariance matrix may specification may then be influential.

### Example: Hypertension data

- Whether smoking is a prognostic factor for high blood pressure at follow-up, adjusting for baseline blood pressure and treatment group.

Approach	Model specification in regard adjustment factors	Summary mean difference (smokers versus non-smokers), 95% CI
Two-stage	Distinct per trial	1.763 (1.146 to 2.380)
One-stage	Distinct per trial	1.756 (1.043 to 2.469)

61

### Example: Hypertension data

- Whether smoking is a prognostic factor for high blood pressure at follow-up, adjusting for baseline blood pressure and treatment group.

Approach	Model specification in regard adjustment factors	Summary mean difference (smokers versus non-smokers), 95% CI
Two-stage	Distinct per trial	1.763 (1.146 to 2.380)
One-stage	Distinct per trial	1.756 (1.043 to 2.469)
One-stage	Fixed (common)	1.689 (0.951 to 2.426)

62

### Example: Hypertension data

- Whether smoking is a prognostic factor for high blood pressure at follow-up, adjusting for baseline blood pressure and treatment group.

Approach	Model specification in regard adjustment factors	Summary mean difference (smokers versus non-smokers), 95% CI
Two-stage	Distinct per trial	1.763 (1.146 to 2.380)
One-stage	Distinct per trial	1.756 (1.043 to 2.469)
One-stage	Fixed (common)	1.689 (0.951 to 2.426)
One-stage	Random (correlated)	1.523 (0.731 to 2.316)

63

### Specification for adjustment terms in IPD meta-analysis: Stata code example

#### One-stage

- Fixed sbpi adjustment term, random treatment effect

```
mixed sbpl i.trialdummy sbpi treat, nocons ///
|| trialdummy: treat, nocons reml var
```

- Stratified sbpi adjustment, random treatment effect

```
mixed sbpl i.trialdummy c.sbp#i.trialdummy treat, nocons ///
|| trialdummy: treat, nocons reml var
```

- Random sbpi adjustment, random treatment effect

```
mixed sbpl i.trialdummy sbpi treat, nocons ///
|| trialdummy: treat sbpi, nocons reml var cov(uns)
```

#### Two-stage

```
ipdmelan, study(trial) keepall re(reml): reg sbpl sbpi treat64
```

### Key Reason 3: Choice of specification for the residual variances

#### Two-stage

- Continuous outcomes, automatically assumes distinct residual variances in each trial.

#### One-stage

- Should allow residual variances to be distinct in each trial, such that  $\epsilon_i \sim N(0, \sigma_i^2)$ .
- Usually tends to simplify this and assume same residual variance in each trial, such that  $\epsilon_i \sim N(0, \sigma^2)$ .

65

### Example: Hypertension data

- Effect of hypertension treatment on systolic blood pressure (REML estimation)

Approach	Assumption	Summary mean difference (95% CI)	$\tau^2$
Two-stage	Different residual variance	-10.17 (-12.27 to -8.07)	7.10
One-stage	Different residual variance	-10.16 (-12.27 to -8.06)	7.13

66

### Example: Hypertension data

- Effect of hypertension treatment on systolic blood pressure (REML estimation)

Approach	Assumption	Summary mean difference (95% CI)	$r^2$
Two-stage	Different residual variance	-10.17 (-12.27 to -8.07)	7.10
One-stage	Different residual variance	-10.16 (-12.27 to -8.06)	7.13
One-stage	Same residual variance	-10.34 (-12.55 to -8.13)	8.19

Larger treatment effect and larger estimate of  $r^2$  when assume same residual variance

67

### Key Reason 4: Different estimation method for $r^2$

- Availability of estimation method
- Convergence issues
- Well-known that ML tends to underestimate  $r^2$  and REML is preferred
  - For non-continuous outcomes, one-stage models typically use ML estimation as it is the only option.
- Method-of-moments estimator of DerSimonian and Laird
  - most common estimation method in a two-stage approach
  - the only option in RevMan

68

### Example: Hypertension data

- Effect of hypertension treatment on systolic blood pressure

Approach	Estimation method	Summary mean difference (95% CI)	$r^2$
Two-stage	REML	-10.17 (-12.27 to -8.07)	7.10
One-stage	REML	-10.16 (-12.27 to -8.06)	7.13

69

### Example: Hypertension data

- Effect of hypertension treatment on systolic blood pressure

Approach	Estimation method	Summary mean difference (95% CI)	$r^2$
Two-stage	REML	-10.17 (-12.27 to -8.07)	7.10
One-stage	REML	-10.16 (-12.27 to -8.06)	7.13
Two-stage	ML	-10.10 (-12.03 to -8.16)	5.84
One-stage	ML	-10.03 (-11.83 to -8.23)	4.94

70

### Example: Hypertension data

- Effect of hypertension treatment on systolic blood pressure

Approach	Estimation method	Summary mean difference (95% CI)	$r^2$
Two-stage	REML	-10.17 (-12.27 to -8.07)	7.10
One-stage	REML	-10.16 (-12.27 to -8.06)	7.13
Two-stage	ML	-10.10 (-12.03 to -8.16)	5.84
One-stage	ML	-10.03 (-11.83 to -8.23)	4.94
Two-stage	MoM	-9.85 (-11.13 to -8.57)	3.07

- Choice of best estimator for  $r^2$  is ongoing issue in the meta-analysis field.

71

### Key Reason 5: Accounting for correlation amongst model parameters

#### Does the model account for correlation?

- One-stage model ✓
- Two-stage approach ✗ unless use multivariate meta-analysis
- Particularly relevant when dealing with longitudinal outcomes
- Usually there are missing data
  - Not all patients provide data at all time-points.
  - Not all studies will measure the same time-points

72

### Example: Jones et al. 2009

- Outcome over time was mini-mental state examination

Time point
1 month
2 months
4 months
6 months
9 months
12 months

73

### Example: Jones et al. 2009

- Outcome over time was mini-mental state examination

Time point	One-stage model (fixed effect)
1 month	0.31 (0.47)
2 months	-0.48 (0.62)
4 months	0.34 (0.48)
6 months	0.20 (0.49)
9 months	0.35 (0.53)
12 months	-0.02 (0.56)

- Borrowing of strength from available data to inform missing data and increase precision.

74

### Example: Jones et al. 2009

- Outcome over time was mini-mental state examination

Time point	One-stage model (fixed effect)	Standard two-stage approach (univariate, fixed effect)
1 month	0.31 (0.47)	0.43 (0.54)
2 months	-0.48 (0.62)	-0.84 (0.97)
4 months	0.34 (0.48)	0.75 (0.57)
6 months	0.20 (0.49)	0.31 (0.50)
9 months	0.35 (0.53)	0.69 (0.63)
12 months	-0.02 (0.56)	0.29 (0.66)

75

### Example: Jones et al. 2009

- Outcome over time was mini-mental state examination

Time point	One-stage model (fixed effect)	Standard two-stage approach (univariate, fixed effect)	Two-stage model (multivariate, fixed effect)
1 month	0.31 (0.47)	0.43 (0.54)	0.30 (0.47)
2 months	-0.48 (0.62)	-0.84 (0.97)	-0.47 (0.59)
4 months	0.34 (0.48)	0.75 (0.57)	0.33 (0.47)
6 months	0.20 (0.49)	0.31 (0.50)	0.19 (0.48)
9 months	0.35 (0.53)	0.69 (0.63)	0.34 (0.52)
12 months	-0.02 (0.56)	0.29 (0.66)	-0.03 (0.55)

76

### Five other keys reasons

6. Clustering and the choice of specification for the intercept
7. Likelihood-based one-stage versus alternative weighting schemes in two-stage
8. Choice of fixed effect or random effects meta-analysis model
9. Derivation of confidence intervals for summary estimates
10. Ecological bias when investigating treatment-covariate interactions

77

### Summary of one- vs. two-stage approach

**Most differences between one-stage and two-stage approaches occur due to:**

1. Different modelling assumptions
2. And/or different estimation methods for  $\tau^2$  and 95% confidence intervals

**When 1 and 2 are the same, the two approaches give very similar results**

**Major benefit of one-stage approach is when there are rare events (can use exact likelihood) but even then, estimation methods may be restrictive (e.g. no REML)**

78

## Part 5:

### Reporting IPD meta-analyses

79

#### Reporting IPD meta-analyses

- IPD meta-analyses often poorly reported  
e.g. see Simmonds et al., 2005
- Often not transparent in regards:
  - Process of obtaining IPD
  - Inclusions/exclusion
  - Missing IPD
  - Analysis methods, handling of missing patient data
  - Risk of bias of included studies
  - Display of individual study results, ...
- To address this, Stewart et al. published PRISMA-IPD.

80

### Reporting IPD meta-analyses

Clinical Review & Education

Special Communication

#### Preferred Reporting Items for a Systematic Review and Meta-analysis of Individual Participant Data The PRISMA-IPD Statement

Lesley A. Stewart, PhD; Mike Clarke, DPhil; Maroeska Rovers, PhD; Richard D. Riley, PhD; Mark Simmonds, PhD; Gavin Stewart, PhD; Jayne F. Timmer, PhD, for the PRISMA-IPD Development Group

JAMA 2015;313(16):1657-1665. doi:10.1001/jama.2015.3656

Extends the original PRISMA statement:

Moher D, Liberati A, Tetzlaff J, Altman DG, (for the PRISMA Group). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339:b2535

81

#### PRISMA-IPD: Example items

Introduction	3	Describe the rationale for the review in the context of what is already known.
Protocol or registration	4	Provide an explicit statement of the questions being addressed with reference, as applicable, to participants, interventions, comparisons, outcomes, and study design (PICO). Include any hypotheses that relate to particular types of participant-level subgroups.
<b>Methods</b>		
Protocol and registration	5	Indicate if a protocol exists and where it can be accessed. If available, provide registration information including registration number and registry name. Provide publication details, if applicable.
Eligibility criteria	6	Specify inclusion and exclusion criteria including those relating to participants, interventions, comparisons, outcomes, study designs, and characteristics (e.g., years when conducted, region or setting, language). Also describe any aspects of study selection that might have influenced the final study set, whether eligible participants were excluded (and ineligible participants excluded) from a study that included a wider population than specified by the review inclusion criteria. The rationale for criteria should be stated.
Identifying studies—Information sources	7	Describe all methods of identifying published and unpublished studies, including, as applicable, which databases were searched and dates of coverage, details of any hand searching (including of conference proceedings), use of study registers and agency or company databases; contact with the original research team and experts in the field; open access journals; and surveys. Give the date of last search or citation.
Identifying studies—Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.
Study selection processes	9	State the process for determining which studies were eligible for inclusion.
Data collection processes	10	Describe how IPD were requested, collected, and managed, including any processes for querying and confirming data with investigators. If IPD were not sought from any eligible study, the reason for this should be stated (for each such study).
If applicable, describe how any studies for which IPD were not available were dealt with. This should include whether, how, and what aggregate data were sought or extracted from study reports and publications (such as extracting data independently in duplicate) and any processes for obtaining and confirming these data with investigators.		

