Physiotherapy Papers and Journal Articles

School of Physiotherapy

2019

Reviews may overestimate the effectiveness of medicines for back pain: Systematic review and meta-analysis

Matthew K. Bagg

Edel O'Hagan

Pauline Zahara

Benedict Wand
The University of Notre Dame Australia, benedict.wand@nd.edu.au

Markus Hubscher

See next page for additional authors

Authors

Matthew K. Bagg, Edel O'Hagan, Pauline Zahara, Benedict Wand, Markus Hubscher, G. Lorimer Moseley, and James H. McAuley

This is the accepted manuscript version of an article published as:

Journal Pre-proof

Reviews may overestimate the effectiveness of medicines for back pain: systematic review and meta-analysis

Matthew K. Bagg, Edel O'Hagan, Pauline Zahara, Benedict M. Wand, Markus Hübscher, G Lorimer Moseley, James H. McAuley
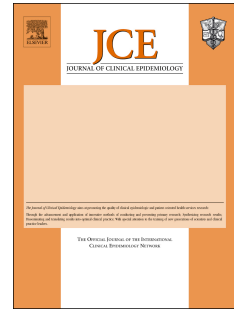
Please cite this article as: Bagg MK, O'Hagan E, Zahara P, Wand BM, Hübscher M, Moseley GL, McAuley JH, Reviews may overestimate the effectiveness of medicines for back pain: systematic review and meta-analysis, *Journal of Clinical Epidemiology* (2020), doi: https://doi.org/10.1016/j.jclinepi.2019.12.006.

Reviews may overestimate the effectiveness of medicines for back pain: systematic review and meta-analysis

Matthew K Bagg [1,2,3], Edel O'Hagan [1,2], Pauline Zahara [1], Benedict M Wand [4], Markus Hübscher [1], G Lorimer Moseley [1,5], James H McAuley [1,6]

1. Neuroscience Research Australia, Randwick, NSW 2031, Australia m.bagg@neura.edu.au; e.ohagan@neura.edu.au; p.zahara@neura.edu.au; m.huebscher@neura.edu.au; J.McAuley@neura.edu.au

2. Prince of Wales Clinical School, University of New South Wales, Kensington, NSW 2052, Australia

3. New College Village, University of New South Wales, Kensington, NSW 2052, Australia

4. School of Physiotherapy, The University of Notre Dame Australia, Fremantle, WA 6959, Australia, benedict.wand@nd.edu.au

5. School of Health Sciences, University of South Australia, SA 5001, Australia, Lorimer.Moseley@unisa.edu.au

6. School of Medical Sciences, University of New South Wales, Kensington, NSW 2052, Australia

Correspondence to:

Matthew K Bagg

Neuroscience Research Australia

PO Box 1165, Randwick, NSW 2031, Australia

Tel: +61 2 9399 1870

Email: m.bagg@neura.edu.au

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

2

Abstract

Objective: Systematic-reviews of analgesics for low back pain generally include published data only. Obtaining data from unpublished trials is potentially important because they may impact effect sizes in meta-analyses. We determined whether including unpublished data from trial registries changes the effect sizes in meta-analyses of analgesics for low back pain.

Study Design and Setting: Trial registries were searched for unpublished data that conformed to the inclusion criteria of n=5 individual source systematic-reviews. We reproduced the meta-analyses using data available from the original reviews then re-ran the same analyses with the addition of new unpublished data.

Results: Sixteen completed, unpublished, trials were eligible for inclusion in four of the source reviews. Data were available for five trials. We updated the analyses for two of the source reviews. The addition of data from two trials reduced the effect size of muscle relaxants, compared to sham, for recent-onset low back pain from -21.71 (95%CI -28.23 to -15.19) to -2.34 (95%CI -3.34 to -1.34) on a 0-100 scale for pain intensity. The addition of data from three trials (one enriched design) reduced the effect size of opioid analgesics, compared to sham, for chronic low back pain from -10.10 (95%CI -12.81 to -7.39) to -9.31 (95%CI -11.51 to -7.11). The effect reduced in the subgroup of enriched design studies, from -12.40 (95%CI -16.90 to -7.91) to -11.34 (95%CI -15.36 to -7.32), and in the subgroup of non-enriched design studies; from -7.27 (95%CI -9.97 to -4.57) to -7.19 (95%CI -9.24 to -5.14).

Conclusion: Systematic-reviews should include reports of unpublished trials. The result for muscle relaxants conflicts with the conclusion of the published review and recent international guidelines. Adding unpublished data strengthens the evidence that opioid analgesics have small effects on persistent low back pain and more clearly suggests these effects may not be clinically meaningful.

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

3

Key words: Back pain, analgesics, publication bias, clinical trials (as topic), meta-analysis (as topic)

Word count: 3624

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

4

What is New?

What is already known:

- Systematic-reviews and meta-analyses that include unpublished data from trial registries may reach different conclusions to those that exclude unpublished data.

What are the new findings:

- Including unpublished data from trial registries changes the effect sizes reported in recently published meta-analyses of common analgesic medicines for low back pain

- Including unpublished data reduced the estimate of treatment effect in all comparisons for which new data were available.

- Current systematic-reviews of medicines for low back pain likely overestimate the benefits of these interventions

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

5

1. Introduction

Evidence from systematic-review and meta-analysis of data from controlled trials has an important role in guideline development and clinical decision-making. It is well-known that not all trials are published promptly, or indeed at all [1–3]. Unpublished data may substantially impact effect sizes when included in published meta-analyses [4–9]. For example, including unpublished data changed the effect size by up to 29% in 14 meta-analyses across cardiovascular medicine, oncology, neurology and rheumatology [5]. Relatedly, published and unpublished data may provide different estimates for the same comparison [10,11]. For example, the effect sizes for comparisons amongst anti-depressant medicines may differ depending on whether the meta-analyses use data that are published, or data from regulatory submissions [10,11].

Analgesic medicines are a common treatment for the management of low back pain (LBP) [12–16] and recent systematic-reviews of analgesic medicines [17–25], with a single exception [23], only include published data. Clinical trial registries provide publicly accessible records of unpublished trials and are recommended information sources for systematic-review teams [26,27]. Despite this, a systematic literature search indicates the effect of including unpublished data on the effect size of published meta-analyses of analgesic medicines for LBP remains unknown.

The aim of this study was to evaluate the effect of including publicly available unpublished data, identified through ClinicalTrials.gov and the World Health Organisation International Clinical Trial Registry Platform (ICTRP), on the effect sizes of published meta-analyses of the most commonly prescribed analgesic medicines for LBP.

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

6

2. Methods

This review represents a supplementary analysis of five previously published systematic-reviews.

2.1 Published meta-analyses used in this study

We investigated systematic-reviews of paracetamol [19], non-steroidal anti-inflammatory drugs (NSAIDs) [20,21], muscle relaxant medicines [18], and opioid analgesics [17] for people with LBP. These were the five most recently published systematic-reviews at the time of writing the protocol. Furthermore, these medicines are the most commonly prescribed for LBP [12,13,15,16,28–30] and, with the exception of paracetamol, are recommended in recent international guidelines for the management of LBP [31].

2.2 Search and selection process for unpublished data

Two authors independently extracted details of the inclusion criteria used in each of the previous systematic-reviews (see eTable 1). Disagreements were resolved by discussion. Arbitration by a third author was available if required. We developed a catalogue of analgesic medicines listed on the World Health Organisation Anatomical Therapeutic Chemical classification system and currently licensed in either Australia, Europe or the United States [32–34] (eTable 2). We adapted the Saragiotto [19] and Enthoven [21] reviews' search strategies for the ICTRP and ClinicalTrials.gov, adding the medicines from the catalogue. For three reviews that did not search trial registries [17,18,20], we developed separate search strategies *de novo,* using search terms for LBP and the catalogued medicines. We used each strategy to search the ICTRP and ClinicalTrials.gov for all trial records registered to the 6th of June 2017 – the most recent search date from the included source systematic-reviews. A summary of the search strategies can be found in eTable 3.

