A random walk through outcomes based research for healthcare

Jamie Kirkham (University of Liverpool)

Acknowledgements: *Doug Altman*
Mike Clarke
Carrol Gamble
Richard Riley
Paula Williamson …

Why are outcomes important?

- Interventions are compared in RCTs by measuring differences in patient outcomes between the groups

- Selection of **appropriate** outcomes is crucial

- ‘Clinical trials are only as credible as their outcomes’ (Tugwell 1993)

Hubs for Trials Methodology Research

Outcomes:

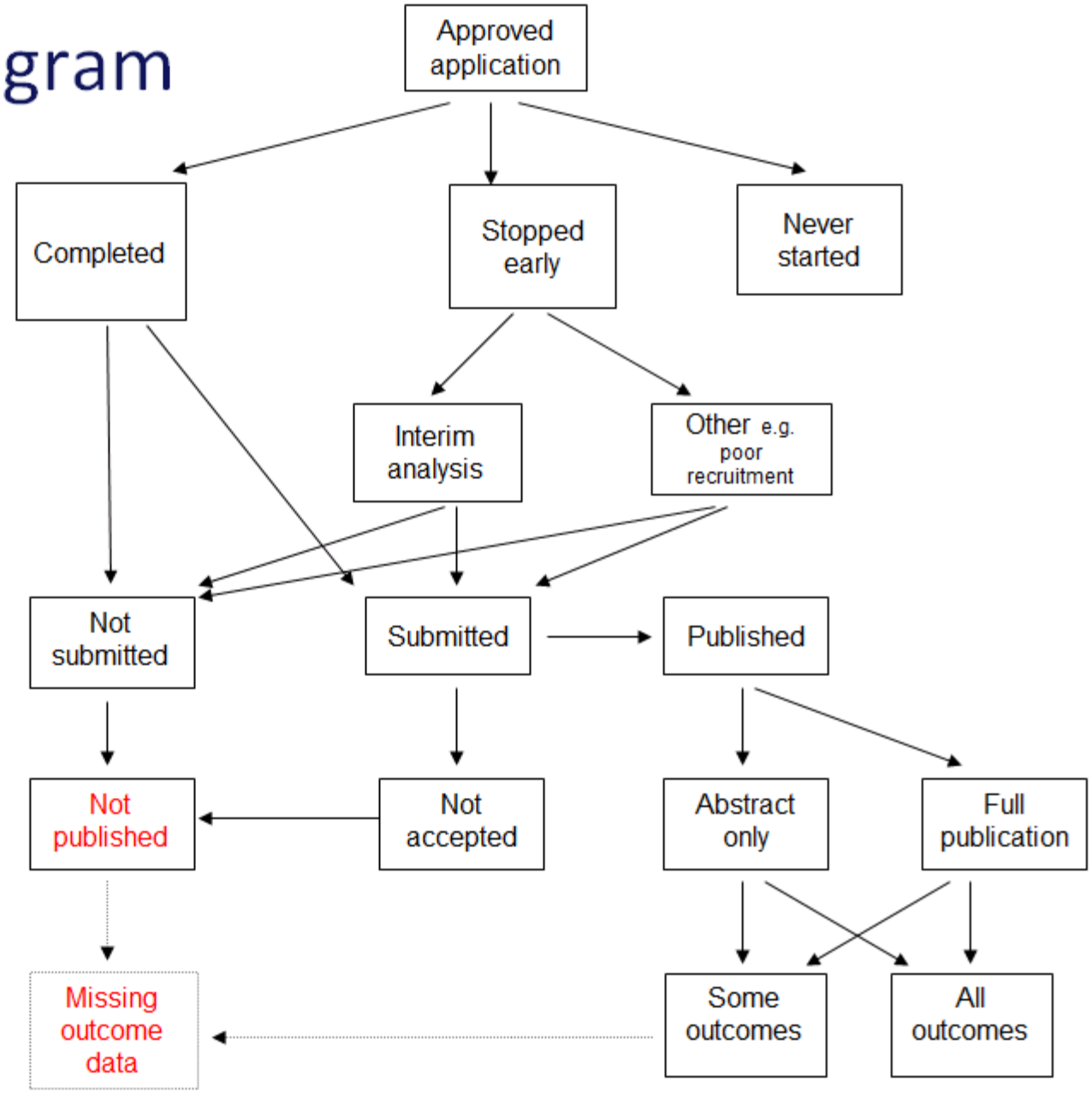
- Core outcome sets
- Ordinal
- Surrogate
- Harms
- Patient orientated
- Outcome reporting

MRC | Hubs for Trials Methodology Research

DMARD trials for Rheumatoid Arthritis

Study	Year	Tender Joints	Swollen Joints	Pain	Patient Global	Physician Global	Function	Acute Phase Reactant	Radiological Damage
ERC	1960	✓	✓	✓	✓	✓	✓	✓	✓
CCC	1973	✓	✓				✓	✓	✓
Huskisson	1976								
Woodland	1981	✓	✓	✓	✓		✓	✓	✓
Palmer	1982	✓	✓		✓		✓	✓	
Ward	1983		✓		✓	✓	✓	✓	
Williams	1983	✓	✓	✓	✓	✓	✓	✓	
Skosey	1988	✓	✓		✓	✓	✓	✓	
Morgan	1990	✓	✓		✓	✓			
Willkins	1992	✓	✓		✓	✓			
Pinheiro	1993	✓		✓			✓	✓	
Rozman	1994	✓	✓		✓	✓			
Farr	1995	✓		✓			✓	✓	
Willkins	1999	✓	✓		✓	✓			
Dougados	1999								
Cohen	2001	✓	✓	✓	✓	✓	✓	✓	✓
Kalden	2001						✓		
Kremer	2002	✓	✓	✓	✓	✓	✓	✓	
Bao	2003	✓	✓	✓	✓	✓	✓	✓	
Mariette	2004						✓		
Dougados	2005	✓	✓	✓			✓	✓	
Hetland	2006						✓		✓
Karanikolas	2006	✓	✓	✓	✓	✓	✓	✓	
Capell	2007	✓	✓	✓	✓	✓	✓	✓	
Ogrendik	2007	✓	✓	✓	✓	✓	✓	✓	