#### PRISMA-IPD: Example items

Synthesis methods	14	Describe the meta-analysis methods used to synthesize IPD. Specify any statistical methods and models used. If necessary, should include those that are not restricted to: <ul style="list-style-type: none"><li>Use of 1-stage or 2-stage approach</li><li>How individual patient data were generated separately within each study and combined across studies (where applicable)</li><li>Specification of any stage models (where applicable) including how clustering of patients within studies was accounted for<ul style="list-style-type: none"><li>Use of fixed- or random-effects models, and any other model assumptions, such as ignorability/hazards</li><li>How (heterogeneity) variances were generated (where applicable)</li><li>Model specification (e.g., random effects, random intercepts, random slopes)</li><li>How studies providing IPD and not providing IPD were analyzed together (where applicable)</li><li>How unique data within IPD were dealt with (e.g., aggregation)</li></ul></li></ul>
Exploration of variation in effects	15	If applicable, describe any exploration of heterogeneity between studies and how it was resolved. State all participant-level characteristics that were analyzed as potential effect modifiers and whether these were pre-specified.
Risk of bias across studies	15	State for any analysis of risk of bias among the included studies of evidence, including any processes for assessing risk of bias and assessing IPD for specific studies, outcomes, or other variables.
Additional analyses	16	Describe methods of any additional analyses, including sensitivity analyses. State where or how these were pre-specified.
Risk of bias within studies	17	Present data in tabular form if feasible, showing the number of patients included in the analysis and the number of events (adverse and non-adverse) for each study, estimates, and 95% confidence intervals, for each intervention compared to control. Consider how any potential bias affects the interpretation of results, through a funnel plot.
Results of individual studies	18	For each comparison and for each main outcome measure or NNT, for each individual study report the number of eligible patients for whom treatment was administered, the number of patients included in the analysis (denominator), and the number of events (adverse and non-adverse), and 95% confidence intervals. These may be calculated or estimated on a forest plot.
Results of synthesis	21	Present summary effects for each meta-analysis undertaken, including confidence intervals and measures of statistical heterogeneity. If applicable, present results of funnel plots based on these data.
When exploring estimates due to patient or study characteristics, present summary interaction estimates for each characteristic examined, including confidence intervals and measures of statistical heterogeneity. State whether the analysis was pre-specified. State whether any interactions were statistically significant.		
Provide a description of the direction and size of effects in terms meaningful to those who would put them into practice.		

85

#### Reporting recommendations for analysis

- Be transparent about all choices
- Pre-specify in the study protocol the approach, estimation method, and modelling assumptions, and why you chose them
- Could pre-specify that you will do both one- and two-stage, and see if conclusions are vulnerable to the choice of method
- If you do both approaches, present both and explain why they are different (if indeed they are)

84

## Reporting considerations for results

- Two-stage approach is more amenable to forest plots
- One-stage approach 'hides' the contribution of each study and does not provide study weights as naturally
- Lecture 4 explains how to derive weights for one-stage analysis (Riley et al., Stat Meth Med Res, in-press)
- Could present study-specific estimates and CIs on a forest plot but with the one-step pooled result forced at the bottom (see practicals), with the one-stage weights shown

85

## Summary

- One-stage approach to IPD meta-analysis becoming popular.
- One-stage models allow more flexible, multi-parameter modelling – but increased computational time, occasional lack of convergence.
- But please - do not ignore clustering!
- Suggest stratify the model by trial (separate control group risk in each trial).
- One-stage vs. two-stage: usually produce similar results.
- Occasionally they differ: different assumptions and/or different estimation methods.
- Use the PRISMA-IPD for reporting guidelines of IPD meta-analyses.

86

## References

- Statistical methods for a meta-analysis of IPD*
- Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: conduct, rationale and reporting. *BMJ* 2010; 340: c221
  - Higgins JP, Whitehead A, Turner RM, Omar RZ, Thompson SG. Meta-analysis of continuous outcome data from individual patients. *Stat Med* 2001; 20: 2219-4
  - Turner RM, Omar RZ, Yang M, Goldstein H, Thompson SG. A multilevel model framework for meta-analysis of clinical trials with binary outcomes. *Stat Med* 2000; 19: 3417-3432
  - Tudur-Smith C, Williamson PR, Mason AG. Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. *Stat Med* 2005; 24: 1307-1319
  - Crowther MJ, Riley RD, Stanssen JA, Wang J, Gueyffier F, Lambert PC. Individual patient data meta-analysis of survival data using Poisson regression models. *BMC MedResMeth* 2012; 12: 34
  - Riley RD, Kausar I, Blend M, Wang J, Gueyffier F, Thijssen J, Deeks JJ. Meta-analysis of continuous outcomes according to baseline imbalance and availability of individual participant data. *Stat Med* 2013; 32(16):2747-66. doi: 10.1002/sim.5728
- One-stage versus two-stage*
- Olkin I, Sampson A. Comparison of meta-analysis versus analysis of variance of individual patient data. *Biometrics* 1998; 54(1):317-22
  - Mathew T, Nondam K. On the equivalence of meta-analysis using literature and using individual patient data. *Biometrics* 1999; 55(4):1221-3
  - Debray TPA, Moons KGM, Abo-Zaid GMA, Koffijberg H, Riley RD (2013) Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? *PLoS ONE* 8(4): e60550
  - Stewart, G.B., et al., Statistical analysis of individual participant data meta-analyses: a comparison of methods and recommendations for practice. *PLoS One*, 2012. 7(10): p. e46042
  - Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage. <sup>87</sup> approaches, and why they may differ. *Stat Med* 2016 doi: 10.1002/sim.7141. [Epub ahead of print]

## **PRACTICAL 2: One-stage IPD meta-analysis**

Please follow the instructions in the Stata do files labelled Practical 2a, 2b and 2c

Practical 2(a): One-stage binary outcomes

Practical 2(b): One-stage continuous outcomes

Practical 2(c): One-stage survival outcomes

### **Learning objectives:**

- Gain experience of fitting one-stage IPD meta-analysis models.
- Analyse binary, continuous and survival data using IPD from multiple studies.
- Interpret Stata output after fitting IPD meta-analysis models.
- Gain experience of using key Stata commands, including ‘logistic’, ‘meqrlogit’, ‘regress’, ‘mixed’, ‘stcox’, and ‘stmixed’.
- Gain experience of computational and estimation issues for IPD meta-analysis.

## LECTURE 4: Special topics

### Summary:

So far we have focused mainly on IPD meta-analysis of randomised trials to summarise a treatment effect. However, one-stage and two-stage IPD meta-analysis approaches can be applied to other topic areas, such as diagnostic test studies, prognostic factor studies, and clinical prediction models. Here we illustrate this, and give examples that show why IPD is important in these settings. We also introduce two emerging topics: percentage study weights in a one-stage IPD meta-analysis and the power of a planned IPD meta-analysis. Percentage study weights are important to reveal the contribution of each study toward the overall meta-analysis result, but these are hidden within a one-stage analysis. We will describe how to obtain them, based on our recent paper in *Statistical Methods in Medical Research*, and compare them to two-stage weights. Statistical power is important to consider in advance of an IPD meta-analysis beginning, as we know IPD projects can be time-consuming and costly. We introduce the concept of simulation-based power calculations for IPD meta-analysis, and provide an example from a recent IPD meta-analysis project in pregnancy.

### Learning objectives:

- Understand how percentage study weights can be derived for a one-stage models,
- Understand why meta-analysis of test accuracy studies are best undertaken in a one-stage approach,
- Appreciate the common problems facing an aggregate data meta-analysis of prognostic factor studies,
- Understand why IPD offers many benefits for meta-analysis of prognostic factor studies, especially for modelling of continuous factors and adjustment factors,
- Appreciate that even an IPD meta-analysis of prognostic factor studies will encounter many problems, such as different methods of measurement, missing data, and poor quality primary studies,
- Recognise why IPD helps to validate clinical prediction models, in terms of checking their predictive performance across multiple settings,
- Understand the benefit of simulation-based power calculations for an IPD meta-analysis, and the general step-by-step process for performing these,
- Recognise that power calculations can be tailored for the IPD that is potentially available, and may help decide whether an IPD project is worth the effort.

### Key references:

1. Abo-Zaid G, Sauerbrei W, Riley RD. IPD meta-analysis of prognostic factor studies: state of the art? *BMC Med Res* 2012;12:56
2. Riley RD, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med* 2013; 10(2): e1001380.
3. Riley RD, Ensor J, Snell KIE, Debray TPA, Altman DG, Moons KGM, Collins GS. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. *BMJ* 2016;353:i3140
4. Kontopantelis E, Springate DA, Parisi R and Reeves D. Simulation-Based Power Calculations for Mixed Effects Modeling: ipdpower in Stata. *2016*. 2016; 74: 25
5. Riley RD, Ensor JE, Jackson D, Burke DL. Deriving percentage study weights in multi-parameter meta-analysis models: with application to meta-regression, network meta-analysis, and one-stage individual participant data models. *Stat Meth Med Res* (2017, in-press).



## LECTURE 4: Special Topics

**Prof Richard Riley**

Professor of Biostatistics  
Keele University

e-mail: r.riley@keele.ac.uk

## Aims

- Percentage study weights for one-stage models
- Cover novel applications of one-stage & two-stage meta-analyses outside of trials
  - diagnostic tests
  - prognostic factors
  - prognostic models
- Consider the power of an IPD meta-analysis

2

### Part 1:

#### Percentage study weights for one-stage and multi-parameter IPD models

3

#### % study weights

- % study weights reveal the contribution of each study
- Routinely presented on a forest plot
- Especially useful when:
  - some studies are at **high risk of bias**
  - some studies are potential **outliers**
  - need to check impact of **modelling assumptions**
- PRISMA recommends that:
  - "... it is preferable also to include, for each study, the numerical group-specific summary data, the effect size and confidence interval, and **the percentage weight**"

4

#### Weights for two-stage approach

- Pooled estimate: often use inverse-variance (MLE) solution
- e.g. for the random effects model:

$$\hat{\theta} = \frac{\sum_{i=1}^k \hat{\theta}_i w_i}{\sum_{i=1}^k w_i} \quad \text{where} \quad w_i = \frac{1}{\text{var}(\hat{\theta}_i) + \tau^2}$$

$$\% \text{ weight for study } i = 100\% \times \frac{w_i}{\sum_{i=1}^k w_i}$$

5

#### Weights for two-stage approach

- You all know this!
- But may not recognise that percentage study weights relates to a decomposition of the variance, via Fisher's information
- For independent studies, we can decompose the total Fisher information of the meta-analysis result using:

$$I_{total}(\hat{\theta}) = (\text{var}(\hat{\theta}))^{-1} = \sum_{i=1}^K w_i = \sum_{i=1}^K I_i(\hat{\theta})$$