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

7

We screened records against the previously extracted inclusion criteria for each review. Two authors independently screened the title, keywords, and record description and subsequently independently screened the full-record; defined as the full web-page, and any downloadable documents. Disagreements were resolved by discussion. Arbitration by a third author was available if required.

MKB verified the additional criteria for inclusion in this study that trials must have been i) registered and ii) unpublished prior to the date of the last search by the respective source systematic-review. Trials were judged unpublished when: i) the full record contained no listed publications, ii) the downloaded dataset (where available) contained no listed publications and iii) independent searches of PubMed [www.ncbi.nlm.nih.gov/pubmed], Google Scholar [scholar.google.com.au/] and Google Search [www.google.com.au/webhp] using the record registration number returned no results. The search and selection process is displayed in adapted Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagrams [35] in eFigures 1-5. We also noted records of trials that were terminated or on-going, at the time of the relevant review, in these diagrams.

2.3 Extraction and management of unpublished data

Two authors independently extracted all available data from full-records into a customised spreadsheet using the same methods and data items as the previous reviews. This included imputing missing standard deviations. The data availability status of eligible records is described in the adapted PRISMA diagrams (eFigures 1-5). We did not contact the sponsors listed on trial registry records to request additional data.

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

8

2.4 Outcomes

The primary outcome in this study is the percentage change in effect size for the effect of an analgesic medicine on a pain outcome in the previous review when available unpublished data are included. We also reported the change in heterogeneity variance with the addition of new unpublished data.

2.5 Data analysis

Two authors independently extracted data from each source systematic-review. We used these data to reproduce the published meta-analyses for the effect of an analgesic medicine on a pain outcome. We checked the forest plots against those in the original articles to verify accuracy. We added all available new data from the trial registry records and calculated a pooled effect for the unpublished data subgroup and a total pooled effect for the published and unpublished data combined. We calculated the percentage change in effect size between the original meta-analysis and the meta-analysis that combined published and unpublished data, following previously described methods [5]. We tested the difference between the original and the combined meta-analyses using the Z-test implemented in Revman 5.3 [36,37]. We described heterogeneity in each meta-analysis with the heterogeneity variance, Tau^2.

2.6 Changes to the protocol and imputed data.

We used the full catalogue of analgesic medicines because the previous reviews [17–21] did not report their inclusion criteria for interventions in greater detail than medicine class, e.g. muscle relaxants. We also included hydrocodone, which was not on the World Health Organisation Anatomical Therapeutic Chemical classification system, because it was included in the review of opioid analgesics [17]. Searches for records of hydrocodone trials were conducted on the 14th of June 2018. We searched 'back pain' in *Condition* instead of in *Title* [19,38] for the paracetamol search strategy.

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

9

We imputed standard deviations for the following studies: EUCTR2012-001920-36, NCT01358526 and NCT01455519. NCT00671502 reported between group differences and 95% confidence intervals for the outcome score. We attempted to impute standard deviations by calculating the standard error and then the standard deviation. However, the confidence intervals were not symmetrical, suggesting the original data were skewed. We used the change score and the standard deviation from NCT00671879 and adjusted the difference in change scores to match the between group differences in outcome in NCT00671502. We tested the influence of excluding NCT00671502 in sensitivity analyses. We determined *post hoc* to report the change in the heterogeneity parameter and the change in sample size, in addition to the change in effect size, for each of the meta-analyses.

The original meta-analysis of muscle relaxants [18] includes a single trial titled *Berry 1988*. Whereas, two trials, Berry 1988a [39] and Berry 1988b [40], are reported in the 'Characteristics of included studies' table. We obtained the paper for Berry 1988a [39] and followed the original reviews' [18] methods to add Berry 1988a [39] to the meta-analysis. We also evaluated the impact of unpublished data on this meta-analysis.

2.7 Ethics and transparency

This study did not require ethical approval. It is a re-analysis of published systematic-reviews, using data publicly available online. Our protocol is publicly accessible on the Open Science Framework [38] and any discrepancies from the study as originally planned have been explained in the preceding section. We have omitted the secondary outcome reported in the protocol for clarity of presentation.

3. Results

3.1 Paracetamol – Saragiotto et al.

We retrieved 66 records via the ICTRP and 76 via ClinicalTrials.gov. We screened 54 unique trial records, excluded 52 and screened the full-record of the remaining two trials. We excluded both records as one trial was published and the other was registered after the date of the last search (eFigure 1).

3.2 NSAIDs – Enthoven et al.

We retrieved 103 records via the ICTRP and 94 via ClinicalTrials.gov. We screened 117 unique trial records, excluded 82 and screened the full-record of the remaining 35 trials. Five records were eligible for inclusion in the Enthoven review [21]. We did not include any of these trials as, at the time of the Enthoven review [21], two trials had been terminated, a single trial was on-going, a single trial had unknown status and the single completed trial did not have any data available (eFigure 2).

3.3 NSAIDs – Machado et al.

The same 117 records were screened as for the Enthoven review. We excluded 64 records and screened the full-record of the remaining 53 trials. Ten records were eligible for inclusion in the Machado review [20]. We did not include any of these trials as, at the time of the Machado review [20], four trials were on-going, a single trial had been terminated, a single trial had unknown status and numeric outcome data were not available for the four completed trials (eFigure 3).

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

11

3.4 Muscle Relaxants – Abdel Shaheed et al.

3.4.1 Results of the search

We retrieved 45 records via the ICTRP and 43 via ClinicalTrials.gov. We screened 57 unique trial records, excluded 31 and screened the full-record of the remaining 26 trials. Eight records were eligible for inclusion in the Abdel Shaheed review [18]. We excluded 6 records as, at the time of the Abdel Shaheed review [18], two trials were on-going, two trials had been terminated, a single trial had unknown status and one of the three completed trials did not have data available. Data were available for two of the three completed trials (NCT00671879 and NCT00671502) (eFigure 4). We included these two records in the updated meta-analysis.

3.4.2 Planned analysis

NCT00671879 (n=805) and NCT00671502 (n=840) contributed an additional 1645 participants, an increase of 238.59%. We reproduced the meta-analysis of the effect (mean difference) of muscle relaxant medicines on a 0-100 mm pain intensity scale, compared to a placebo medicine. The inclusion of unpublished data from NCT00671879 and NCT00671502 reduced the pooled effect size by 89.22%, from -21.71 (95%CI -28.23 to -15.19) to -2.34 (95%CI -3.34 to -1.34). Tau^2 reduced from 32.79 to 1.11 (Figure 1). The Chi^2 test for a difference in effect size between the published and unpublished subgroups was statistically significant (Chi^2=39.36, p<.00001). These data are summarised in Table 1.

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

12

Study or Subgroup	Muscle relaxant			Placebo			Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
	Mean	SD	Total	Mean	SD	Total			
1.1.1 Published data included in SR									
Berry 1988b	29	43.3	51	33	32.9	54	0.5%	-4.00 [-18.77, 10.77]	
Chandanwale 2011	20.7	15.1	112	44.9	23	113	3.3%	-24.20 [-29.28, -19.12]	
Hindle 1972	-14	20.4	14	0	20.8	14	0.4%	-14.00 [-29.26, 1.26]	
Ketenci 2005_a	18.6	16.6	32	43.7	27.9	13	0.4%	-25.10 [-41.32, -8.88]	
Ketenci 2005_b	6.3	11.7	38	43.7	27.9	14	0.4%	-37.40 [-52.48, -22.32]	
Tuzun 2003	25.1	20.9	73	47.4	19.8	68	2.0%	-22.30 [-29.02, -15.58]	
Subtotal (95% CI)			320			276	7.1%	-21.71 [-28.23, -15.19]	
Heterogeneity: Tau² = 32.79; Chi² = 11.49, df = 5 (P = 0.04); I² = 56%									
Test for overall effect: Z = 6.53 (P < 0.00001)									
1.1.2 Unpublished data from trial registries									
NCT00671502_a	-16.3	1.34	280	-15.2	1.32	140	23.2%	-1.10 [-1.37, -0.83]	
NCT00671502_b	-15.8	1.3	281	-15.2	1.32	139	23.2%	-0.60 [-0.87, -0.33]	
NCT00671879_a	-15.5	1.34	271	-15.2	1.32	132	23.2%	-0.30 [-0.58, -0.02]	
NCT00671879_b	-16.4	1.3	270	-15.2	1.32	132	23.2%	-1.20 [-1.47, -0.93]	
Subtotal (95% CI)			1102			543	92.9%	-0.80 [-1.21, -0.39]	
Heterogeneity: Tau² = 0.16; Chi² = 27.76, df = 3 (P < 0.00001); I² = 89%									
Test for overall effect: Z = 3.80 (P = 0.0001)									
Total (95% CI)			1422			819	100.0%	-2.34 [-3.34, -1.34]	
Heterogeneity: Tau² = 1.11; Chi² = 182.69, df = 9 (P < 0.00001); I² = 95%									
Test for overall effect: Z = 4.58 (P < 0.00001)									
Test for subgroup differences: Chi² = 39.36, df = 1 (P < 0.00001), I² = 97.5%									