Trial Flow Diagram



Publication Bias ⇦

ORB ⇦

The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews

Jamie J Kirkham,[1] Kerry M Dwan,[1] Douglas G Altman,[2] Carrol Gamble,[1] Susanna Dodd,[1] Rebecca Smyth,[3] Paula R Williamson[1]

	Clear that the outcome was measured and analysed		
A	States outcome analysed but only reported that result not significant (typically stating p-value >0.05).	Partial	High risk
B	States outcome analysed but only reported that result significant (typically stating p-value <0.05).	Partial	Low risk
C	States outcome analysed but insufficient data presented to be included in meta-analysis or to be considered to be fully tabulated.	Partial	Low risk
D	States outcome analysed but no results reported.	None	High risk
	Clear that the outcome was measured		
E	Clear that outcome was measured but not necessarily analysed. Judgment says likely to have been analysed but not reported because of non-significant results	None	High risk
F	Clear that outcome was measured but not necessarily analysed. Judgment says unlikely to have been analysed but not reported because of non-significant results	None	Low risk
	Unclear that the outcome was measured		
G	Not mentioned but clinical judgment says likely to have been measured and analysed.	None	High risk
H	Not mentioned but clinical judgment says unlikely to have been measured.	None	Low risk
	Clear that the outcome was NOT measured		
I	Clear that outcome was not measured.	N/A	No risk

The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews

Jamie J Kirkham,[1] Kerry M Dwan,[1] Douglas G Altman,[2] Carrol Gamble,[1] Susanna Dodd,[1] Rebecca Smyth,[3] Paula R Williamson[1]

ORBIT – key messages

BMJ (2010); **340**:c356

- ORB suspected in at least one trial in 34% of 283 Cochrane reviews
- Based on 88% sensitivity and 80% specificity for detecting the bias

ELSEVIER

CrossMark

Journal of Clinical Epidemiology 68 (2015) 542–550

Journal of Clinical Epidemiology

Multivariate meta-analysis helps examine the impact of outcome reporting bias in Cochrane rheumatoid arthritis reviews

Giacomo Frosi[a,*], Richard D. Riley[b], Paula R. Williamson[a], Jamie J. Kirkham[a]

[a]*Department of Biostatistics, University of Liverpool, 1st Floor Duncan Building, Daulby Street, Liverpool, L69 3GA, United Kingdom*

[b]*School of Health and Population Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom*

Accepted 24 November 2014; Published online 28 November 2014

- All 21 reviews contained missing data on at least one of the eight outcomes.

- ORB was highly suspected in 247 (22%) of the 1,118 evaluable outcomes from 155 assessable trials

Outcome Reporting Bias in RA Reviews

(Tocilizumab for rheumatoid arthritis)

Study ID (author, study start date)	OMERACT Core Outcome Set							
	Tender joints	Swollen joints	Pain	Physician global	Patient global	Functional status	Acute phase reactants	Radiological damage
Study 1	✔	✔	✔	✔	✔	✔	✔	✘
Study 2	✔	✔	✔	✔	✔	✔	✔	✘
Study 3	✔	✔	✔	✔	✔	✔	✔	✘
Study 4	✔	✔	✔	✔	✔	✔	✔	✔
Study 5	✔	✔	✔	✔	✔	✔	✔	✘
Study 6	✔	✔	✔	✔	✔	✔	✔	✘
Study 7	✔	✔	✔	✔	✔	✔	✔	✘
Study 8	✔	✔	✔	✔	✔	✔	✔	✘

Measured

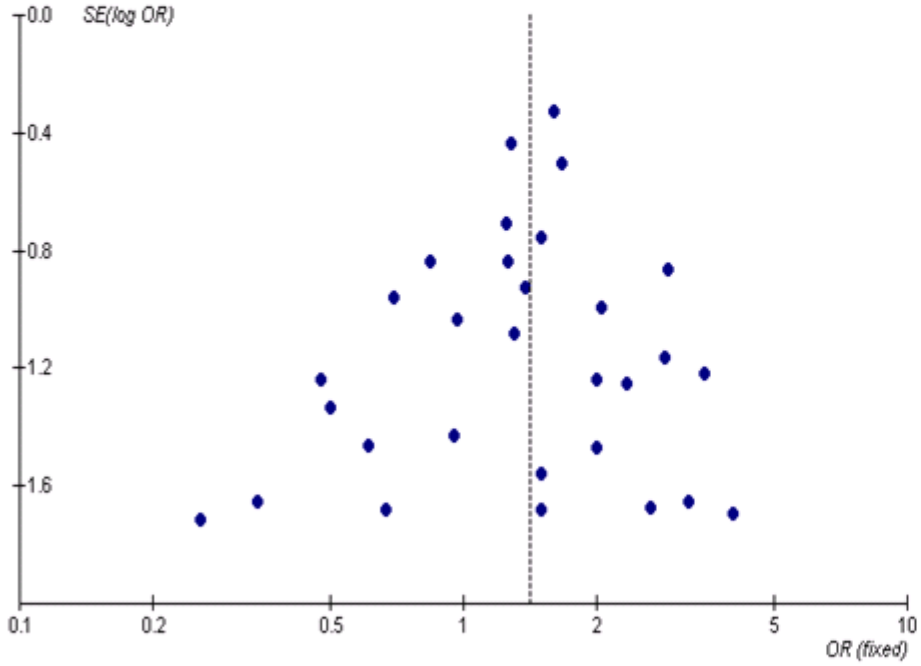
Outcome Reporting Bias in RA Reviews

(Tocilizumab for rheumatoid arthritis)