6

## Weights for two-stage approach

- This allows us to decompose the variance

$$\begin{aligned}
 \text{var}(\hat{\theta}) &= \text{var}(\hat{\theta}) \times I_{\text{total}}(\hat{\theta}) \times \text{var}(\hat{\theta}) \\
 &= \text{var}(\hat{\theta}) \times \sum_{i=1}^K I_i(\hat{\theta}) \times \text{var}(\hat{\theta}) \\
 &= \text{var}(\hat{\theta}) \times \sum_{i=1}^K w_i \times \text{var}(\hat{\theta}) \\
 &= \sum_{i=1}^K \text{var}(\hat{\theta}) \times w_i \times \text{var}(\hat{\theta}) \\
 &= \sum_{i=1}^K W_i
 \end{aligned}$$

7

## Weights for two-stage approach

- So we have

$$W_i = \text{var}(\hat{\theta}) \times w_i \times \text{var}(\hat{\theta}) = \text{var}(\hat{\theta}) \times I_i(\hat{\theta}) \times \text{var}(\hat{\theta})$$

- Percentage study weights are thus

$$\% \text{ weight of study } i = 100\% \times \frac{W_i}{\sum_{i=1}^K W_i} = 100\% \times \frac{W_i}{\text{var}(\hat{\theta})}$$

8

## Weights for two-stage approach

- So we have

$$W_i = \text{var}(\hat{\theta}) \times w_i \times \text{var}(\hat{\theta}) = \text{var}(\hat{\theta}) \times I_i(\hat{\theta}) \times \text{var}(\hat{\theta})$$

- Percentage study weights are thus

$$\% \text{ weight of study } i = 100\% \times \frac{W_i}{\sum_{i=1}^K W_i} = 100\% \times \frac{W_i}{\text{var}(\hat{\theta})}$$

equivalent to:

$$\% \text{ weight of study } i = 100\% \times \frac{w_i}{\sum_{i=1}^K w_i}$$

9

## Weights for two-stage approach

- So we have

$$W_i = \text{var}(\hat{\theta}) \times w_i \times \text{var}(\hat{\theta}) = \text{var}(\hat{\theta}) \times I_i(\hat{\theta}) \times \text{var}(\hat{\theta})$$

- Percentage study weights are thus

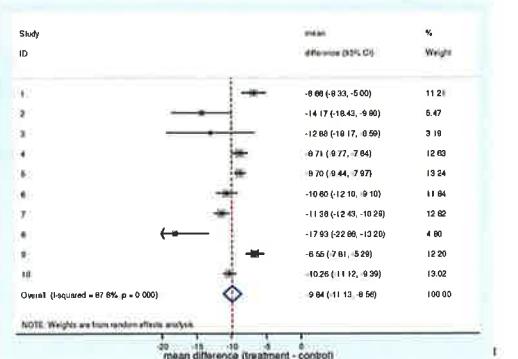
$$\% \text{ weight of study } i = 100\% \times \frac{W_i}{\sum_{i=1}^K W_i} = 100\% \times \frac{W_i}{\text{var}(\hat{\theta})}$$



But only this approach generalises to a one-stage (multi-parameter) model

10

## Example: meta-analysis of 10 hypertension trials



## One-stage IPD meta-analysis

- An alternative approach is a one-stage meta-analysis
- This analyses all the IPD from all studies simultaneously
- Basically a multi-level (mixed effects) regression model; e.g. linear mixed model:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

- Usually one-stage and two-stage approaches give very similar results ... but crucially not always (see lecture 3)!

12

## One-stage approach

Continuous outcomes: e.g. ANCOVA (random-effects model):

- Written in matrix notation for 3 studies, each with 2 patients

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

$$\begin{pmatrix} y_{F11} \\ y_{F12} \\ y_{F21} \\ y_{F22} \\ y_{F31} \\ y_{F32} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & y_{B11} & 0 & 0 & x_{11} \\ 1 & 0 & 0 & y_{B12} & 0 & 0 & x_{12} \\ 0 & 1 & 0 & 0 & y_{B21} & 0 & x_{21} \\ 0 & 1 & 0 & 0 & y_{B22} & 0 & x_{22} \\ 0 & 0 & 1 & 0 & 0 & y_{B31} & x_{31} \\ 0 & 0 & 1 & 0 & 0 & y_{B32} & x_{32} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \boldsymbol{\beta}_3 \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \boldsymbol{\theta} \end{pmatrix} + \begin{pmatrix} x_{11} & 0 & 0 & 0 & 0 & 0 & u_1 \\ 0 & x_{12} & 0 & 0 & 0 & 0 & u_1 \\ 0 & 0 & x_{21} & 0 & 0 & 0 & u_2 \\ 0 & 0 & 0 & x_{22} & 0 & 0 & u_2 \\ 0 & 0 & 0 & 0 & x_{31} & 0 & u_3 \\ 0 & 0 & 0 & 0 & 0 & x_{32} & u_3 \end{pmatrix} \begin{pmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \\ e_{31} \\ e_{32} \\ \boldsymbol{\epsilon} \end{pmatrix}$$

13

## One-stage approach

Continuous outcomes: e.g. ANCOVA (random-effects model):

- Written in matrix notation for 3 studies, each with 2 patients

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

$$\begin{pmatrix} y_{F11} \\ y_{F12} \\ y_{F21} \\ y_{F22} \\ y_{F31} \\ y_{F32} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & y_{B11} & 0 & 0 & x_{11} \\ 1 & 0 & 0 & y_{B12} & 0 & 0 & x_{12} \\ 0 & 1 & 0 & 0 & y_{B21} & 0 & x_{21} \\ 0 & 1 & 0 & 0 & y_{B22} & 0 & x_{22} \\ 0 & 0 & 1 & 0 & 0 & y_{B31} & x_{31} \\ 0 & 0 & 1 & 0 & 0 & y_{B32} & x_{32} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \boldsymbol{\beta}_3 \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \boldsymbol{\theta} \end{pmatrix} + \begin{pmatrix} x_{11} & 0 & 0 & 0 & 0 & 0 & u_1 \\ 0 & x_{12} & 0 & 0 & 0 & 0 & u_1 \\ 0 & 0 & x_{21} & 0 & 0 & 0 & u_2 \\ 0 & 0 & 0 & x_{22} & 0 & 0 & u_2 \\ 0 & 0 & 0 & 0 & x_{31} & 0 & u_3 \\ 0 & 0 & 0 & 0 & 0 & x_{32} & u_3 \end{pmatrix} \begin{pmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \\ e_{31} \\ e_{32} \\ \boldsymbol{\epsilon} \end{pmatrix}$$

Of key interest: the pooled treatment effect, but we must retain entire matrix when deriving weights

14

## One-stage approach

Continuous outcomes: e.g. ANCOVA (random-effects model):

- Written in matrix notation for 3 studies, each with 2 patients

$$\begin{pmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \\ e_{31} \\ e_{32} \end{pmatrix} \sim N(\mathbf{0}, \mathbf{R}) \quad \mathbf{R} = \begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_1^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_1^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_2^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_3^2 \end{pmatrix}$$

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{pmatrix} \sim N(\mathbf{0}, \mathbf{G}) \quad \mathbf{G} = \begin{pmatrix} \tau^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \tau^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \tau^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tau^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \tau^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \tau^2 \end{pmatrix}$$

15

## One-stage IPD meta-analysis

PRISMA-IPD says: "In common with the standard PRISMA Statement **the display of forest plots for key outcomes is advocated**, irrespective of the type of approach to statistical analysis"

- One-stage analysis provides meta-analysis results directly
- Study-specific results are hidden
- Thus not immediately amenable to forest plots
- One-stage weights are hidden**
- Further, one-stage models are multi-parameter problems
- Not immediately obvious how to calculate them

16

## ipdforest in Stata

The Stata Journal (2013)  
13, Number 3, pp. 574-587

### A short guide and a forest plot command (ipdforest) for one-stage meta-analysis

Ernesto Kortopardo  
NHS School for Primary Care Research  
Institute of Population Health  
University of Manchester  
Manchester, UK  
er.kortopardo@manchester.ac.uk

David Bannister  
Institute of Population Health  
University of Manchester  
Manchester, UK  
david.bannister@manchester.ac.uk

**Abstract:** In this article, we describe a new individual patient data meta-analysis estimation command, ipdforest. This command provides a forest plot displaying a one-stage meta-analysis with study-specific weights. (These commands have been developed to be used in Stata 11 and 12 and do not yet work in Stata 13, which is currently not compatible with the new syntax.) The overall effect is obtained from the preceding meta-analysis right down to study effects from least to the greatest weight. This command can be used to estimate the overall effect of a particular data meta-analysis model with Stata 12 documented.

**Keywords:** Stata 12, ipdforest, meta-analysis, forest plot, individual patient data, IPD, one-stage

"Patient weights are uniform; therefore, each study's weight is the ratio of its participants over the total number of participants across all studies"

17

## ipdforest in Stata

The Stata Journal (2013)  
13, Number 3, pp. 574-587

### A short guide and a forest plot command (ipdforest) for one-stage meta-analysis

Ernesto Kortopardo  
NHS School for Primary Care Research  
Institute of Population Health  
University of Manchester  
Manchester, UK  
er.kortopardo@manchester.ac.uk

David Bannister  
Institute of Population Health  
University of Manchester  
Manchester, UK  
david.bannister@manchester.ac.uk

**Abstract:** In this article, we describe a new individual patient data meta-analysis estimation command, ipdforest. This command provides a forest plot displaying a one-stage meta-analysis with study-specific weights. (These commands have been revised in Stata 13 to allow and negotiate, respectively; ipdforest is currently not compatible with the new syntax.) The overall effect is obtained from the preceding meta-analysis right down to study effects from least to the greatest weight. This command can be used to estimate the overall effect of a particular data meta-analysis model with Stata 12 documented.

**Keywords:** Stata 13, ipdforest, meta-analysis, forest plot, individual patient data, IPD, one-stage

"Patient weights are uniform; therefore, each study's weight is the ratio of its participants over the total number of participants across all studies"

- We disagree, and here is our new proposal ...

18

Article

**Deriving percentage study weights in multi-parameter meta-analysis models: with application to meta-regression, network meta-analysis and one-stage individual participant data models**

Richard D Riley,<sup>1</sup> Jole Ennor,<sup>1</sup> Dan Jackson<sup>2</sup> and Danielle L Burke<sup>1</sup>

**Abstract**  
Many meta analysis models contain multiple parameters, for example due to multiple outcomes, multiple treatments or multiple regression coefficients. In particular, meta-regression models may contain multiple study-level covariates, and one or more individual parameter. In meta-analysis models, we can compute study-specific weights for each parameter. Here, we propose how to derive percentage study weights for each individual parameter in order to reveal the (otherwise hidden) contribution of each study toward the parameter estimates of interest. We assume that studies are independent and use a decomposition of Fisher's information matrix to decompose the total variance matrix of parameter estimates into study-specific contributions, from which percentage weights are derived. This approach generalises how percentage weights are calculated in a traditional, single parameter meta-analysis model. Application is made to one- and two-stage individual participant data meta-analysis, and network meta-analysis, including the use of random treatment effects. These revised percentage study weights toward equally important estimates, such as summary treatment effects and treatment covariate interactions, and are especially useful when some studies are potential outliers or at high risk of bias. We also derive percentage study weights toward methodologically interesting measures, such as the magnitude of ecological bias (difference between within-study and across-study associations) and the amount of inconsistency (difference between direct and indirect evidence in a network meta-analysis).