Favours muscle relaxant Favours placebo

Figure 1. Effect of muscle relaxants on pain; short term; acute low back pain. Effect size is mean difference, calculated using random effects inverse variance model in RevMan 5.3. Outcome is 0-100 pain-intensity scale. 95% CIs for the unpublished data subgroup are not visible at this resolution. Placebo group Total was split to incorporate multi-arm trials as per [18]. Medicines in published data subgroup are as per [18]. NCT00671502_a and NCT00671879_a are carisoprodol 500mg SR, _b are carisoprodol 700mg SR.

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

13

Table 1. Number of additional, unpublished trials retrieved from trial registries and change in effect size when available data are added to the published meta-analysis. Methods for searching, selection and analysis of unpublished trials from trial registries followed exactly those used in the original reviews.

Medicine	Number of trials (Number of subjects randomised)				Outcome (scale)	Effect Size			Weight of new trials (%)	Change in effect size (%)	Direction of change in effect size	Change in statistical significance	Change in clinical meaningful-ness	Test for difference between subgroups
	Included in original SR (all)	Unpublished trials eligible for inclusion in original SR	Unpublished, completed	Unpublished, completed, data available		Original SR (95%CI)	Original SR [reproduced] (95%CI)	Unpublished data added						
Paracetamol [19]	3 (n=1825)	0	NA	NA	NA	NA	NA	NA	0	0	nil	nil	nil	NA
NSAIDs [21]	13 (n=4807)	5	1 (n=127)	0	NA	NA	NA	NA	0	0	nil	nil	nil	NA
NSAIDs [20]	35 (n=6065)	10	4 (n=1699)	0	NA	NA	NA	NA	0	0	nil	nil	nil	NA
Muscle Relaxants [18]	15 (n=3362)	8	3 (n=1841)	2 (n=1680)	Pain intensity, short term (0-100)	-21.3 (-29 to -13.5)	-21.71 (-28.23 to -15.19)	-2.34 (-3.34 to -1.34)	92.9	89.22	Decrease efficacy	nil	Yes	Chi²=39.36 p<.00001
Opioids [17] Combination	20 (n=6869)	11	8 (n=3251)	3 (n=1273)	Pain intensity, short term (0-100)	-10.10 (-12.8 to -7.4)	-10.10 (-12.81 to -7.39)	-9.31 (-11.51 to -7.11)	21.1	7.82	Decrease efficacy	nil	nil	Chi²=4.98 p=.03
Opioids [17] Enriched	9 (n=2996)	6	6 (n=2579)	1 (n=601)	Pain intensity, short term (0-100)	-12.40 (-16.9 to -7.9)	-12.40 (-16.90 to -7.91)	-11.34 (-15.36 to -7.32)	13	8.55	Decrease efficacy	nil	nil	Chi²=7.67 p=.006
Opioids [17] Nonenriched	11 (n=3873)	5	2 (n=672)	2 (n=672)	Pain intensity, short	-7.3 (-10 to -4.6)	-7.27 (-9.97 to -4.57)	-7.19 (-9.24 to -5.14)	18.8	1.1	Decrease efficacy	nil	nil	Chi²=0.15 p=.70

					term (0-100)									

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

15

3.4.3 Sensitivity analysis – inclusion of NCT00671879 only

The inclusion of unpublished data from NCT00671879 only, reduced the pooled effect size by 69.09% to -6.71 (95%CI -8.86 to -4.56). Tau^2 reduced from 32.79 to 3.29.

3.4.4 Additional analyses

3.4.4.1 Inclusion of Berry 1988a in original meta-analysis

The inclusion of Berry 1988a reduced the pooled effect size of the original meta-analysis from -21.71 (95%CI -28.23 to -15.19) to -17.98 (95%CI -26.72 to -9.25). Tau^2 increased from 32.79 to 102.53.

3.4.4.2 Inclusion of unpublished data to meta-analysis containing Berry 1988a

The inclusion of unpublished data from NCT00671879 and NCT00671502 reduced the pooled effect size by 87.15% to -2.31 (95%CI -3.30 to -1.31). Tau^2 reduced from 102.53 to 1.10.

3.4.4.3 Sensitivity analysis – inclusion of NCT00671879 only to meta-analysis containing Berry 1988a

The addition of unpublished data from NCT00671879 only, reduced the effect size by 64.46% to -6.39 (95%CI -8.48 to -4.29). Tau^2 reduced from 102.53 to 3.27.

3.5 Opioid analgesics – Abdel Shaheed et al.

3.5.1 Results of the search

We retrieved 114 records via the ICTRP and 153 via ClinicalTrials.gov. We screened 138 unique trial records, excluded 43 records and screened the full-record of the remaining 95 trials. Eleven trials were eligible for inclusion in the Abdel Shaheed review [17]. We excluded eight records as, at the time of the Abdel Shaheed review [17], a single trial was on-going, a single trial had been terminated, a single trial had unknown status and data were not available for five of the eight completed trials. Data were available for three of the eight completed trials (EUCTR2012-

001920-36, NCT01358526, NCT01455519) (eFigure 5). We included these three records in the updated meta-analyses.

3.5.2 Planned analysis

The original review [17] conducted meta-analyses of the short- and intermediate-term effects of single-ingredient opioid analgesics and of the intermediate-term effect of opioid/simple analgesic combinations on a 0-100 mm pain intensity scale, compared to a placebo medicine. These analyses considered enriched-enrolment (participants are selected based on response to open-label intervention, following which they are randomised to continue or withdraw from the intervention [41]) and non-enriched (conventional) designs as separate subgroups as well as a combined analysis of both enriched and non-enriched designs. We did not retrieve any new data for effects in the intermediate term, nor for combination formulations. Thus, we only reproduced the meta-analyses of the short-term effect of single-ingredient opioid analgesic medicines.

NCT01358526 contributed an additional 600 participants to the enriched designs analysis, an increase of 135.52%. The inclusion of unpublished data from NCT01358526 reduced the pooled effect size of the enriched designs meta-analysis by 8.55% from -12.40 (95%CI -16.90 to -7.91) to -11.34 (95%CI -15.36 to -7.32). Tau^2 reduced from 33.51 to 29.71 (Figure 2). The difference in effect size between the published and unpublished trials was statistically significant (Chi^2=7.67, p=.006) (Table 1).