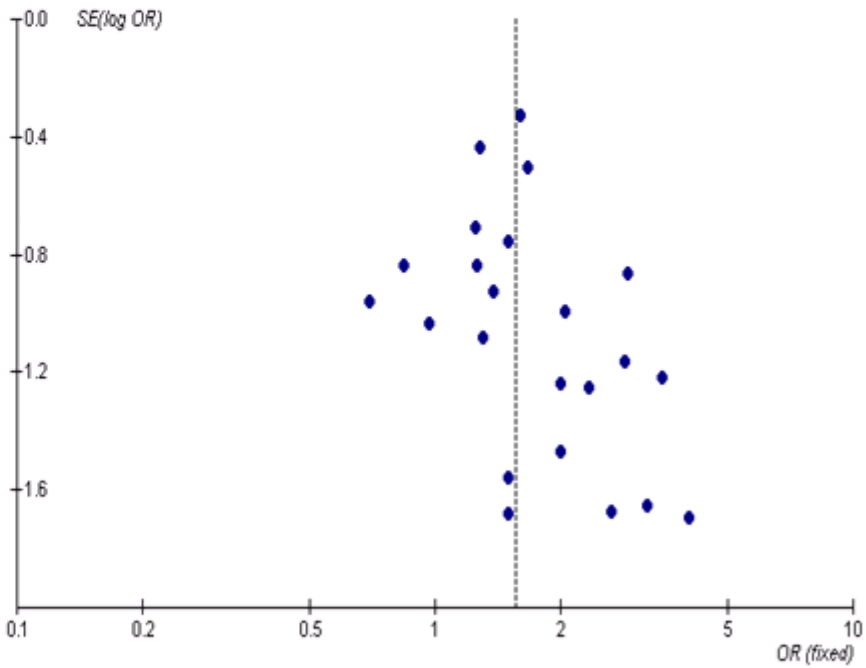
Study ID (author, study start date)	OMERACT Core Outcome Set							
	Tender joints	Swollen joints	Pain	Physician global	Patient global	Functional status	Acute phase reactants	Radiological damage
Study 1	✗	✗	✗	✗	✗	✗	✗	✗
Study 2	✗	✗	✗	✗	✗	✗	✗	✗
Study 3	✗	✗	✗	✗	✗	✗	✗	✗
Study 4	✗	✗	✗	✗	✗	✗	✗	✓
Study 5	✓	✓	✓	✓	✓	✓	✓	✗
Study 6	✓	✓	✗	✗	✗	✓	✓	✗
Study 7	✓	✓	✗	✗	✗	✓	✗	✗
Study 8	✗	✗	✗	✗	✗	✗	✗	✗

Measured and Reported

Impact of ORB / Publication Bias



OR 1.41 (1.04,1.91)



OR 1.55 (1.13,2.14)

The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews

Jamie J Kirkham,[1] Kerry M Dwan,[1] Douglas G Altman,[2] Carrol Gamble,[1] Susanna Dodd,[1] Rebecca Smyth,[3] Paula R Williamson[1]

ORBIT – key messages

BMJ (2010); **340**:c356

- ORB suspected in at least one trial in 34% of 283 Cochrane reviews

- 42 significant meta-analyses
  - 8 (19%) would not have remained significant
  - 11 (26%) would have overestimated the treatment effect by > 20%

Which is the bigger problem?

Outcome reporting bias

Publication bias

OPEN ACCESS Freely available online

PLOS ONE

Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias — An Updated Review

Kerry Dwan*, Carrol Gamble, Paula R. Williamson, Jamie J. Kirkham, for the Reporting Bias Group¶

Department of Biostatistics, University of Liverpool, Liverpool, England

- Fully reported: OR 2.2 to 4.7 if statistically significant
- Reports vs protocols: 40–62% at least one primary outcome changed, newly introduced or omitted

Whose fault is poor reporting?

- Poor reporting indicates a collective failure of authors, peer reviewers, and editors

  ….on a massive scale

- Researchers (authors) may not know what information to include in a report of research