19

## Recall two-stage weights

- We used a decomposition of the variance

$$W_i = \text{var}(\hat{\theta}) \times I_i(\hat{\theta}) \times \text{var}(\hat{\theta})$$

- Percentage study weights were thus:

$$\% \text{ weight of study } i = 100\% \times \frac{W_i}{\text{var}(\hat{\theta})}$$

We now mimic this for the one-stage approach but in matrix form

20

## In brief, we propose ...

- For the parameter in the one-stage model corresponding to row  $r$  of  $\hat{\beta}$  then

$$\% \text{ weight of study } i = 100\% \times \frac{\mathbf{W}_i(\hat{\beta})_{r,r}}{\sum_{i=1}^K \mathbf{W}_i(\hat{\beta})_{r,r}} = 100\% \times \frac{\mathbf{W}_i(\hat{\beta})_{r,r}}{\text{var}(\hat{\beta})_{r,r}}$$

where the ' $r,r$ ' notation refers to the element  $(r,r)$  of the corresponding matrix, and

$$\mathbf{W}_i(\hat{\beta}) = \text{var}(\hat{\beta}) \times I_i(\hat{\beta}) \times \text{var}(\hat{\beta})$$

and  $\text{var}(\hat{\beta})$  is the variance of  $\hat{\beta}$

- Weights also apply to meta-regression & multivariate m-a<sup>21</sup>

## A note on multiple parameters...

- Usually one-stage models contain multiple parameters

e.g. for blood pressure outcome & 10 IPD trials, we suggest a linear regression with:

- 10 intercepts (one per trial)
- 10 adjustment terms for baseline BP (one per trial)
- a treatment effect

- Therefore 21 parameters in  $\hat{\beta}$

- We can calculate % weights for each parameter
- They may be different!
- Stata not yet available; SAS Proc Mixed best option
- See Appendix A for more on one-stage weights

22

## EXAMPLE 1: Overall treatment effect

23

## IPD meta-analysis for overall effect

- IPD from 10 trials to reduce high systolic blood pressure
- Hypertension treatment versus control

### TWO-STAGE APPROACH:

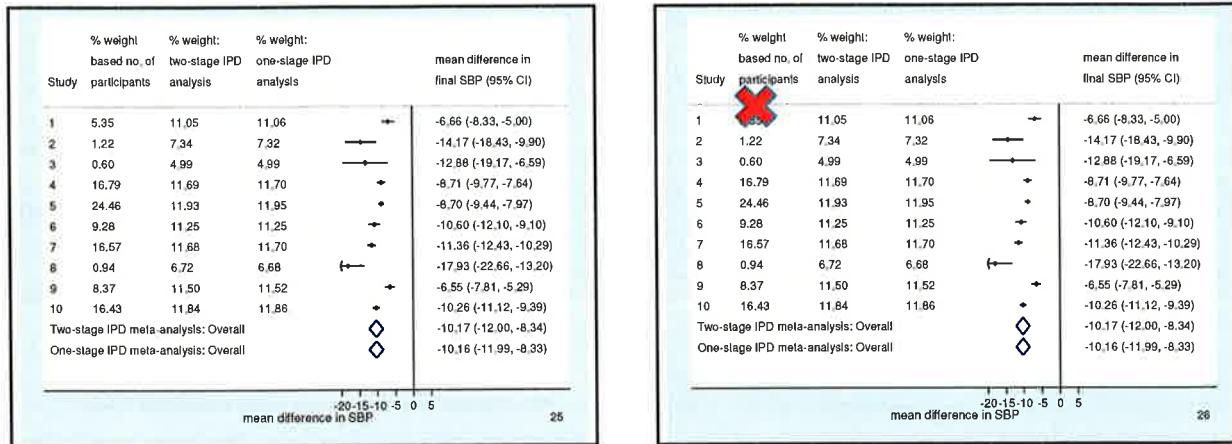
STEP 1: ANCOVA in each trial, to obtain mean difference in SBP at follow-up after adjusting for baseline

STEP 2: Random effects meta-analysis of the treatment effect estimates (using REML)

### ONE-STAGE APPROACH (REML):

ANCOVA with separate intercept, residual variance & baseline adjustment per trial, & random treatment effect

24



## EXAMPLE 2:

### Rare outcomes (diagnostic accuracy of a test)

27

 ELSEVIER

Journal of Clinical Epidemiology 61 (2008) 41–51

The binomial distribution of meta-analysis was preferred to model within-study variability

Taye H. Hamza<sup>a,b</sup>, Henk C. van Houwelingen<sup>b</sup>, Theo Stijnen<sup>a</sup>

<sup>a</sup>Department of Epidemiology and Biostatistics, Erasmus MC – University Medical Center Rotterdam, P.O.Box 2040, 3000 CA Rotterdam, The Netherlands

<sup>b</sup>Department of Medical Statistics and Biostatistics, Leiden University Medical Center, Leiden, The Netherlands

Accepted 21 March 2007

**Abstract**

**Objective:** When studies report proportions such as sensitivity or specificity, it is common to meta-analyze them using the Poisson distribution, which is often incorrect. This method assumes the within-study variability will be proportional to a model for the proportion, which may lead to bias for several reasons. Alternatively an exact likelihood approach based on the binomial within-study distribution can be used. This method can easily be performed in standard statistical packages. We investigate the performance of the standard method and the alternative approach.

**Study Design and Setting:** We compare the two approaches through a simulation study, in terms of bias, mean squared error, and coverage probabilities. We also study the effect of varying or specifying the between studies variance, the within study sample size, and the number of studies. The methods are illustrated using a published meta-analysis data set.

**Results:** The exact likelihood approach performs always better than the approximate approach and gives unbiased estimates. The coverage probability in particular for the profile likelihood is also reasonably acceptable. In contrast, the approximate approach has large bias with very poor coverage probabilities in many cases.

**Conclusion:** The exact likelihood approach is the method of preference and should be used whenever feasible. © 2008 Elsevier Inc. All rights reserved.

**Keywords:** Meta-analysis; Diagnostic test; Random effects; Sensitivity; Specificity; Exact likelihood

### Example (Hamza et al., 2008)

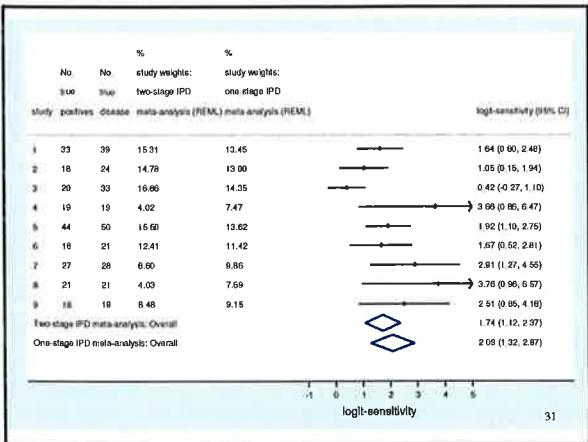
- An IPD meta-analysis to examine the accuracy of PET for diagnosis of Alzheimer's Disease
  - 9 studies identified containing diseased people with a binary response of test positive or test negative
- Sensitivity,  $p_i = \text{no. patients test positive}_i / \text{no. patients with disease}_i = r_i / n_i$
- We can usually get  $r$  and  $n$  from publications
  - This is essentially our IPD in this setting, when patient-level covariates are not of interest

29

### Example (Hamza et al., 2008)

- Very simple one-stage model needed
  - A random intercept logistic regression model
- $$r_i \sim \text{Binomial}(n_i, p_i)$$
- $$\text{logit}(p_i) = \beta_i \quad \beta_i \sim N(\beta, \tau^2)$$
- Pooled logit-sensitivity
- Small numbers of patients in each study, with zero cells in two of the studies ... not a problem for one-stage approach
  - But, an alternative two-stage approach requires continuity corrections in the first stage to obtain logit-sensitivity estimates for all studies, to then allow a random-effects meta-analysis model in second stage.

30



## Aside: Sensitivity & specificity

- Usually interested in both sensitivity & specificity

Sensitivity,  $p_{1i} = \text{no. patients test positive}_i / \text{no. patients with disease}_i = r_{11i} / n_{1i}$

Specificity,  $p_{0i} = \text{no. patients test positive}_i / \text{no. patients without disease}_i = r_{00i} / n_{0i}$

- Can extend the one-stage model to jointly examine sensitivity and specificity, whilst accounting for their between-study correlation
- More on this in the multivariate lecture tomorrow
- For completeness, model is also shown in Appendix B here

32

## Some references

- Riley R, Eensor J, Jackson D, Burke DL. Deriving percentage study weights in meta-analysis and meta-regression models: with application to interactions, network meta-analysis and one-stage IPD models (Stat Meth Med Res, in-press)
- Jackson D, White IR, Price M, Copas J and Riley RD. Borrowing of strength and study weights in multivariate and network meta-analysis. Stat Methods Med Res. 2016 (in-press).
- Fisher D. IPDMETAN: Stata module for performing two-stage IPD meta-analysis
- Kontopantelis E, Reeves D. IPDFOREST: Stata module to produce forest plot for IPD meta-analysis (one-stage)

33

## Part 2:

### IPD meta-analysis of prognostic factor studies

34

## Prognosis and prognostic factors

- The study of the risk (probability, likelihood) of future outcomes and events
- Greek: *πρόγνωσις* (fore-knowing, foreseeing)

### A prognostic factor (PF) is ...

- A biological, behavioural, symptomatic, psychological or environmental measure
- that, among people with a given health disease or condition (i.e. a startpoint), is *associated* with a future health outcome (i.e. an endpoint)

35

## What are prognostic factors?

### o Examples

Cancer:

Stage of disease is associated with a worse time to death or recurrence, & is thus a PF

Low back pain:

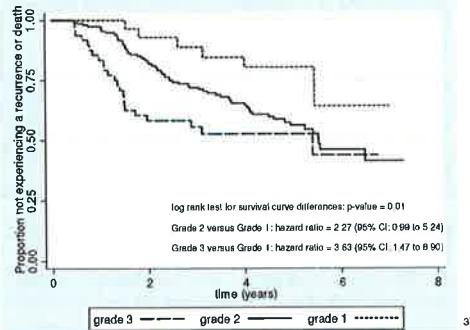
PFs for a worse outcome include co-morbid depression & higher levels of functional limitation

AIDS:

CD4 count is associated with disease progression, & is thus a PF

36

**Example:** In breast cancer, tumour grade is a PF as it is associated with a worse time to disease recurrence or death



37

OPEN ACCESS Freely available online

PLOS MEDICINE

Guidelines and Guidance

## Prognostic Research Strategy (PROGRESS) 2: Prognostic Factor Research

Richard D. Riley<sup>1,\*</sup>, Jill A. Hayden<sup>2†</sup>, Ewout W. Steyerberg<sup>3</sup>, Karel G. M. Moons<sup>4</sup>, Keith Abrams<sup>5</sup>, Panayiotis A. Kyza<sup>6</sup>, Núria Maitz<sup>7</sup>, Andrew Briggs<sup>8</sup>, Sara Schroter<sup>9</sup>, Douglas G. Altman<sup>10</sup>, Harry Hemingway<sup>11</sup>, for the PROGRESS Group

<sup>1</sup> School of Health and Population Sciences, University of Birmingham, Birmingham, United Kingdom; <sup>2</sup> Department of Community Health and Epidemiology, Dalhousie University, Halifax, Nova Scotia, Canada; <sup>3</sup> Department of Public Health, Erasmus MC, Rotterdam, Netherlands; <sup>4</sup> John雪花 Centre for Health Services and Primary Care, UMC Utrecht, Utrecht, Netherlands; <sup>5</sup> Centre for Biostatistics & Genetic Epidemiology, Department of Health Sciences, School of Medicine, University of Leicester, Leicester, United Kingdom; <sup>6</sup> Department of Oral and Maxillofacial Surgery, Hospital Universitario de La Princesa, Madrid, Spain; <sup>7</sup> Charles River Research, Boston, Massachusetts, United States; <sup>8</sup> National Cancer Research Institute, Bethesda, Maryland, United States; <sup>9</sup> Health Economics & Health Technology Assessment, Centre for Population & Health Sciences, University of Glasgow, Glasgow, United Kingdom; <sup>10</sup> MRC Epidemiology Unit, Medical School, Edinburgh, United Kingdom; <sup>11</sup> Centre for Statistics in Medicine, University of Oxford, Oxford, United Kingdom; <sup>\*</sup> Department of Epidemiology and Public Health, University College London, London, United Kingdom

Prognostic factor research aims to identify factors associated with subsequent clinical outcome in people with a particular disease or health condition. In this article, the second in the PROGRESS series, the authors discuss the role of prognostic factors in current clinical practice, randomised trials, and developing new interventions, and explain why and how prognostic factor research should be improved.

A prognostic factor is any measure that, among people with a

Prognostic factor research aims to discover and evaluate factors that might be useful as modifiable targets for interventions to improve outcomes, building blocks for prognostic models, or predictors of differential treatment response. Prognostic factor

Citation: Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyza P, Maitz N, Briggs A, Schroter S, Altman DG, Hemingway H (2013) Prognostic Research Strategy (PROGRESS) 2: Prognostic Factor Research. PLoS Med 10(2): e1001301. doi:10.1371/journal.pmed.1001301

38

## Why meta-analysis of PF studies?

- o Primary studies to identify PFs clearly important
- o Use of PFs ideally based on **overall evidence**
- o This is difficult because:
  - Large number of primary studies
  - Conflicting results
  - Small patient numbers
- o Formal evidence-based reviews and synthesis of PF studies needed
- o Meta-analysis (potentially) crucial in this context
  - *statistical synthesis of quantitative evidence*

39

## Meta-analysis using aggregate data

- o Traditional meta-analysis uses **aggregate data**
- o Obtainable from publications or study authors
- o Meta-analysis of PF studies usually requires data from each study:

**an estimate** of the relationship between the factor and outcome

e.g. hazard ratio for overall survival

**& the standard error** of this estimate

e.g. standard error of log hazard ratio

40

## Methods of Parmar et al. are essential for extracting HRs

STATISTICS IN MEDICINE  
 Statist. Med. 17, 2103–2331 (1998)

### EXTRACTING SUMMARY STATISTICS TO PERFORM META-ANALYSES OF THE PUBLISHED LITERATURE FOR SURVIVAL ENDPOINTS

MAHESH K. D. PARMAR<sup>1\*</sup>, VALTER TORRI<sup>2</sup> AND LESLEY STEWART<sup>1</sup>

<sup>1</sup> MRC Cancer Trials Office, 3 Shattock's Way, Cambridge, U.K.

<sup>2</sup> Istituto Mario Negri, Milan, Italy

#### SUMMARY

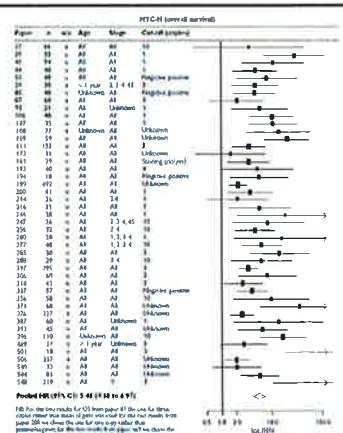
Meta-analyses aim to provide a full and comprehensive summary of related studies which have addressed a similar question. When the studies involve time-to-event survival-type data the most appropriate statistics to use are the log hazard ratio and its variance. However, these are not always explicitly presented for each study. In this paper a number of methods of extracting estimates of these parameters in a variety of situations are presented. One of these methods should improve the efficiency and reliability of meta-analyses of the published literature with survival-type endpoints. © 1998 John Wiley & Sons, Ltd.

41

**Example:**  
**Meta-analysis of PF studies using aggregate data**

Q: Is MYCN a PF for overall survival in neuroblastoma?

45 studies provided HR estimates via the methods of Parmar et al



## Meta-analysis issues: e.g. MYCN ...

- o Despite exhaustive attempts, only 35% of relevant studies provided hazard ratio estimates
  - o Poor reporting endemic; publication bias
  - o Large heterogeneity between-studies
    - Type of estimate, unadjusted and adjusted;
    - choice of adjustment factors
    - statistical analysis method
    - study design (retrospective versus prospective)
    - study quality (lots of small studies, no protocol)
    - purpose of study (often not clear)
    - treatments used; stage of disease
    - method of measurement; choice of cut-off level
- etc

43

## Is meta-analysis using IPD the solution?

### Example of IPD in prognostic factors in neuroblastoma:

no	Marker levels			Adjustment factors		Survival and disease status		
	TH	LDH	MYCN	Age	Stage	Time of recurrence	Final survival time	Final disease status
1	Pos	200	5	3 yrs	1		150 days	ALIVE
2	Neg	350	3	2 yrs	4	330 days	390 days	DEAD
3	Neg	120	1	2 yrs	3	230 days	250 days	ALIVE with disease
4	Neg	320	1	6 yrs	4	27 days	48 days	DEAD
...	...	...	...	...	...	...	...	...
								44

## General strategy for IPD meta-analysis of PF studies

- o Either one-stage or two-stage approach
- o As detailed earlier for IPD from trials, but now want prognostic effect of a factor (not treatment)

### Two-stage

- o Fit a logistic or Cox regression model in each study
- o Pool the logHR estimates using a fixed-effect or random-effects meta-analysis

### One-stage

- o Fit a logistic or Cox regression model to all studies combined, with a separate intercept (baseline hazard) per study and possibly random-effects on the prognostic effects in the model

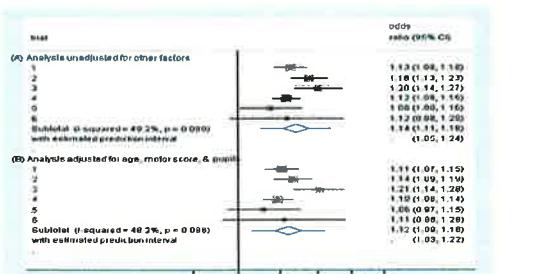
45

## Potential advantages of IPD

- o Calculate aggregate data directly in first stage
  - o Potentially longer follow-up time
  - o Check modelling assumptions (e.g. proportional hazards) and adapt if necessary (e.g. allow prognostic effect to vary: non-proportional hazards)
  - o Deal with & examine missing data
  - o Adjust for a more consistent set of variables
  - o Choose consistent cut-off level
  - o Analyse continuous variables on original scale
  - o Model non-linear trends
  - o Examine causes of heterogeneity
- etc

46

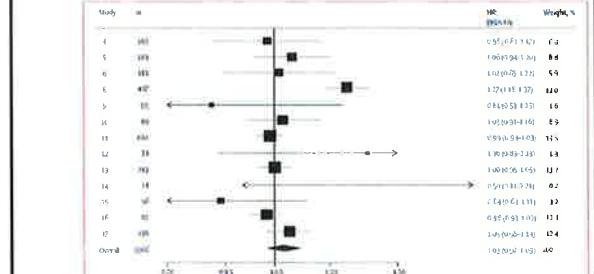
### Example: IPD meta-analysis in traumatic brain injury (Maas et al.) Is glucose a prognostic factor for poor outcome?



- IPD allows glucose to be modelled on its **continuous scale** (linear trend)
- IPD allows a **consistent set of adjustment factors**
- Results show that glucose has prognostic value in addition to age, motor score & pupil reactivity

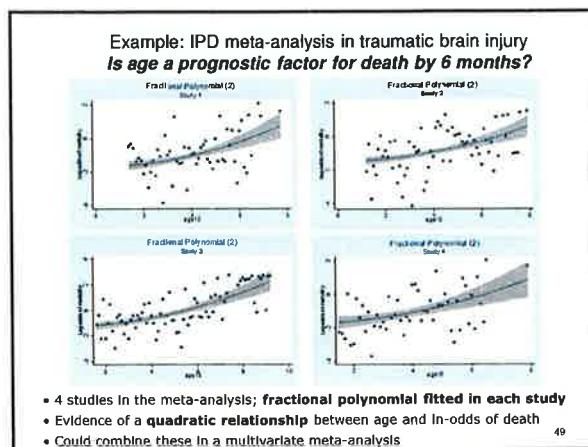
47

### Example: IPD meta-analysis in non-small-cell lung carcinoma (Trivelia et al.) Is microvessel density a prognostic factor for death?



- Microvessel density analysed as continuous
- IPD from **unpublished studies included**
- Results contradict previous meta-analysis of published summary data

48



49

Abu-Zaid et al. BMC Medical Research Methodology 2012, 12:56  
<http://www.biomedcentral.com/1471-2288/12/56>

**BMC Medical Research Methodology**  
Open Access

**RESEARCH ARTICLE**

**Individual participant data meta-analysis of prognostic factor studies: state of the art?**

Gözde Abu-Zaid<sup>1</sup>, Will Sauerbrei<sup>2</sup> and Richard D Riley<sup>3</sup>

**Abstract**

**Background:** Prognostic factor studies are often conducted individually with a great focus on the health condition. Although many studies include participants from different studies, there has been little effort to synthesize prognostic factor studies. We review the feasibility and conduct of this approach.

**Methods:** A systematic review to identify published IPD meta-analyses of prognostic factor studies, followed by detailed assessment of a random sample of 20 studies published from 2006. Six of them were from the IMPACT (International Prospective Individual Patient Data) study, in most cases in a prospective collaboration. For the remaining 14 studies, we used the IMPACT methodology to pool individual patient data.