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

17

Study or Subgroup	Opioid Mean	Opioid SD	Opioid Total	Placebo Mean	Placebo SD	Placebo Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
2.1.1 Published data included in SR									
Hale 2005_a	6.5	23.8	75	27.5	23.8	33	7.7%	−21.00 [−30.74, −11.26]	
Hale 2005_b	8.5	23.9	71	27.5	23.8	34	7.7%	−19.00 [−28.74, −9.26]	
Hale 2007	31	16.8	49	55	15.4	18	8.6%	−24.00 [−32.53, −15.47]	
Hale 2010	31	19.4	133	41	18.8	133	12.0%	−10.00 [−14.59, −5.41]	
Katz 2007	33	24.6	105	45	27.9	100	9.7%	−12.00 [−19.21, −4.79]	
Rauck 2014	4.8	15.6	151	9.6	15.5	151	12.8%	−4.80 [−8.31, −1.29]	
Schnitzer 2000	35	27.9	127	51	29.8	127	9.8%	−16.00 [−23.10, −8.90]	
Vorsanger 2008_a	35	26	129	40	25.9	63	9.2%	−5.00 [−12.81, 2.81]	
Vorsanger 2008_b	34	23.7	127	40	25.9	63	9.4%	−6.00 [−13.61, 1.61]	
Subtotal (95% CI)			967			722	87.0%	−12.40 [−16.90, −7.91]	
Heterogeneity: Tau² = 33.51; Chi² = 32.64, df = 8 (P < 0.0001); I² = 75%									
Test for overall effect: Z = 5.40 (P < 0.00001)									
2.1.2 Unpublished data from trial registries									
NCT01358526	38.6	20.03	298	43.2	19.99	302	13.0%	−4.60 [−7.80, −1.40]	
Subtotal (95% CI)			298			302	13.0%	−4.60 [−7.80, −1.40]	
Heterogeneity: Not applicable									
Test for overall effect: Z = 2.82 (P = 0.005)									
Total (95% CI)			1265			1024	100.0%	−11.34 [−15.36, −7.32]	
Heterogeneity: Tau² = 29.71; Chi² = 40.21, df = 9 (P < 0.00001); I² = 78%									
Test for overall effect: Z = 5.53 (P < 0.00001)									
Test for subgroup differences: Chi² = 7.67, df = 1 (P = 0.006), I² = 87.0%									

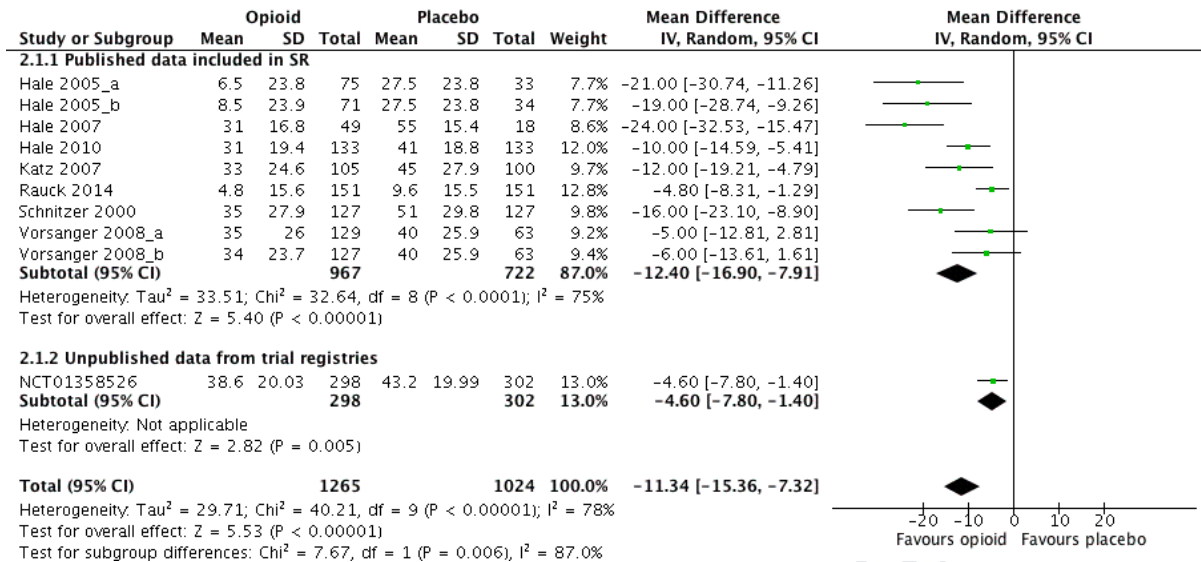Favours opioid   Favours placebo (axis: −20, −10, 0, 10, 20)

Figure 2. Single ingredient opioid analgesic vs placebo. Short term. Enriched design. Effect size is mean difference, calculated using random effects inverse variance model in RevMan 5.3. Outcome is 0-100 pain-intensity scale. Placebo group Total was split to incorporate multi-arm trials as per [17]. Medicines in published data subgroup are as per [17]. NCT01358526 is oxycodone/naloxone.

EUCTR2012-001920-36 (n=607) and NCT01455519 (n=26) contributed an additional 633 participants to the non-enriched designs analysis, an increase of 136.59%. The inclusion of unpublished data from EUCTR2012-001920-36 and NCT01455519 reduced the pooled effect size of the non-enriched designs meta-analysis by 1.1% from -7.27 (95%CI -9.97 to -4.57) to -7.19 (95%CI -9.24 to -5.14). Tau² reduced from 3.63 to zero (Figure 3). There was no difference identified in effect size between the published and unpublished trials (Chi²=0.15, p=.70) (Table 1).

Study or Subgroup	Opioid Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
2.2.1 Published data included in SR									
Buynak 2010_a	46.2	17.5	171	53.8	19	88	18.5%	-7.60 [-12.36, -2.84]	
Buynak 2010_b	47.4	20.6	201	53.8	19	89	17.7%	-6.40 [-11.27, -1.53]	
Chu 2012	-21.1	15.9	48	-12.5	19.2	55	9.1%	-8.60 [-15.38, -1.82]	
Cloutier 2013	48.6	23.1	32	55.9	25.4	31	2.9%	-7.30 [-19.30, 4.70]	
Gordon 2010	40.9	20.2	37	48	22.5	42	4.7%	-7.10 [-16.52, 2.32]	
Ueberall 2012	-21	20	107	-20	18	110	16.3%	-1.00 [-6.07, 4.07]	
Webster 2006_a	42	25.5	206	54	28.7	34	4.0%	-12.00 [-22.26, -1.74]	
Webster 2006_b	41	25.1	206	54	28.7	33	3.9%	-13.00 [-23.37, -2.63]	
Webster 2006_c	39	25.3	206	54	28.7	34	4.0%	-15.00 [-25.25, -4.75]	
Subtotal (95% CI)			1214			516	81.2%	-7.27 [-9.97, -4.57]	

Heterogeneity: Tau² = 3.63; Chi² = 10.26, df = 8 (P = 0.25); I² = 22%
Test for overall effect: Z = 5.27 (P < 0.00001)

Study or Subgroup	Opioid Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
2.2.2 Unpublished data from trial registries									
EUCTR2012-001920-36_a	-29.5	25.64	122	-21.6	23.96	31	4.6%	-7.90 [-17.48, 1.68]	
EUCTR2012-001920-36_b	-29.5	27.11	120	-21.6	23.96	31	4.4%	-7.90 [-17.63, 1.83]	
EUCTR2012-001920-36_c	-31.8	28.7	117	-21.6	23.96	32	4.4%	-10.20 [-20.00, -0.40]	
EUCTR2012-001920-36_d	-30.5	25.46	123	-21.6	23.96	31	4.6%	-8.90 [-18.46, 0.66]	
NCT01455519	-10.67	29.32	18	-10.57	26.98	8	0.8%	-0.10 [-23.19, 22.99]	
Subtotal (95% CI)			500			133	18.8%	-8.35 [-13.08, -3.62]	

Heterogeneity: Tau² = 0.00; Chi² = 0.66, df = 4 (P = 0.96); I² = 0%
Test for overall effect: Z = 3.46 (P = 0.0005)

Study or Subgroup	Opioid Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
Total (95% CI)			1714			649	100.0%	-7.19 [-9.24, -5.14]	

Heterogeneity: Tau² = 0.00; Chi² = 11.20, df = 13 (P = 0.59); I² = 0%
Test for overall effect: Z = 6.88 (P < 0.00001)
Test for subgroup differences: Chi² = 0.15, df = 1 (P = 0.70), I² = 0%