- Editors may not know what information should be included

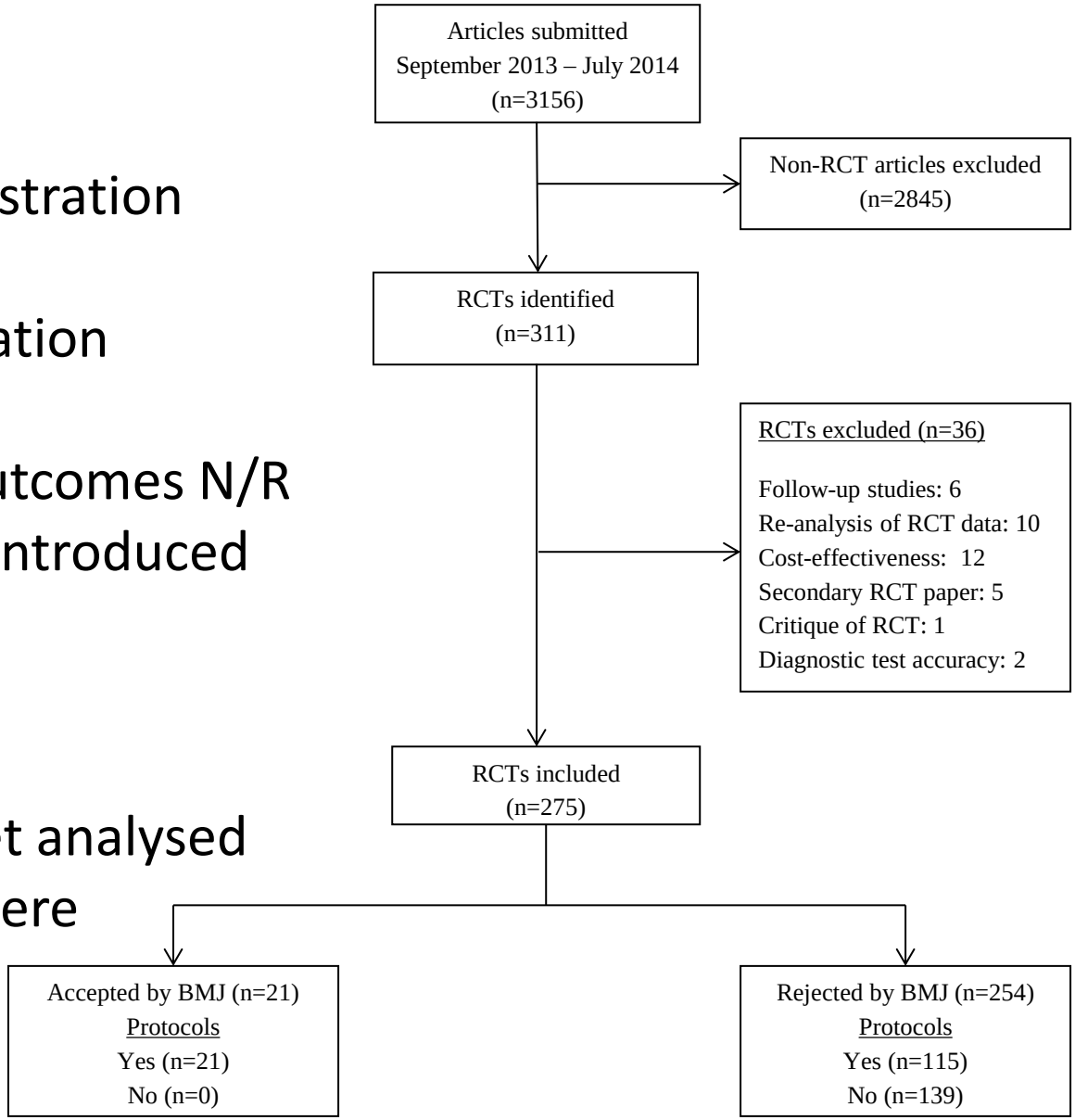
What help can be given to authors?

What help can be given to editors?

- Prospective trial registration
- Protocol required
- Transparency declaration

- 20% pre-specified outcomes N/R
- 10% new outcomes introduced

- Reasons:
  - Space limitation
  - Outcomes not yet analysed
  - Reported elsewhere
  - Errors



Selective Reporting

BMJ

RESEARCH

Frequency and reasons for outcome reporting bias in clinical trials: interviews with trialists

"When I take a look at the data I see what best advances the story, and if you include too much data the reader doesn't get the actual important message, so sometimes you get data that is either not significant or doesn't show anything, and so you, we, just didn't include that". *Smyth et al., 2011*

BMJ 2011; 342:c7153.

BMJ

RESEARCH

Frequency and reasons for outcome reporting bias in clinical trials: interviews with trialists

R M D Smyth, research associate,[1,2] J J Kirkham, research associate,[1] A Jacoby, professor of medical sociology,[2] D G Altman, professor of statistics in medicine,[3] C Gamble, senior lecturer,[1] P R Williamson, professor of medical statistics[1]

- 4/17(24%), trials in which pre-specified outcomes had been measured but not analysed (the "direction" of the main findings influenced the investigators' decision not to analyse the remaining data collected.

- In 14 (67%) of the 21 randomly selected PubMed trials, there was at least one unreported efficacy or harm outcome.

thebmj

CrossMark
click for updates

BMJ 2014;349:g6501 doi: 10.1136/bmj.g6501 (Published 21 November 2014)

RESEARCH

Selective reporting bias of harm outcomes within studies: findings from a cohort of systematic reviews

- Specific harm outcomes were reported in only 38% (92/243) of Cochrane reviews.

- Overall, 76% (705/931) of primary studies in the Cochrane reviews did not report the relevant review's primary harm outcome, or did not report it in full.

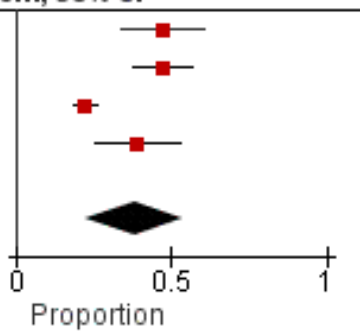
Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions (Review)

Page MJ, McKenzie JE, Kirkham J, Dwan K, Kramer S, Green S, Forbes A

Study or Subgroup	Proportion	SE	Weight	Proportion IV, Random, 95% CI
Silagy 2002 (2002)	0.47	0.07	23.3%	0.47 [0.33, 0.61]
Parmelli 2007 (2005-2006)	0.47	0.05	25.5%	0.47 [0.37, 0.57]
Kirkham 2010b (2006-2007)	0.22	0.02	27.9%	0.22 [0.18, 0.26]
Dwan 2013a (2006-2009)	0.39	0.07	23.3%	0.39 [0.25, 0.53]
Total (95% CI)			**100.0%**	**0.38 [0.23, 0.54]**