**Results:** Four people published IPD meta-analyses of prognostic factors were identified up to March 2009. Only three were published after 2006 but illustrate a range of five studies within one year, and traumatic brain injury the most recent research field. A total of 11 IPD and many additional data sets, including modelling assumptions, analyses available on their continuous scale with the possibility of pooling for each level (individual, study, and study by study), were included. The IMPACT studies had the largest number of variables and the most complete data, and thus required the fewest IPD meta-analyses. For these studies, different sets of prognostic factors were used, and the validity of study methods of measurement. The IMPACT studies is a linking example, and had generally similar designs, methodology and treatment protocols. Therefore, standardization was always as high as possible. In contrast, the other studies had different designs, methodology and treatment protocols. Therefore, standardization was often compromised without losing publication bias and availability bias are often considered as being adequately explained.

**Conclusions:** IPD meta-analysis of prognostic factors is feasible and offers many advantages, as displayed mostly by the IMPACT studies. However, with respect to their inherent biological and methodological characteristics, and their conduct and therefore cannot be automatically assumed.

50

### Potential challenges even with IPD

- Poor quality IPD
- Time-consuming to obtain and clean IPD
- Dealing with missing data
- Dealing with systematically missing predictors (may not adjust for all variables in all studies?)
- Heterogeneity!
  - Different methods of measurement
  - Different outcome definitions
  - Different treatment strategies, etc
- Publication bias and unavailable IPD

**A rich area for methodology research!**

51

### Prospectively planned IPD meta-analyses

Thakkinstian et al note:

*"since the IPD meta-analysis is a retrospective collaboration, it is difficult to get clinical variables that have been assessed and measured using similar methods across all studies; standardization is best done as a prospective collaboration"*

- Prospectively planned analyses involve primary studies agreeing to pool their IPD at the end of their study
- Can agree common methodology standards and the same factors and endpoints to measure to substantially reduce heterogeneity

52

### Some references

- Abo-Zaid G et al. IPD meta-analysis of prognostic factor studies: state of the art? *BMC Med Res 2012;12:56*
- Kyzas PA, Loizou KT, Ioannidis JP. Selective reporting biases in cancer prognostic factor studies. *J Natl Cancer Inst 2005;97(14):1043-55.*
- Trivella M, et al. MicrovesSEL density as a prognostic factor in non-small-cell lung carcinoma: a meta-analysis of individual patient data. *Lancet Oncology. 2007(6):488-499.*
- Maas AI, et al. Prognosis and clinical trial design in traumatic brain injury: the IMPACT study. *J Neurotrauma 2007;24(2):232-8*
- Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: conduct, rationale and reporting. *BMJ 2010;340:c221*
- Riley RD, et al. Prognostic Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med 2013; 10(2): e1001380.*
- Sauerbrei W, et al. Evidence-based assessment and application of prognostic markers: the long way from single studies to meta-analysis. *Communications in Statistics 2006;35:1333-1342*

53

### A related area: clinical prediction models

- Clinical prediction models use multiple prognostic/risk factors ('predictors') to estimate the:
  - risk of developing future outcomes (prognostic model)
  - risk of a disease/condition being present (diagnostic model)
- Understanding individual risk is crucial to clinical decision-making
- Based on risk, individuals can be targeted for early preventative strategies and/or different treatment strategies
- Hundreds of models developed each year (often poorly)
- All require validation of their predictive performance (rarely done)

54

## Logistic prediction model example

$$\log \left( \frac{P_{\text{success}}}{1 - P_{\text{success}}} \right) = 2.66 + 1.48 \text{ IncompMisc} \\ - 1.63 \text{ NilBleeding} - 0.07 \text{ Age}$$

where  $P_{\text{success}}$  denotes the probability for a patient to have a successful expectant management. IncompMisc has value of 1 if the diagnosis at primary scan is incomplete miscarriage and 0 otherwise. NilBleeding is 1 if there is neither vaginal bleeding nor clots and 0 otherwise. Alternatively, the model can be represented in the following form for calculating the predictive probability for a patient to have a successful expectant management:

$$P_{\text{success}} = \frac{e^{2.66 + 1.48 \text{ IncompMisc} - 1.63 \text{ NilBleeding} - 0.07 \text{ Age}}}{1 + e^{2.66 + 1.48 \text{ IncompMisc} - 1.63 \text{ NilBleeding} - 0.07 \text{ Age}}}$$

Ref: Casikar J, et al. Prediction of successful expectant management of first trimester miscarriage: development and validation of a new mathematical model. *Aust N Z J Obstet Gynaecol.* 2013;53(1):58-63, 55

## IPD meta-analysis opportunity

- IPD from multiple studies (or multiple practices within large databases) provides a unique opportunity
- Can now develop models using large data (many thousands of patients)
- Immediately validate the developed model (even multiple times)
- Can also evaluate how to implement the model in new populations (e.g. does the model need to be recalibrated with the baseline risk/hazard in each population of interest?)
- This is a hot topic ...

56

## Overview of different questions and approaches for IPD meta-analysis in prediction research



### GUIDELINES AND GUIDANCE

#### Individual Participant Data (IPD) Meta-analyses of Diagnostic and Prognostic Modeling Studies: Guidance on Their Use

Thomas P. A. Debray,<sup>1,\*</sup> Richard D. Riley,<sup>2</sup> Alessandra M. Moons,<sup>3</sup> Jukkaus  
B. Reitsma,<sup>2</sup> Karel G. M. Moons,<sup>4</sup> Catherine W. G. M. Moons,<sup>5</sup> Mette H. M. Moons,<sup>6</sup>

<sup>1</sup>Julian Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands, <sup>2</sup>The Dutch Cochrane Center, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands, <sup>3</sup>Haaswijk Institute for Primary Care and Health Sciences, University of Groningen, Groningen, The Netherlands, <sup>4</sup>Erasmus MC, Rotterdam, The Netherlands, <sup>5</sup>Radboudumc Nijmegen, The Netherlands

\* Membership of the Cochrane IPD Meta-analysis Methods group is listed in the Acknowledgments.

<sup>1–5</sup>https://doi.org/10.1371/journal.pmed.1008791

### OPEN ACCESS

Citation: Debray TP, Riley RD, Moens AM, Reitsma JB, Moons KG, Moons AHM, et al. Cochrane IPD Meta-analysis Methods group (2018) Individual Participant Data (IPD) Meta-analysis of Diagnostic and Prognostic Modeling Studies: Guidance on Their Use. *PLOS Med* 15(1): e1008791. <https://doi.org/10.1371/journal.pmed.1008791>

57

## Examining discrimination and subgroups, and added value of markers

Journal of Biomathematics. <https://doi.org/10.1089/biom.2013.2810>. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial license, which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1089/biom.2013.2810

### Practice of Epidemiology

#### Assessing Risk Prediction Models Using Individual Participant Data From Multiple Studies

Lisa Pennell,<sup>1</sup> Stephen Kaplorge,<sup>1</sup> Ian R. White,<sup>2</sup> Simon G. Thompson,<sup>3</sup> Angela M. Wood,<sup>4</sup> and the Emerging Risk Factors Collaboration

<sup>1</sup> Correspondence to Dr. Angela M. Wood, Strategic Research Laboratory, Department of Public Health and Primary Care, School of Clinical Medicine, University of Cambridge, West Cambridge CB1 8RN, United Kingdom (e-mail: amw76@medsch.cam.ac.uk)

Initially submitted January 29, 2013; accepted for publication November 12, 2013.

58

## Cross-validation approach

### Statistics in Medicine

#### Research Article

Received 24 June 2012 Accepted 15 October 2012 Published online 11 December 2012 in Wiley Online Library

https://doi.org/10.1002/sim.5232

## A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis

Thomas P. A. Debray,<sup>1,\*</sup> Karel G. M. Moons,<sup>2</sup> Ikhlaaq Ahmed,<sup>3</sup> Hendrik Koffijberg,<sup>4</sup> and Richard David Riley<sup>5</sup>

The use of individual participant data (IPD) from multiple studies is an exciting popular approach to developing and validating clinical prediction models, however, typically data in longitudinal designs, such as time-to-event data, has driven the adoption of such methods approaches for prognostic modeling with heterogeneity between study populations. Although these approaches are strong predictors of survival and other outcomes, they have been limited to the development of models within individual or study populations outside the derivation data. We consider several approaches to develop a stable model that can be used to predict an outcome in an external dataset. We also propose strategies for choosing a valid model to test for when the model is to be validated or applied to new individuals or study populations. These strategies are implemented by the IPD-MDA software to facilitate model development and validation. We illustrate these approaches using a real-world example of a prognostic model lacking strong internal external cross-validation and extend our framework to account for time-to-event data for model external validation. Our results demonstrate that the proposed approach to prognostic modeling, which utilizes cross-validation, is able to predict the model's performance in a population that is not represented by included studies. In summary, our framework allows the development of strong internal and external validation of a single, integrated prediction model from an IPD-MDA to achieve improved model performance and generalizability. Copyright © 2013 John Wiley & Sons, Ltd.

Keywords: prediction research, IPD, prediction models, meta-analysis, time-to-event, survival analysis, individual participant data (IPD), external validation

59

## Our recent BMJ paper on external validation

### RESEARCH METHODS AND REPORTING

#### OPEN ACCESS



## External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges

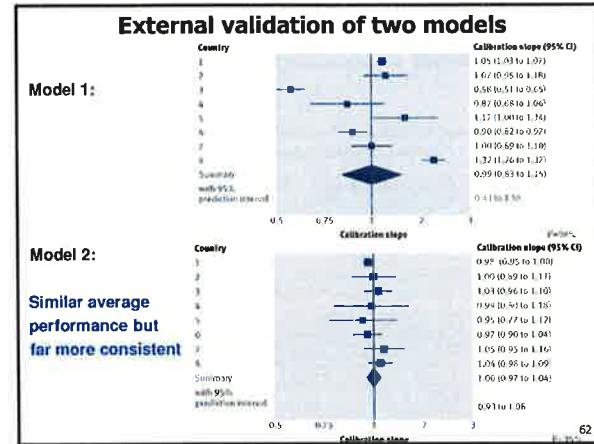
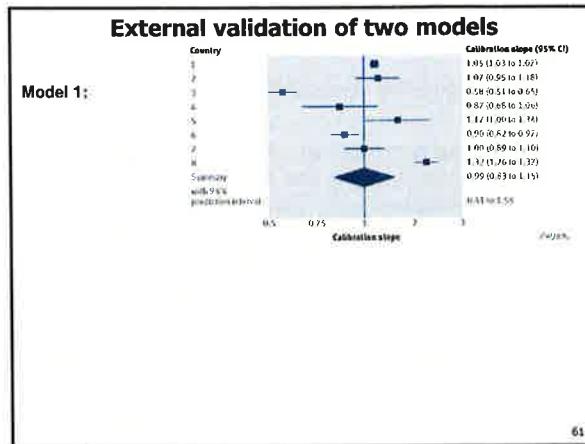
Richard D Riley,<sup>1</sup> Joie Ensor,<sup>1</sup> Kym E Snell,<sup>2</sup> Thomas P A Debray,<sup>1,4</sup> Doug G Altman,<sup>3</sup> Karel G M Moons,<sup>1,4</sup> Gary S Collins<sup>5</sup>

<sup>1</sup>Medical Institute for Technology and Health Sciences, Radboud University, Nijmegen, The Netherlands, <sup>2</sup>Salford Royal NHS Foundation Trust, Salford, UK, <sup>3</sup>Sussex Medical School, University of Brighton, Brighton, UK, <sup>4</sup>Department of Primary Care and Population Sciences, University College London, London, UK, <sup>5</sup>Nuffield Centre for Health Sciences, University of Oxford, University of Oxford, Oxford, UK, <sup>6</sup>Utrecht University, Utrecht, The Netherlands, <sup>7</sup>Centre for Statistical Bioinformatics, University of Leiden, Leiden, The Netherlands, <sup>8</sup>Centre for Statistics in Medicine, University of Oxford, Oxford, UK

Access to big datasets from e-health records and individual participant data (IPD) meta-analysis is signaling a new advent of external validation studies for clinical prediction models. In this article, the authors illustrate novel opportunities for external validation in big, combined datasets, while drawing attention to methodological challenges and reporting issues.

based ratios). Well-known examples are the Framingham risk score and QRSQ20+, which estimate the 10-year risk of developing cardiovascular disease; the Nottingham prognostic index, which predicts the five-year survival probability of a woman with newly diagnosed breast cancer;<sup>1,2</sup> and the Wells score for predicting the presence of a pulmonary embolism.<sup>3,4</sup> In the last decade, the use of big datasets and methods of machine learning for building prediction models have increased.<sup>5,6</sup> These articles all emphasized three fundamental components of prediction model research: model development, external validation, and impact evaluation.