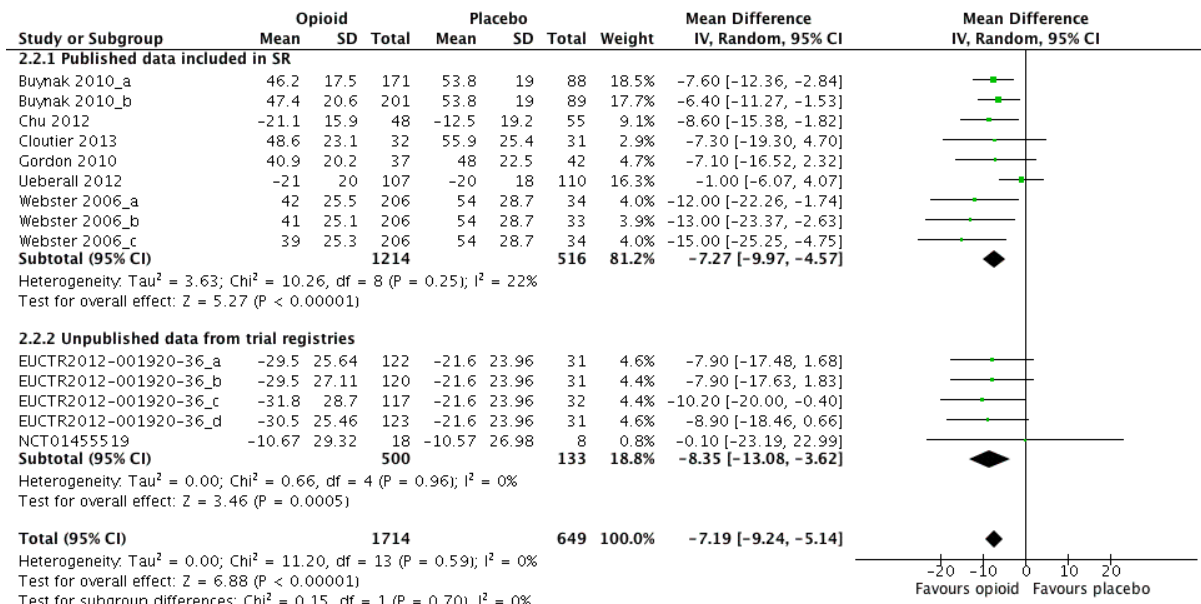-20 -10 0 10 20
Favours opioid Favours placebo

Figure 3. Single ingredient opioid analgesic vs placebo. Short term. Nonenriched design. Effect size is mean difference, calculated using random effects inverse variance model in RevMan 5.3. Outcome is 0-100 pain-intensity scale. Placebo group Total was split to incorporate multi-arm trials as per [17]. Medicines in published data subgroup are as per [17]. EUCTR2012-001920-36_a is cebranopadol 200mg, _b is cebranopadol 400mg, _c is cebranopadol 600mg, _d is tapentadol and NCT01455519 is hydromorphone ER.

NCT01358526 (n=600), EUCTR2012-001920-36 (n=607) and NCT01455519 (n=26) contributed an additional 1233 participants to the combined designs analysis, an increase of 136.1%. For the combined analysis of both enriched and non–enriched designs the inclusion of unpublished data from EUCTR2012-001920-36, NCT01358526, and NCT01455519 reduced the pooled effect size by 7.82% from -10.10 (95%CI -12.81 to -7.39) to -9.31 (95%CI -11.51 to -7.11). Tau² reduced from 19.73 to 14.46 (Figure 4). The difference in effect size between the published and unpublished trials was statistically significant (Chi²=4.98, p=.03) (Table 1).
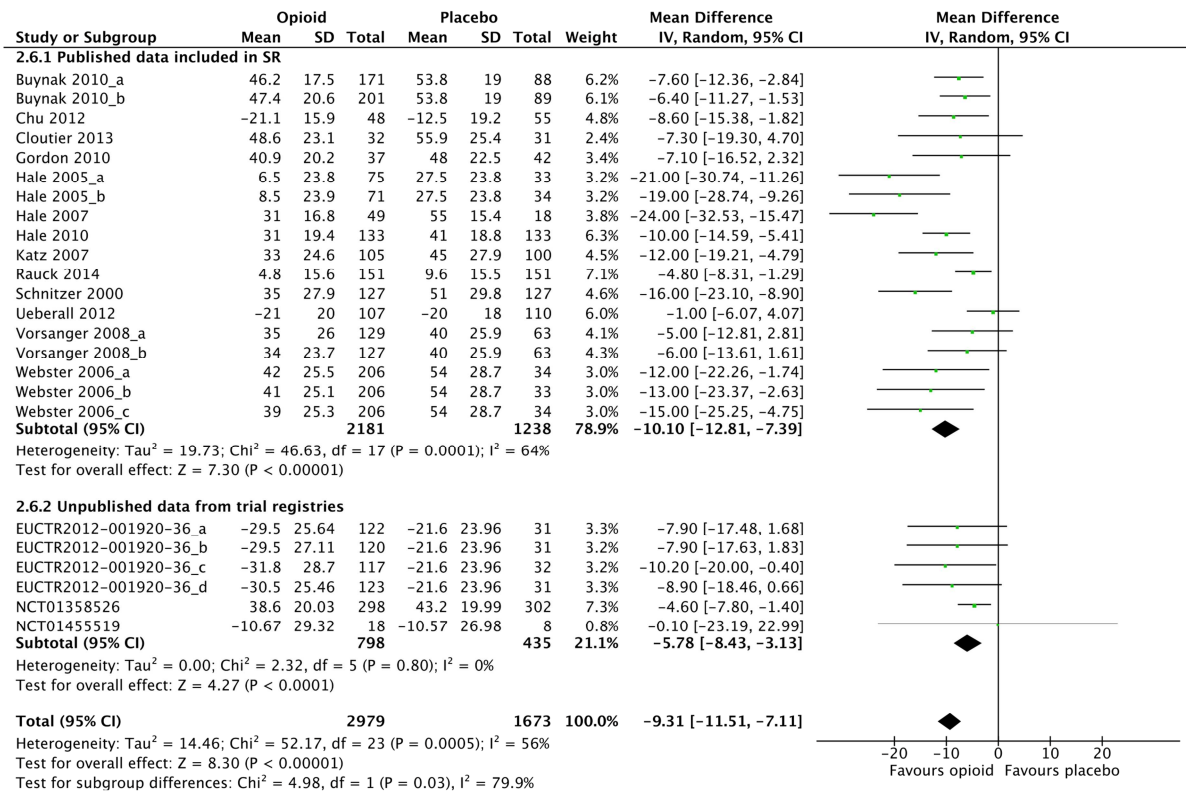
Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

19

Study or Subgroup	Opioid Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
2.6.1 Published data included in SR									
Buynak 2010_a	46.2	17.5	171	53.8	19	88	6.2%	−7.60 [−12.36, −2.84]	
Buynak 2010_b	47.4	20.6	201	53.8	19	89	6.1%	−6.40 [−11.27, −1.53]	
Chu 2012	−21.1	15.9	48	−12.5	19.2	55	4.8%	−8.60 [−15.38, −1.82]	
Cloutier 2013	48.6	23.1	32	55.9	25.4	31	2.4%	−7.30 [−19.30, 4.70]	
Gordon 2010	40.9	20.2	37	48	22.5	42	3.4%	−7.10 [−16.52, 2.32]	
Hale 2005_a	6.5	23.8	75	27.5	23.8	33	3.2%	−21.00 [−30.74, −11.26]	
Hale 2005_b	8.5	23.9	71	27.5	23.8	34	3.2%	−19.00 [−28.74, −9.26]	
Hale 2007	31	16.8	49	55	15.4	18	3.8%	−24.00 [−32.53, −15.47]	
Hale 2010	31	19.4	133	41	18.8	133	6.3%	−10.00 [−14.59, −5.41]	
Katz 2007	33	24.6	105	45	27.9	100	4.5%	−12.00 [−19.21, −4.79]	
Rauck 2014	4.8	15.6	151	9.6	15.5	151	7.1%	−4.80 [−8.31, −1.29]	
Schnitzer 2000	35	27.9	127	51	29.8	127	4.6%	−16.00 [−23.10, −8.90]	
Ueberall 2012	−21	20	107	−20	18	110	6.0%	−1.00 [−6.07, 4.07]	
Vorsanger 2008_a	35	26	129	40	25.9	63	4.1%	−5.00 [−12.81, 2.81]	
Vorsanger 2008_b	34	23.7	127	40	25.9	63	4.3%	−6.00 [−13.61, 1.61]	
Webster 2006_a	42	25.5	206	54	28.7	34	3.0%	−12.00 [−22.26, −1.74]	
Webster 2006_b	41	25.1	206	54	28.7	33	3.0%	−13.00 [−23.37, −2.63]	
Webster 2006_c	39	25.3	206	54	28.7	34	3.0%	−15.00 [−25.25, −4.75]	
Subtotal (95% CI)			2181			1238	78.9%	−10.10 [−12.81, −7.39]	