Heterogeneity: $Tau^2 = 0.02$; $Chi^2 = 33.23$, df = 3 ($P < 0.00001$); $I^2 = 91\%$

Test for overall effect: $Z = 4.78$ ($P < 0.00001$)



Prevalence of systematic reviews that added, omitted, upgraded or downgraded at least one outcome between the protocol and published systematic review was 38% (95% CI 23% to 54%

Solutions to ORB

Non-Statistical Solutions

- Obtain the missing outcome data

Statistical Solutions (sensitivity analyses)

- Bound for maximum bias (*Trials* 2007; 8:9)
- Multivariate meta-analysis (*Statistics in Medicine* (2012); 31(20): 2179-2195)
- Explicit modelling techniques (*Biostatistics* 2014; 15(2): 370-383)
- Other methods, e.g. regression approaches

Which is the bigger problem?

Outcome reporting bias

Publication bias

Aims and Objectives

Aim: To estimate the proportion of missing participant data due to lack of publication of the study and the proportion due to missing outcome data within a published study.

Objective:

- Compute the proportion of fully reported outcome data
- Compute the proportion of partially reported data
- Compute the proportion of missing data from published studies (selective reporting)
- Compute the proportion of missing data from unpublished studies (publication bias)
- Compute the proportion of missing data from all studies (published and unpublished)

Data sources: Protocols of clinical research projects submitted to the research ethics committee of the University of Freiburg (Germany) and associated full published articles

Results

- **Study cohort:** 308 studies; 167 (54%) published
  - Increased risk of commercially funded studies being published
    [Relative risk 1.20, 95% CI (0.86, 1.67)]

- **Outcomes**: 3407 (from 308 studies)
  - Commercially funded studies less likely to publish all outcomes
    [Relative risk 0.64, 95% CI (0.30, 1.38)]

- **Total participant data:** 2,618,116 (*sample size x outcomes)

*For published studies the sample size was taken from the study publication (actual sample size achieved); for unpublished studies this was taken as the planned sample size from the study protocol.

Results

Proportions of reporting/missingness	
Proportion of fully published data	47%
Proportion of partially reported data	34%
Proportion of missing data from published studies (within-study selective outcome reporting)	4%
Proportion of missing data from unpublished studies (publication bias)	15%
Proportion of missing data from all studies	19%
Sensitivity analyses	
Proportion of missing data from all studies (partially reported = unpublished)	53%

Multivariate meta-analysis

UNIVERSITY OF LIVERPOOL

Research Article

Statistics in Medicine

Received 7 July 2011, Accepted 8 February 2012 Published online 25 April 2012 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.5356

A multivariate meta-analysis approach for reducing the impact of outcome reporting bias in systematic reviews

Jamie J. Kirkham,[a*†] Richard D. Riley[b] and Paula R. Williamson[a]

Bound for maximum bias

UNIVERSITY OF LIVERPOOL

Bivariate fixed effect meta-analysis model

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \sim N\left(\begin{pmatrix} \beta_X \\ \beta_Y \end{pmatrix}, \begin{pmatrix} s^2_{X_i} & \rho_i s_{X_i} s_{Y_i} \\ \rho_i s_{X_i} s_{Y_i} & s^2_{Y_i} \end{pmatrix}\right)$$

β_x :	treatment effects outcome 1
β_Y :	treatment effects outcome 2
s^2x_i:	within study variances for outcome 1 (assumed known)
s^2y_i:	within study variances for outcome 2 (assumed known)
$\rho_{i:}$	within-study correlations, indicating the association between outcome estimates within a study (assumed known)

Multivariate meta-analysis
(Example)

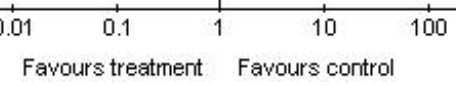
UNIVERSITY OF LIVERPOOL

- Two binary outcomes of interest: TF and mortality
- 15 eligible trials all reported TF: pooled RR (95% CI) of 1.11 (1.02, 1.21) [favours combination therapy]

Review: Example
Comparison: 01 Treatment versus control
Outcome: 01 Mortality

Study or sub-category	Treatment n/N	Control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
1	22/364	32/369		46.56	0.70 [0.41, 1.18]
2	2/77	7/83		9.87	0.31 [0.07, 1.44]
3	5/45	5/45		7.32	1.00 [0.31, 3.22]
4	1/22	2/22		2.93	0.50 [0.05, 5.12]
5	1/39	2/46		2.69	0.59 [0.06, 6.26]
6	1/35	1/42		1.33	1.20 [0.08, 18.50]
7	2/42	3/82		2.98	1.30 [0.23, 7.49]
8	7/95	5/94		7.36	1.39 [0.46, 4.21]
9	0/19	4/21		6.28	0.12 [0.01, 2.13]
10	8/50	9/54		12.68	0.96 [0.40, 2.29]
Total (95% CI)	788	858		100.00	0.74 [0.53, 1.06]

Total events: 49 (Treatment), 70 (Control)
Test for heterogeneity: Chi² = 5.29, df = 9 (P = 0.81), I² = 0%
Test for overall effect: Z = 1.66 (P = 0.10)



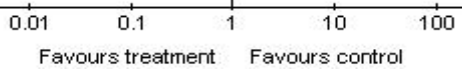
Multivariate meta-analysis

(Example)

UNIVERSITY OF LIVERPOOL

Review: Example
Comparison: 02 Treatment versus control
Outcome: 01 Treatment Failure