60



## Part 4:

### Power of an IPD meta-analysis

63

### Rationale for power calculations

- IPD meta-analyses often time-consuming & costly
- Researchers and funders should consider the statistical power of planned IPD projects.
- Is it worth the effort?

e.g. If it was known in advance that IPD from a particular number of studies would only increase power to 50%, then researchers and funders may not proceed.

- Conversely, if a potential IPD meta-analysis increases power to over 80%, then funders will be more reassured that the project is worth resourcing.<sup>84</sup>

### Rationale for power calculations

- Power calculations are non-trivial
- Power depends on the choice and specification of analysis model (e.g. covariates to be included, number of parameters, magnitude of effects), and the parameter estimation method, amongst other factors.
- Simple algebraic solutions are not straightforward unless simplifying conditions are made, as proposed by Kovalchik

Kovalchik SA. Aggregate-data estimation of an individual patient data linear random effects meta-analysis with a patient covariate-treatment interaction term. *Biostatistics*. 2013; 14: 273-83.

Kovalchik SA and Cumberland WG. Using aggregate data to estimate the standard error of a treatment-covariate interaction in an individual patient data meta-analysis. *Biom J*. 2012; 54: 370-84.

### ipdpower

- Kontopantelis et al. propose simulation-based power calculations utilising a one-stage IPD meta-analysis framework
- Excellent package 'ipdpower' in Stata
- User inputs number of studies, number of patients, effect sizes and heterogeneity
- Module simulates potential IPD meta-analyses thousands of times, and evaluates power across results
- Very clever idea (avoids needed to know a formula)
- In Stata: ssc install ipdpower

66

**Journal of Statistical Software**  
October 2016, Volume 71, Issue 12.  
doi:10.1863/jss.v071.i12

**Simulation-Based Power Calculations for Mixed Effects Modeling: ipdpower in Stata**

Evangelos Kontopantelis  
NIHR School for Primary Care Research  
University of Manchester

David A Springate  
University of Manchester

Brent Pritchard  
University of Manchester

David Bannerman  
University of Manchester

**Abstract**  
Standard power calculations are practical and straightforward to power calculations, especially for model-level model. Mixed models often have the analysis which can be very messy. In addition, power calculations are model specific and under certain mixed effects models are often not available. ipdpower is a Stata command for power calculations for mixed effects models. ipdpower is new, considerably reduced command than ipdpower from Stata's mixed effects command. Also, ipdpower is designed to be used with individual patient data (IPD) meta-analysis and primary care databases such as PCD. In total, when patients are nested within studies, and general practice respectively, the methods apply to any two-level structure.

**Keywords:** Stata, ipdpower, power, coverage, meta-analysis, multilevel, mixed effects, random effects, individual patient data, IPD, primary care databases, PCD.

67

## ipdpower

- However, ipdpower is sometimes limited by:
  - the computational difficulty of fitting one-stage models (i.e. it can be very slow and not run certain models)
  - not removing ecological bias (see tomorrow for more on this)
  - not having methods such as Hartung-Knapp to inflate confidence intervals

e.g. in help file: with multiple random effects  
"non-convergence is more frequent than convergence."

68

### Simulation using a two-stage approach

- Ensor et al. (submitted) propose a two-stage approach
  - Very quick
  - Removes ecological bias
  - Utilises methods such as Hartung-Knapp to inflate confidence intervals
  - Code available in forthcoming paper
  - But can be tailored for each particular question

69

### Simulation using a two-stage approach

- Step by step guide
  - specify an underlying (data generating) statistical model for trials in the IPD meta-analysis (e.g. ANCOVA model);
  - use readily available information (e.g. from publications) and prior knowledge (e.g. number of studies promising IPD, number of patients within) to specify model parameter values
    - e.g. control group mean, intervention effect, treatment-covariate interaction, magnitude of heterogeneity, size of residual variances

70

### Simulation using a two-stage approach

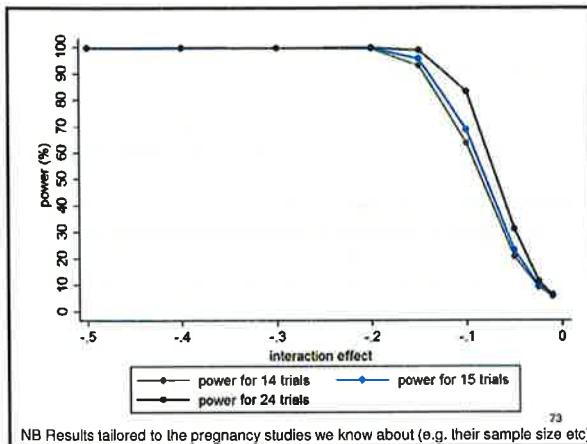
- simulate an IPD meta-analysis dataset of a particular size from the model, and apply a two-stage IPD meta-analysis to obtain the summary estimate of interest (e.g. interaction effect) and its associated *p*-value;
- repeat the previous step (e.g. thousands of times), then estimate the power to detect a genuine effect by the proportion of summary estimates with a significant *p*-value.

71

### Example

- An IPD meta-analysis of lifestyle interventions in pregnancy to reduce weight gain
  - 14 trials (1183 patients) promised their IPD
  - Aggregate data available to inform parameter values
  - What is the power to detect a treatment-BMI interaction?
- Using our simulation-based approach, we find
  - <60% power to detect a treatment effect improvement of 1kg for every 10 unit increase in BMI.
  - IPD from ten additional trials would improve power to over 80%, but only if heterogeneity in interaction is negligible.

72



### Summary of Special Topics Lecture

- Covered novel topics beyond trials & the 'basics'
  - One-stage and two-stage IPD models can be used to examine test accuracy, prognostic effects, model performance, and other pertinent measures
  - New methods emerging to derive percentage weights in complex one-stage models
  - Simulation-based methods like ipdpower and Ensor work are exciting ways to evaluate power in advance of an IPD meta-analysis
- (could save two years of your life!)

74

### Appendix A: One-stage weights

75

### Utilise Fisher's information matrix

- Let  $\hat{\beta}$  be the vector of parameter estimates from 1-stage m-a
- $I_{total}(\hat{\beta})$  is the total Fisher information matrix  
(the inverse of  $\text{var}(\hat{\beta})$  from the meta-analysis)

76

### Utilise Fisher's information matrix

- Let  $\hat{\beta}$  be the vector of parameter estimates from 1-stage m-a
- $I_{total}(\hat{\beta})$  is the total Fisher information matrix  
(the inverse of  $\text{var}(\hat{\beta})$  from the meta-analysis)
- Assume total Fisher information matrix is the sum of the study-specific independent information matrices, after holding fixed variance terms

$$I_{total}(\hat{\beta}) = \sum_{i=1}^K I_i(\hat{\beta})$$

77

### Utilise Fisher's information matrix

- Let  $\hat{\beta}$  be the vector of parameter estimates from 1-stage m-a
- $I_{total}(\hat{\beta})$  is the total Fisher information matrix  
(the inverse of  $\text{var}(\hat{\beta})$  from the meta-analysis)
- Assume total Fisher information matrix is the sum of the study-specific independent information matrices, after holding fixed variance terms

$$\text{where } I_{total}(\hat{\beta}) = \sum_{i=1}^K I_i(\hat{\beta})$$

$$I_i(\hat{\beta}) = (\mathbf{X}_i^T \mathbf{V}_i^{-1} \mathbf{X}_i)$$

$\mathbf{X}_i$  is the block of the design matrix from study  $i$  and  $\mathbf{V}_i$  contains residual variances and taus as from full meta-analysis

- We utilise this decomposition, to decompose  $\text{var}(\hat{\beta})$  itself

78

### Our proposal: decompose variance

$$\begin{aligned}
 \text{Decompose variance: } \mathbf{W}_{total}(\hat{\beta}) &= \mathbf{V}_{total}(\hat{\beta}) \\
 &= \mathbf{V}_{total}(\hat{\beta}) \times \mathbf{I}_{total}(\hat{\beta}) \times \mathbf{V}_{total}(\hat{\beta}) \\
 &= \sum_i \mathbf{V}_{total}(\hat{\beta}) \times \mathbf{I}_i(\hat{\beta}) \times \mathbf{V}_{total}(\hat{\beta}) \\
 &= \sum_i \mathbf{W}_i(\hat{\beta})
 \end{aligned}$$

79

### Our proposal: decompose variance

$$\begin{aligned}
 \text{Decompose variance: } \mathbf{W}_{total}(\hat{\beta}) &= \mathbf{V}_{total}(\hat{\beta}) \\
 &= \mathbf{V}_{total}(\hat{\beta}) \times \mathbf{I}_{total}(\hat{\beta}) \times \mathbf{V}_{total}(\hat{\beta}) \\
 &= \sum_i \mathbf{V}_{total}(\hat{\beta}) \times \mathbf{I}_i(\hat{\beta}) \times \mathbf{V}_{total}(\hat{\beta})
 \end{aligned}$$

**Step 1:** Fit the full meta-analysis model

$$- \text{Obtain } \mathbf{I}_{total}(\hat{\beta}) \text{ and } \mathbf{V}_{total}(\hat{\beta}) = (\mathbf{I}_{total}(\hat{\beta}))^{-1}$$