Heterogeneity: Tau² = 19.73; Chi² = 46.63, df = 17 (P = 0.0001); I² = 64%
Test for overall effect: Z = 7.30 (P < 0.00001)

Study or Subgroup	Opioid Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
2.6.2 Unpublished data from trial registries									
EUCTR2012-001920-36_a	−29.5	25.64	122	−21.6	23.96	31	3.3%	−7.90 [−17.48, 1.68]	
EUCTR2012-001920-36_b	−29.5	27.11	120	−21.6	23.96	31	3.2%	−7.90 [−17.63, 1.83]	
EUCTR2012-001920-36_c	−31.8	28.7	117	−21.6	23.96	32	3.2%	−10.20 [−20.00, −0.40]	
EUCTR2012-001920-36_d	−30.5	25.46	123	−21.6	23.96	31	3.3%	−8.90 [−18.46, 0.66]	
NCT01358526	38.6	20.03	298	43.2	19.99	302	7.3%	−4.60 [−7.80, −1.40]	
NCT01455519	−10.67	29.32	18	−10.57	26.98	8	0.8%	−0.10 [−23.19, 22.99]	
Subtotal (95% CI)			798			435	21.1%	−5.78 [−8.43, −3.13]	

Heterogeneity: Tau² = 0.00; Chi² = 2.32, df = 5 (P = 0.80); I² = 0%
Test for overall effect: Z = 4.27 (P < 0.0001)

Study or Subgroup	Opioid Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
Total (95% CI)			2979			1673	100.0%	−9.31 [−11.51, −7.11]	

Heterogeneity: Tau² = 14.46; Chi² = 52.17, df = 23 (P = 0.0005); I² = 56%
Test for overall effect: Z = 8.30 (P < 0.00001)
Test for subgroup differences: Chi² = 4.98, df = 1 (P = 0.03), I² = 79.9%

−20 −10 0 10 20
Favours opioid Favours placebo

Figure 4. Single ingredient opioid analgesic vs placebo. Short term. Enriched and nonenriched designs. Effect size is mean difference, calculated using random effects inverse variance model in RevMan 5.3. Outcome is 0-100 pain-intensity scale. Placebo group Total was split to incorporate multi-arm trials as per [17]. Medicines in published data subgroup are as per [17]. NCT01358526 is oxycodone/naloxone, EUCTR2012-001920-36_a is cebranopadol 200mg, _b is cebranopadol 400mg, _c is cebranopadol 600mg, _d is tapentadol and NCT01455519 is hydromorphone ER.

4. Discussion

We aimed to determine the effect of including publicly available, unpublished data from trial registries on the effect sizes of published meta-analyses of the most commonly used analgesic medicines for LBP. We clearly show that such data reduced the effect sizes of muscle relaxant

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

20

and opioid-analgesic medicines for LBP. These results have implications for future evidence synthesis and for the interpretation of evidence in this field.

The biggest impact of including unpublished trial registry data was observed for muscle relaxant medicines - an 87% decrease in effect size. The minimal clinically important difference between groups for back pain intensity is considered to be 10 points on a 0-100 point scale [24,42]. The original review reported an effect size greater than this [-21.71 (95%CI -28.23 to -15.19)] [18]. However, the addition of unpublished data yielded a revised effect size that, although statistically significant, was well short of the clinically meaningful threshold [-2.34 (95%CI -3.34 to -1.34)] and with much less heterogeneity. Clearly, adding unpublished data changes the precision and clinical meaning of the effect size. The Abdel Shaheed study [18] is the largest and most recent review of these medicines for LBP. Updates of this review or future syntheses of evidence for these medicines should include unpublished data.

These implications extend to future clinical guidelines. Clinical guidelines have an important role in directing the use of appropriate interventions in practice. The recommendations are often informed by systematic-reviews. The current American College of Physicians/American Pain Society clinical guideline for the management of LBP makes a strong recommendation for the use of muscle relaxant medicines for recent-onset LBP [43]. This recommendation is based on data from a systematic-review, published in 2003 [44], and additional update searches [24]. Although data were analysed differently, the 2003 review reported an effect of a similar magnitude as that reported in the recent review sampled in our study. We note that the updated search included unpublished data sources, although none were identified for muscle relaxants. Clearly, guideline developers should search for and include unpublished data if conducting systematic-reviews or be more cautious of their interpretation of pooled estimates taken from systematic-reviews that do not include unpublished data.

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

21

We have also shown that the effect estimates for opioid analgesic medicines reduced with the inclusion of unpublished data. The authors of the original review concluded that opioid analgesics probably do not provide clinically meaningful effects at recommended doses based on the results of the combined designs analysis [-10.10 (95%CI -12.81 to -7.39)] [17]. The reduced size and heterogeneity and improved precision of the effect estimate [-9.31 (95%CI -11.51 to -7.11)] with the inclusion of unpublished data substantiate the conclusion of the original review.

Our results also have implications for the judgements of confidence in evidence that are made by systematic review teams. These judgements interpret the meta-analytic effect estimates in the context of possible bias, observed heterogeneity and missing data. A well-known framework for making these judgements is GRADE [45]. The original analyses of muscle relaxants and opioid analgesics were both judged using the GRADE framework; whereby the review teams judged the evidence to be high quality for the effect of muscle relaxant medicines [18] and moderate quality for the effect of opioid analgesic medicines [17] on short term pain intensity, compared to a placebo medicine. In both cases, the additional unpublished trials have implications for the judgement of confidence in the effect estimates from the updated analyses. However, examination of these implications is beyond the scope of this study.

Our work has a number of strengths. We prospectively registered and followed a protocol [38], now recommended practice in the clinical pain sciences [46]. We have reported all protocol deviations and *post hoc* decisions, to facilitate clear understanding of our work. We used sensitive search strategies for a large number of currently licensed analgesic medicines. We reproduced the methods of the previous reviews as precisely as possible, which increases the likelihood that our results are due to the influence of the unpublished data and not to the reproduction process. Additionally, we extracted data in duplicate and cross-checked all of our analyses.

Our work has some limitations. We restricted the scope of unpublished data sources to trial registrations. There are multiple sources of data from controlled clinical trials; websites, regulatory agencies, direct contact with trialists, and data sharing networks [26,47,48]. The sample we have drawn may not be representative of all unpublished data. However, trial registries are free to access, require the least resources to search and obtain data and do not invoke issues of propriety or confidentiality. These are relevant considerations for review production teams. Further, we aimed only to determine the influence of unpublished data, rather than update the previous reviews. For these reasons, we are confident that the restriction of our sampling to trial registries is justified. We welcome further work to evaluate the influence of other sources of unpublished data.

A related limitation of this study is that the majority (n=16) of, completed, unpublished trials (n=24) that we identified as eligible for inclusion in the original reviews did not have data available on the registry. This is a known problem [49] and there is the potential that the results would have been different if more data were included. For example, we could not update any analyses in Machado et al. [20], although 4 trials were eligible. Further, we observed a small change in effect size with the addition of 3 trials of opioid analgesics, yet there were a further 5 eligible trials. We may have been able to obtain more data had we contacted the sponsors of the relevant trials. However, access to these data is not universal, may require significant effort to obtain and may incur restrictions on use [47,48,50,51]. Regardless, our results demonstrate that it will be important for future work to include these records.

We developed a catalogue of medicines for this study. We note that eperisone is a muscle relaxant medicine included in the Abdel Shaheed review [18] but not in our catalogue. We searched *post hoc* for trials of eperisone for back pain and identified 14 trials. Three of these trials

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

23

do not meet the inclusion criteria for this study and 11 do not have data available. Thus, the exclusion of eperisone does not impact our conclusions.