Study or sub-category	Treatment n/N	Control n/N	RR (fixed) 95% CI	Weight %	RR (fixed) 95% CI
01 Both outcomes					
1	188/370	188/384		33.72	1.04 [0.90, 1.20]
2	48/77	59/83		10.38	0.88 [0.70, 1.09]
3	19/36	18/33		3.43	0.97 [0.62, 1.50]
4	5/22	6/22		1.10	0.83 [0.30, 2.33]
5	23/39	26/46		4.36	1.04 [0.72, 1.50]
6	3/39	6/45		1.02	0.58 [0.15, 2.16]
7	13/38	29/76		3.53	0.90 [0.53, 1.52]
8	47/95	39/94		7.16	1.19 [0.87, 1.63]
9	9/19	10/21		1.74	0.99 [0.52, 1.91]
10	13/49	11/48		2.03	1.16 [0.58, 2.33]
Subtotal (95% CI)	784	852		68.47	1.01 [0.91, 1.12]
Total events: 368 (Treatment), 392 (Control)					
Test for heterogeneity: Chi² = 4.02, df = 9 (P = 0.91), I² = 0%					
Test for overall effect: Z = 0.22 (P = 0.82)					
02 Single outcome					
11	14/46	5/45		0.92	2.74 [1.08, 6.98]
12	10/35	3/32		0.57	3.05 [0.92, 10.10]
13	67/77	50/69		9.64	1.20 [1.01, 1.42]
14	130/378	99/372		18.24	1.29 [1.04, 1.61]
15	14/35	12/36		2.16	1.20 [0.65, 2.22]
Subtotal (95% CI)	571	554		31.53	1.33 [1.15, 1.55]
Total events: 235 (Treatment), 169 (Control)					
Test for heterogeneity: Chi² = 5.75, df = 4 (P = 0.22), I² = 30.5%					
Test for overall effect: Z = 3.78 (P = 0.0002)					
Total (95% CI)	1355	1406		100.00	1.11 [1.02, 1.21]
Total events: 603 (Treatment), 561 (Control)					
Test for heterogeneity: Chi² = 17.00, df = 14 (P = 0.26), I² = 17.7%					
Test for overall effect: Z = 2.50 (P = 0.01)					



Practical Example: Head and Neck Cancer

*Tandon et al. (2010) assess the prognostic association of mutant p53 for OS & DFS in patients with head & neck cancer

For cancers at the Laraynx, Oral cavity, Oropharynx and Hypopharynx site, they extracted hazard ratios for:

- Overall survival
- Disease free survival

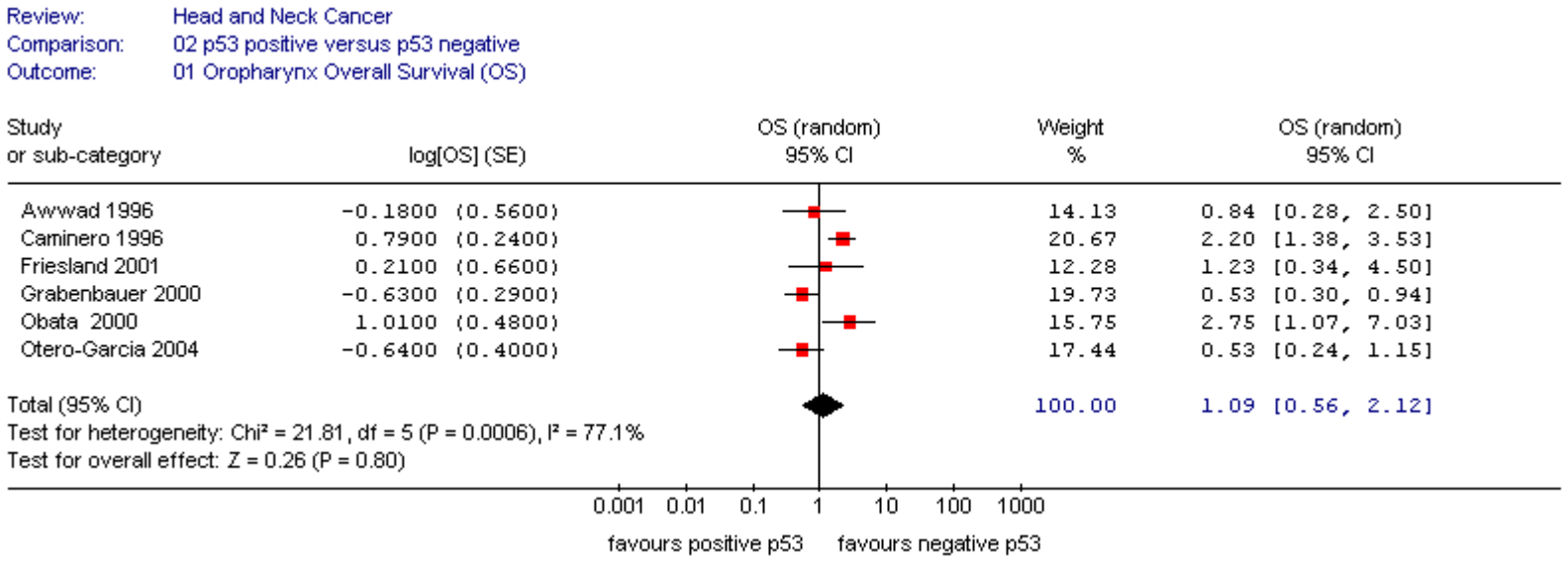
*_Cancer Epidemiology, Biomarkers & Prevention_ (2010) 19; 574