80

### Our proposal: decompose variance

$$\begin{aligned}
 \text{Decompose variance: } \mathbf{W}_{total}(\hat{\beta}) &= \mathbf{V}_{total}(\hat{\beta}) \\
 &= \mathbf{V}_{total}(\hat{\beta}) \times \mathbf{I}_{total}(\hat{\beta}) \times \mathbf{V}_{total}(\hat{\beta}) \\
 &= \sum_i \mathbf{V}_{total}(\hat{\beta}) \times \mathbf{I}_i(\hat{\beta}) \times \mathbf{V}_{total}(\hat{\beta})
 \end{aligned}$$

**Step 1:** Fit the full meta-analysis model

$$- \text{Obtain } \mathbf{I}_{total}(\hat{\beta}) \text{ and } \mathbf{V}_{total}(\hat{\beta}) = (\mathbf{I}_{total}(\hat{\beta}))^{-1}$$

**Step 2:** Obtain information due to study  $i$  using

$$\mathbf{I}_i(\hat{\beta}) = (\mathbf{X}_i^T \mathbf{V}_i^{-1} \mathbf{X}_i)$$

↑  
total variance matrix of the response ( $\mathbf{Y}$ ) in study  $i$

81

### Our proposal: decompose variance

$$\begin{aligned}
 \text{Decompose variance: } \mathbf{W}_{total}(\hat{\beta}) &= \mathbf{V}_{total}(\hat{\beta}) \\
 &= \mathbf{V}_{total}(\hat{\beta}) \times \mathbf{I}_{total}(\hat{\beta}) \times \mathbf{V}_{total}(\hat{\beta}) \\
 &= \sum_i \mathbf{V}_{total}(\hat{\beta}) \times \mathbf{I}_i(\hat{\beta}) \times \mathbf{V}_{total}(\hat{\beta})
 \end{aligned}$$

**Step 1:** Fit the full meta-analysis model

$$- \text{Obtain } \mathbf{I}_{total}(\hat{\beta}) \text{ and } \mathbf{V}_{total}(\hat{\beta}) = (\mathbf{I}_{total}(\hat{\beta}))^{-1}$$

**Step 2:** Obtain information due to study  $i$  using

$$\mathbf{I}_i(\hat{\beta}) = (\mathbf{X}_i^T \mathbf{V}_i^{-1} \mathbf{X}_i)$$

**Step 3:** Calculate the total weight due to study  $i$

$$\mathbf{W}_i(\hat{\beta}) = \mathbf{V}_{total}(\hat{\beta}) \times \mathbf{I}_i(\hat{\beta}) \times \mathbf{V}_{total}(\hat{\beta})$$

82

### Our proposal: decompose variance

**Step 4:** Recognise that:

$$\mathbf{W}_i(\hat{\beta}) = \mathbf{V}_{total}(\hat{\beta}) \times \mathbf{I}_i(\hat{\beta}) \times \mathbf{V}_{total}(\hat{\beta})$$

$$\mathbf{W}_{total}(\hat{\beta}) = \sum_i \mathbf{W}_i(\hat{\beta})$$

- Diagonal entries of  $\mathbf{W}_{total}(\hat{\beta})$  are sum of diagonal entries of  $\mathbf{W}_i(\hat{\beta})$

83

### Our proposal: decompose variance

**Step 4:** Recognise that:

$$\mathbf{W}_i(\hat{\beta}) = \mathbf{V}_{total}(\hat{\beta}) \times \mathbf{I}_i(\hat{\beta}) \times \mathbf{V}_{total}(\hat{\beta})$$

$$\mathbf{W}_{total}(\hat{\beta}) = \sum_i \mathbf{W}_i(\hat{\beta})$$

- Diagonal entries of  $\mathbf{W}_{total}(\hat{\beta})$  are sum of diagonal entries of  $\mathbf{W}_i(\hat{\beta})$

Thus, for a chosen parameter within  $\hat{\beta}$ , say  $\hat{\theta}$ , we take the corresponding diagonal entries of  $\mathbf{W}_{total}(\hat{\beta})$  and  $\mathbf{W}_i(\hat{\beta})$  to give

$$\begin{aligned}
 &\% \text{ of total weight due to study } i \\
 &= 100\% \times \frac{W_i(\hat{\theta})}{W_{total}(\hat{\theta})} = 100\% \times \frac{W_i(\hat{\theta})}{\text{var}(\hat{\theta})}
 \end{aligned}$$

84

## Appendix B: One-stage bivariate meta-analysis of test accuracy studies

85

	Group 1 Diseased	Group 0 Healthy
Test +	$r_{11i}$	$r_{10i}$
Test -	$r_{01i}$	$r_{00i}$

**Within-studies**  
 $r_{11i} \sim \text{Binomial}(n_{1i}, p_{1i}) \quad r_{00i} \sim \text{Binomial}(n_{0i}, p_{0i})$

**Between-studies**

$$\text{logit}(p_{1i}) = \beta_{1i}$$

$$\beta_{1i} = \beta_1 + u_{1i}$$

*Mean logit-sensitivity  
across studies*

$$\text{logit}(p_{0i}) = \beta_{0i}$$

$$\beta_{0i} = \beta_0 + u_{0i}$$

*Mean logit-specificity  
across studies*

86

	Group 1 Diseased	Group 0 Healthy
Test +	$r_{11i}$	$r_{10i}$
Test -	$r_{01i}$	$r_{00i}$

**Within-studies**  
 $r_{11i} \sim \text{Binomial}(n_{1i}, p_{1i}) \quad r_{00i} \sim \text{Binomial}(n_{0i}, p_{0i})$

**Between-studies**

$$\text{logit}(p_{1i}) = \beta_{1i}$$

$$\beta_{1i} = \beta_1 + u_{1i}$$

$$\text{logit}(p_{0i}) = \beta_{0i}$$

$$\beta_{0i} = \beta_0 + u_{0i}$$

**Study-level covariates can also be introduced here,  
e.g. Measurement device, to explain heterogeneity**

87

	Group 1 Diseased	Group 0 Healthy
Test +	$r_{11i}$	$r_{10i}$
Test -	$r_{01i}$	$r_{00i}$

**Within-studies**  
 $r_{11i} \sim \text{Binomial}(n_{1i}, p_{1i}) \quad r_{00i} \sim \text{Binomial}(n_{0i}, p_{0i})$

**Between-studies**

$$\text{logit}(p_{1i}) = \beta_{1i}$$

$$\beta_{1i} = \beta_1 + u_{1i}$$

$$\text{logit}(p_{0i}) = \beta_{0i}$$

$$\beta_{0i} = \beta_0 + u_{0i}$$

$$\begin{pmatrix} u_{1i} \\ u_{0i} \end{pmatrix} \sim N \begin{pmatrix} (0) \\ (0) \end{pmatrix}, \Omega = \begin{pmatrix} \tau_1^2 & \tau_1 \tau_0 \rho \\ \tau_1 \tau_0 \rho & \tau_0^2 \end{pmatrix}$$

88

	Group 1 Diseased	Group 0 Healthy
Test +	$r_{11i}$	$r_{10i}$
Test -	$r_{01i}$	$r_{00i}$

**Within-studies**  
 $r_{11i} \sim \text{Binomial}(n_{1i}, p_{1i}) \quad r_{00i} \sim \text{Binomial}(n_{0i}, p_{0i})$

**Between-studies**

$$\text{logit}(p_{1i}) = \beta_{1i}$$

$$\beta_{1i} = \beta_1 + u_{1i}$$

$$\text{logit}(p_{0i}) = \beta_{0i}$$

$$\beta_{0i} = \beta_0 + u_{0i}$$

$$\begin{pmatrix} u_{1i} \\ u_{0i} \end{pmatrix} \sim N \begin{pmatrix} (0) \\ (0) \end{pmatrix}, \Omega = \begin{pmatrix} \tau_1^2 & \tau_1 \tau_0 \rho \\ \tau_1 \tau_0 \rho & \tau_0^2 \end{pmatrix}$$

between-study heterogeneity

89

	Group 1 Diseased	Group 0 Healthy
Test +	$r_{11i}$	$r_{10i}$
Test -	$r_{01i}$	$r_{00i}$

**Within-studies**  
 $r_{11i} \sim \text{Binomial}(n_{1i}, p_{1i}) \quad r_{00i} \sim \text{Binomial}(n_{0i}, p_{0i})$

**Between-studies**

$$\text{logit}(p_{1i}) = \beta_{1i}$$

$$\beta_{1i} = \beta_1 + u_{1i}$$

$$\text{logit}(p_{0i}) = \beta_{0i}$$

$$\beta_{0i} = \beta_0 + u_{0i}$$

$$\begin{pmatrix} u_{1i} \\ u_{0i} \end{pmatrix} \sim N \begin{pmatrix} (0) \\ (0) \end{pmatrix}, \Omega = \begin{pmatrix} \tau_1^2 & \tau_1 \tau_0 \rho \\ \tau_1 \tau_0 \rho & \tau_0^2 \end{pmatrix}$$

between-study correlation

90



## **LECTURE 5: Using IPD meta-analysis to identify effect modifiers**

### **Summary:**

Compared to an aggregate data meta-analysis, an IPD meta-analysis can potentially produce more clinically relevant results, going beyond the ‘grand mean’ toward individualised or stratified medicine. For example, subgroups of patients with a common characteristic (e.g. those female) can be identified within IPD, and thus meta-analysis results derived specifically for them, with increased power compared to the individual studies themselves. Furthermore, IPD allows one to more powerfully and reliably examine differential treatment effects across individuals (‘effect modifiers’), as one can directly utilise within-trial information to estimate how patients’ characteristics modify treatment benefit (‘treatment-covariate interactions’). In this lecture, we describe how to use one-stage and two-stage IPD meta-analysis approaches for estimating interactions and effect modifiers, and stress the importance of avoiding ecological bias by separating out within-trial and across-trial interactions.

### **Learning objectives:**

- Understand the rationale for stratified medicine
- Recognise the meaning of the terms effect modifier and treatment-covariate interaction
- Understand how to conduct a two-stage or one-stage IPD meta-analysis for estimating treatment-covariate interactions within a linear, logistic and Cox regression framework
- Understand the shortcomings of a meta-regression of aggregate data when IPD are not available, and why the IPD approach is preferred
- Appreciate the meaning of the term ‘ecological bias’ and understand how to remove this issue within a one-stage IPD analysis

### **Key references:**

1. Berlin JA, et al. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Stat Med* 2002;21(3):371-87
2. Lambert PC, et al. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *J Clin Epidemiol* 2002;55(1):86-94.
3. Turner RM, Omar RZ, Yang M, Goldstein H, Thompson SG: A multilevel model framework for meta-analysis of clinical trials with binary outcomes. *Stat Med* 2000, 19: 3417-3432.
4. Tudur-Smith C, Williamson PR, Marson AG: Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. *Stat Med* 2005, 24: 1307-1319.
5. Riley RD, et al. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. *Stat Med* 2008; 27: 1870-93
6. Fisher, D. J., et al.. A critical review of methods for the assessment of patient-level interactions in individual participant data meta-analysis of randomized trials, and guidance for practitioners." *J Clin Epidemiol* 2011; 64(9): 949-967.