We established publication status using a three-step rule. There is the potential that we misclassified studies as unpublished that have in fact been published. This is most likely because the published report does not include the trial registry identification number. We note that alternative methods for identifying publication status have been reported [8,51–56]. These exhibit some overlap with our methods, primarily the use of the trial identification number. Regardless of the methods used, accurate reporting is important to facilitate understanding of these judgements. This standard of reporting is in line with recommendations [35,57–59].

We note that five systematic-reviews of analgesic medicines for LBP have been published [22–25,60] subsequent to the registration of our protocol. Four of these reviews [22–25] searched trial registries for reports of unpublished trials (although, no trials were available for meta-analysis). This indicates that research practice in the field may be changing. The results of our study provide the first evidence to support this change in practice.

5. Conclusion

Most systematic-reviews of analgesic medicines for LBP do not include unpublished data from clinical trial registries. Including unpublished data in recent meta-analyses reduced the estimate of treatment effect in all comparisons for which new data were available. This suggests that current systematic-reviews of medicines for LBP likely overestimate the benefits of these drugs.

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

24

6. Acknowledgements:

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

25

had no role in the conception or writing of the manuscript and there are no other relationships

or activities that could appear to have influenced the submitted work.

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

26

7. References

[1]   Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical

        trials due to statistical significance or direction of trial results. Cochrane Database Syst

        Rev 2009:5–8. https://doi.org/10.1002/14651858.MR000006.pub3.

[2]   Canestaro WJ, Hendrix N, Bansal A, Sullivan SD, Devine EB, Carlson JJ. Favorable and

        publicly funded studies are more likely to be published: a systematic review and meta-

        analysis. J Clin Epidemiol 2017;92:58–68. https://doi.org/10.1016/j.jclinepi.2017.08.004.

[3]   Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic Review of the Empirical

        Evidence of Study Publication Bias and Outcome Reporting Bias - An Updated Review.

        PLoS One 2013;8. https://doi.org/10.1371/journal.pone.0066844.

[4]   Hopewell S, McDonald S, Clarke Mike J, Egger M. Grey literature in meta-analyses of

        randomized trials of health care interventions. Cochrane Database Syst Rev 2007.

        https://doi.org/10.1002/14651858.MR000010.pub3.

[5]   Baudard M, Yavchitz A, Ravaud P, Perrodeau E, Boutron I. Impact of searching clinical

        trial registries in systematic reviews of pharmaceutical treatments: methodological

        systematic review and reanalysis of meta-analyses. BMJ 2017;356.

        https://doi.org/10.1136/bmj.j448.

[6]   Schmucker CM, Blümle A, Schell LK, Schwarzer G, Oeller P, Cabrera L, et al. Systematic

        review finds that study data not published in full text articles have unclear impact on

        meta-analyses results in medical research. PLoS One 2017;12:1–16.

        https://doi.org/10.1371/journal.pone.0176210.

[7]   Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials:

        reanalysis of meta-analyses. BMJ 2012;344:d7202. https://doi.org/10.1136/bmj.d7202.

[8]   Wilson LM, Sharma R, Dy SM, Waldfogel JM, Robinson KA. Searching ClinicalTrials.gov

        did not change the conclusions of a systematic review. J Clin Epidemiol 2017;90:127–35.

        https://doi.org/10.1016/j.jclinepi.2017.07.009.

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

27

[9]   Dechartres A, Atal I, Riveros C, Meerpohl J, Ravaud P. Association Between Publication

        Characteristics and Treatment Effect Estimates. Ann Intern Med 2018;169:385.

        https://doi.org/10.7326/M18-1517.

[10]   Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective Publication of

        Antidepressant Trials and Its Influence on Apparent Efficacy. N Engl J Med

        2008;358:252–60. https://doi.org/10.1056/NEJMc080313.

[11]   Trinquart L, Abbé A, Ravaud P. Impact of reporting bias in network meta-analysis of

        antidepressant placebo-controlled trials. PLoS One 2012;7:1–8.

        https://doi.org/10.1371/journal.pone.0035219.

[12]   Gore M, Tai KS, Sadosky A, Leslie D, Stacey BR. Use and Costs of Prescription

        Medications and Alternative Treatments in Patients with Osteoarthritis and Chronic Low

        Back Pain in Community-Based Settings. Pain Pract 2012;12:550–60.

        https://doi.org/10.1111/j.1533-2500.2012.00532.x.

[13]   Hart OR, Uden RM, McMullan JE, Ritchie MS, Williams TD, Smith BH. A study of

        National Health Service management of chronic osteoarthritis and low back pain. Prim

        Health Care Res Dev 2015;16:157–66. https://doi.org/10.1017/S1463423614000140.

[14]   Williams CM, Maher CG, Hancock MJ, McAuley JH, McLachlan AJ, Britt H, et al. Low

        back pain and best practice care: A survey of general practice physicians. Arch Intern Med

        2010;170:271–7. https://doi.org/10.1001/archinternmed.2009.507.

[15]   Ivanova JI, Birnbaum HG, Schiller M, Kantor E, Johnstone BM, Swindle RW. Real-world

        practice patterns, health-care utilization, and costs in patients with low back pain: the long

        road to guideline-concordant care. Spine J 2011;11:622–32.

        https://doi.org/10.1016/j.spinee.2011.03.017.

[16]   Mathieson S, Valenti L, Maher CG, Britt H, Li Q, McLachlan AJ, et al. Worsening trends

        in analgesics recommended for spinal pain in primary care. Eur Spine J 2018;27:1136–45.

        https://doi.org/10.1007/s00586-017-5178-4.

[17]   Abdel Shaheed C, Maher CG, Williams KA, Day R, McLachlan AJ. Efficacy, Tolerability, and Dose-Dependent Effects of Opioid Analgesics for Low Back Pain. JAMA Intern Med 2016;176:958. https://doi.org/10.1001/jamainternmed.2016.1251.

[18]   Abdel Shaheed C, Maher CG, Williams KA, McLachlan AJ. Efficacy and tolerability of muscle relaxants for low back pain: Systematic review and meta-analysis. Eur J Pain 2017;21:228–37. https://doi.org/10.1002/ejp.907.

[19]   Saragiotto BT, Machado GC, Ferreira ML, Pinheiro MB, Abdel Shaheed C, Maher CG. Paracetamol for low back pain. Cochrane Database Syst Rev 2016. https://doi.org/10.1002/14651858.CD012230.

[20]   Machado GC, Maher CG, Ferreira PH, Day RO, Pinheiro MB, Ferreira ML. Non-steroidal anti-inflammatory drugs for spinal pain: a systematic review and meta-analysis. Ann Rheum Dis 2017;76:1269–78. https://doi.org/10.1136/annrheumdis-2016-210597.

[21]   Enthoven WT, Roelofs PD, Deyo RA, van Tulder MW, Koes BW. Non-steroidal anti-inflammatory drugs for chronic low back pain. Cochrane Database Syst Rev 2016. https://doi.org/10.1002/14651858.CD012087.

[22]   Shanthanna H, Gilron I, Rajarathinam M, AlAmri R, Kamath S, Thabane L, et al. Benefits and safety of gabapentinoids in chronic low back pain: A systematic review and meta-analysis of randomized controlled trials. PLoS Med 2017;14:1–21. https://doi.org/10.1371/journal.pmed.1002369.

[23]   Mathieson S, Kasch R, Maher CG, Pinto RZ, McLachlan AJ, Koes BW, et al. Combination drug therapy for the management of low back pain and sciatica; systematic review and meta-analysis. J Pain 2018. https://doi.org/10.1016/j.jpain.2018.06.005.

[24]   Chou R, Deyo R, Friedly J, Skelly A, Weimer M, Fu R, et al. Systemic Pharmacologic Therapies for Low Back Pain: A Systematic Review for an American College of Physicians Clinical Practice Guideline. Ann Intern Med 2017. https://doi.org/10.7326/M16-2458.