Head and Neck Cancer – outcome matrix (Hazard Ratio)

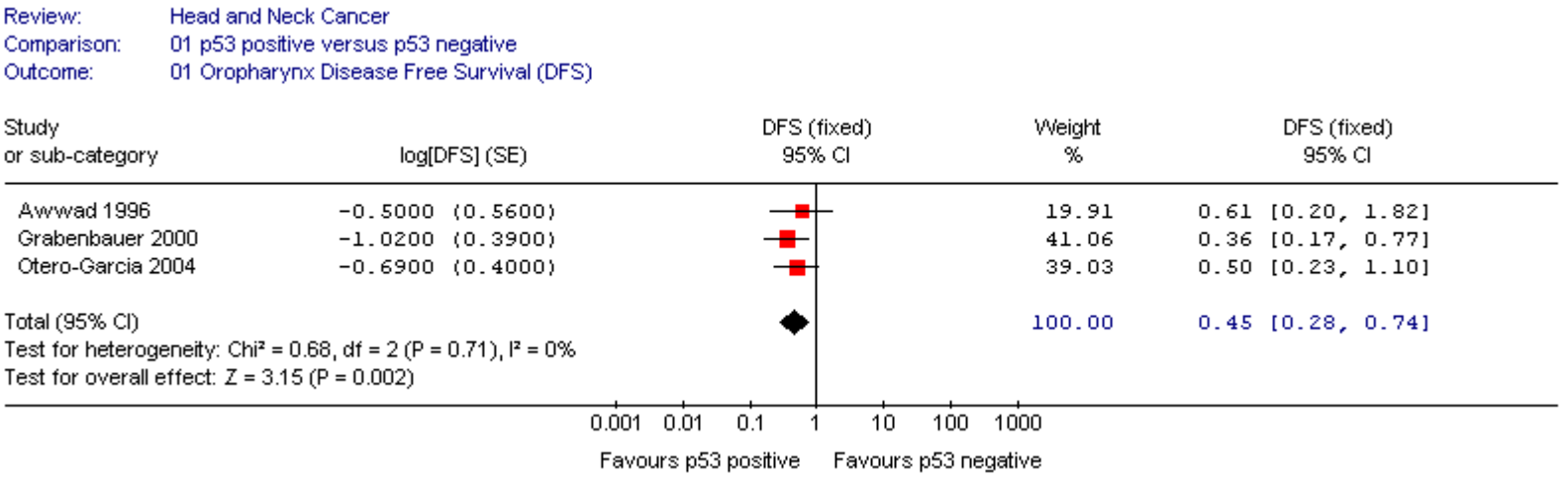
Study	Site	OS	DFS
Awwad	Oropharynx	✓	✓
Caminero	Oropharynx	✓	✗
Friesland	Oropharynx	✓	✗
Grabenbauer	Oropharynx	✓	✓
Obata	Oropharynx	✓	✓
Otero	Oropharynx	✓	✗

Practical Example: Univariate analysis (Oropharynx site)

OS



DFS



Practical Example: Multivariate analysis (Oropharynx site)

For OS (no missing data):

A univariate meta-analysis gives:

- a pooled hazard ratio estimate of 1.09 (p = 0.776)

A bivariate meta-analysis (assume w/s = 0.8) gives:

- a pooled hazard ratio estimate of 1.12 (p = 0.662)

... no difference in conclusions

For DFS (missing data in 3 studies):

A univariate meta-analysis gives:

- a pooled hazard ratio estimate of 0.45 (p = 0.046)

A bivariate meta-analysis (assume w/s = 0.8) gives:

- a pooled hazard ratio estimate of 0.80 (p = 0.47)

... very different conclusions!!!

Van Houwelingen et al., 2002

STATISTICS IN MEDICINE
Statist. Med. 2002; **21**:589–624 (DOI: 10.1002/sim.1040)

TUTORIAL IN BIOSTATISTICS

Advanced methods in meta-analysis: multivariate approach and meta-regression

Hans C. van Houwelingen[1,*,†], Lidia R. Arends[2] and Theo Stijnen[2]

Good overview of multivariate methods for meta-analysis
Provides code on how to implement in SAS

Multivariate meta-analysis
(Example)

Model		All-cause Mortality (b_1)				Treatment Failure (b_2)			
		Pooled Risk Ratio	Standard Error	P-value	95% Confidence Interval	Pooled Risk Ratio	Standard Error	P-value	95% Confidence Interval
UFMA	0	0.778	0.182	0.168	(0.545, 1.111)	1.095	0.040	0.024	(1.012, 1.184)
BFMA	0.021 (impute)	0.782	0.182	0.177	(0.548, 1.117)	1.095	0.040	0.024	(1.012, 1.184)
BFMA	0.2	0.818	0.180	0.264	(0.575, 1.164)	1.097	0.040	0.020	(1.015, 1.187)
BFMA	0.309 (impute)	0.841	0.178	0.330	(0.594, 1.191)	1.099	0.040	0.018	(1.016, 1.188)
BFMA	0.4	0.861	0.175	0.393	(0.611, 1.213)	1.099	0.040	0.017	(1.017, 1.189)
BFMA	0.6	0.909	0.165	0.561	(0.657, 1.256)	1.101	0.039	0.014	(1.020, 1.189)
BFMA	0.7	0.933	0.158	0.662	(0.685, 1.271)	1.101	0.038	0.012	(1.021, 1.188)
BFMA	0.8	0.959	0.147	0.775	(0.718, 1.280)	1.102	0.037	0.010	(1.024, 1.185)

Core outcome set

- Agreed standardised set of most important outcomes

- Disease/condition specific
- All treatment types or a particular intervention

- Both benefits and harms
- The minimum – expect others to be collected

- Focus of effectiveness trials
- Relevant within routine clinical practice

Kirkham *et al. Trials* 2013, **14**:21
http://www.trialsjournal.com/content/14/1/21

TRIALS

RESEARCH Open Access

Can a core outcome set improve the quality of systematic reviews? – a survey of the Co-ordinating Editors of Cochrane review groups

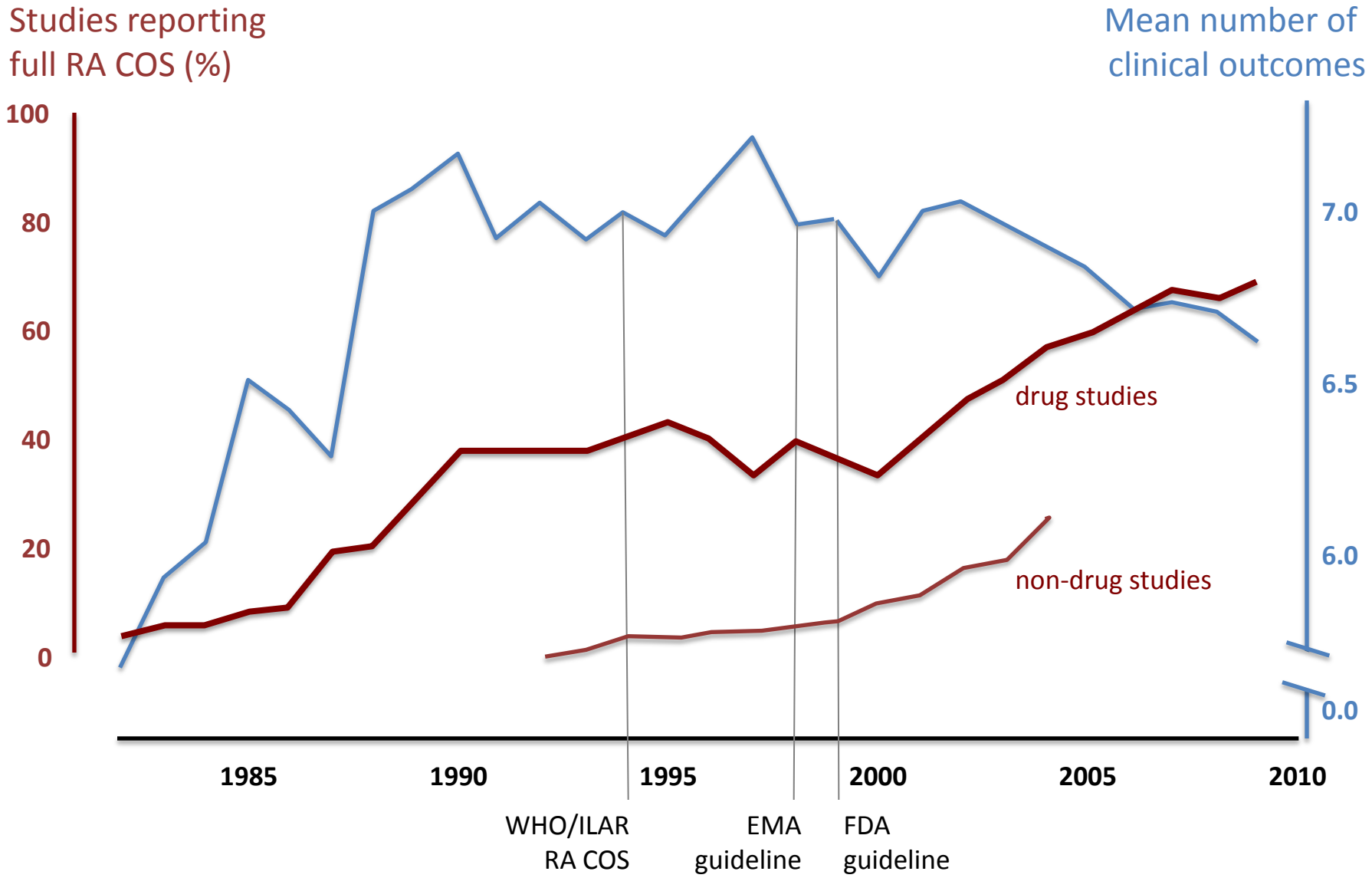
Jamie J Kirkham[1*], Elizabeth Gargon[1], Mike Clarke[2] and Paula R Williamson[1]

- 73% of Co-ordinating Editors thought that a COS for effectiveness trials should be used routinely for a SoF table

- Different CRG's have different policies on outcome selection

- Three main COS challenges were:
  - development of COS,
  - scope of COS and
  - persuading authors/trialists to implement


TO DO


Studies reporting full RA COS (%)
Mean number of clinical outcomes
100
80
60
40
20
0
7.0
6.5
6.0
0.0
drug studies
non-drug studies
1985
1990
1995
2000
2005
2010
WHO/ILAR RA COS
EMA guideline
FDA guideline

OPEN ACCESS Freely available online

PLOS | MEDICINE

Evidence for the Selective Reporting of Analyses and Discrepancies in Clinical Trials: A Systematic Review of Cohort Studies of Clinical Trials

Kerry Dwan[1]*, Douglas G. Altman[2], Mike Clarke[3], Carrol Gamble[1], Julian P. T. Higgins[4,5], Jonathan A. C. Sterne[4], Paula R. Williamson[1], Jamie J. Kirkham[1]

- Per protocol analysis vs. ITT
- Subgroup analyses
- Dichotomising continuous outcomes
- Dealing with missing data