[25]   Enke O, New HA, New CH, Mathieson S, McLachlan AJ, Latimer J, et al.
       Anticonvulsants in the Treatment of Low Back Pain and Lumbar Radicular Pain: a
       Systematic Review and Meta-Analysis. CMAJ 2018;190190:786–93.
       https://doi.org/10.1503/cmaj.171333.

[26]   Isojarvi J, Wood H, Lefebvre C, Glanville J. Challenges of identifying unpublished data
       from clinical trials: Getting the best out of clinical trials registers and other novel sources.
       Res Synth Methods 2018:1–18. https://doi.org/10.1002/jrsm.1294.

[27]   Higgins JPT, Lasserson T, Chandler J, Tovey D, Churchill R. Methodological
       Expectations of Cochrane Intervention Reviews. London: 2016.

[28]   Shmagel A, Ngo L, Ensrud K, Foley R. Prescription Medication Use among Community-
       Based US Adults with Chronic Low Back Pain: a Cross-Sectional Population Based Study.
       J Pain 2018. https://doi.org/10.1016/j.jpain.2018.04.004.

[29]   Thackeray A, Hess R, Dorius J, Brodke D, Fritz J. Relationship of Opioid Prescriptions to
       Physical Therapy Referral and Participation for Medicaid Patients with New-Onset Low
       Back Pain. J Am Board Fam Med 2017;30:784–94.
       https://doi.org/10.3122/jabfm.2017.06.170064.

[30]   Michaleff ZA, Harrison C, Britt H, Lin CWC, Maher CG. Ten-year survey reveals
       differences in GP management of neck and back pain. Eur Spine J 2012;21:1283–9.
       https://doi.org/10.1007/s00586-011-2135-5.

[31]   Oliveira CB, Maher CG, Pinto RZ, Traeger AC, Lin CWC, Chenot JF, et al. Clinical
       practice guidelines for the management of non-specific low back pain in primary care: an
       updated overview. Eur Spine J 2018;27:1–13. https://doi.org/10.1007/s00586-018-5673-
       2.

[32]   Australian Register of Therapeutic Goods n.d. http://tga-
       search.clients.funnelback.com/s/search.html?query=&collection=tga-artg.

[33]   Drugs@FDA n.d. www.accessdata.fda.gov/scripts/cder/daf/index.cfm.

[34] European Medicines Agency search portal n.d. http://www.ema.europa.eu/ema/index.jsp?curl=pages/includes/medicines/medicines_landing_page.jsp&mid=.

[35] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP a, et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. Ann Intern Med 2009;151:W-65-W-94.

[36] Borenstein M, Hedges L V, Higgins JPT, Rothstein HR. Introduction to meta-analysis. Chichester, United Kingdom: John Wiley & Sons, Ltd.; 2009.

[37] Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions 2011. http://handbook.cochrane.org/.

[38] Bagg MK, O'Hagan E, Wand BM, Hubscher M, Moseley GL, McAuley JH. Does data from trial registries influence effect size? A sensitivity analysis of recent meta-analyses of pharmacological interventions for pain in people with low back pain [Study Protocol]. Open Sci Framew 2017.

[39] Berry H, Hutchinson DR. A Multicentre Placebo-controlled Study in General Practice to Evaluate the Efficacy and Safety of Tizanidine in Acute Low-back Pain. J Int Med Res 1988;16:75–82.

[40] Berry H, Hutchinson DR. Tizanidine and ibuprofen in acute low-back pain: results of a double-blind multicentre study in general practice. J Int Med Res 1988;16:83–91.

[41] Katz N. Enriched Enrollment Randomized Withdrawal Trial Designs of Analgesics Focus on Methodology. Clin J Pain 2009;25:797–807.

[42] Busse JW, Bartlett SJ, Dougados M, Johnston BC, Guyatt GH, Kirwan JR, et al. Optimal Strategies for Reporting Pain in Clinical Trials and Systematic Reviews: Recommendations from an OMERACT 12 Workshop. J Rheumatol 2015;42:1–9. https://doi.org/10.3899/jrheum.141440.

[43]   Qaseem A, Wilt TJ, McLean RM, Forciea MA. Noninvasive Treatments for Acute,

        Subacute, and Chronic Low Back Pain: A Clinical Practice Guideline From the American

        College of Physicians. Ann Intern Med 2017. https://doi.org/10.7326/M16-2367.

[44]   Van Tulder MW, Touray T, Furlan AD, Solway S, Bouter LM. Muscle relaxants for non-

        specific low-back pain (Review). Cochrane Database Syst Rev 2003.

        https://doi.org/10.1002/14651858.CD004252.

[45]   Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Schünemann HJ. GRADE : what is "

        quality of evidence " and why is it important to clinicians? Br Med J 2008;336:995–8.

[46]   Lee H, Lamb SE, Bagg MK, Toomey E, Cashin AG, Moseley GL. Reproducible and

        replicable pain research. Pain 2018. https://doi.org/10.1097/j.pain.0000000000001254.

[47]   Goldacre B, Lane S, Mahtani KR, Heneghan C, Onakpoya I, Bushfield I, et al.

        Pharmaceutical companies' policies on access to trial data, results, and methods: Audit

        study. BMJ 2017;358:1–12. https://doi.org/10.1136/bmj.j3334.

[48]   Goldacre B, Gray J. OpenTrials: Towards a collaborative open database of all available

        information on all clinical trials. Trials 2016;17:1–12. https://doi.org/10.1186/s13063-

        016-1290-8.

[49]   Huser V, Cimino JJ. Linking ClinicalTrials.gov and PubMed to Track Results of

        Interventional Human Clinical Trials. PLoS One 2013;8:1–9.

        https://doi.org/10.1371/journal.pone.0068409.

[50]   Herder M, Doshi P, Lemmens T. Precedent pushing practice: Canadian court orders

        release of unpublished clinical trial data. BMJ Opin 2018.

        https://blogs.bmj.com/bmj/2018/07/19/precedent-pushing-practice-canadian-court-

        orders-release-of-unpublished-clinical-trial-data/.

[51]   Bashir R, Bourgeois FT, Dunn AG. A systematic review of the processes used to link

        clinical trial registrations to their published results. Syst Rev 2017;6.

        https://doi.org/10.1186/s13643-017-0518-3.

[52]   Hartung DM, Zarin DA, Guise J-M, McDonagh M, Paynter R, Helfand M. Reporting

Discrepancies Between the ClinicalTrials.gov Results Database and Peer-Reviewed

Publications. Ann Intern Med 2014;160:477–84.

[53]   Altman DG, Furberg CD, Grimshaw JM, Shanahan DR. Linked publications from a

single trial: A thread of evidence. Trials 2014;15:1–3. https://doi.org/10.1186/1745-6215-

15-369.

[54]   Meddings K. Linking clinical trials = enriched metadata and increased transparency.

Crossref Blog 2016. https://www.crossref.org/blog/linking-clinical-trials--enriched-

metadata-and-increased-transparency/.

[55]   Shanahan D, Meddings K. Clinical trial data and articles linked for the first time. Crossref

Blog 2016. https://www.crossref.org/blog/linked-clinical-trials-are-here/.

[56]   Shanahan D. Threaded Publications: one step closer. Med 2014.

http://blogs.biomedcentral.com/on-medicine/2014/01/31/threaded-publications-one-

step-closer/.

[57]   Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred Reporting Items for Systematic

Reviews and Meta-Analyses: The PRISMA Statement. Ann Intern Med 2009;151:264–9.

[58]   Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred

reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015

statement. Syst Rev 2015;4:1. https://doi.org/10.1186/2046-4053-4-1.

[59]   Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred

reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015:

elaboration and explanation. BMJ 2015;349:g7647–g7647.

https://doi.org/10.1136/bmj.g7647.

[60]   Tucker H-R, Scaff K, McCloud T, Carlomagno K, Daly K, Garcia A, et al. Harms and

benefits of opioids for management of non-surgical acute and chronic low back pain: a

systematic review. Br J Sports Med 2019:bjsports-2018-099805.

Journal of Clinical Epidemiology
Systematic-reviews that include only published data overestimate effect sizes | Bagg et al.

33

https://doi.org/10.1136/bjsports-2018-099805.

Author Contributions:

MKB and JHM conceived the study. MKB, BMW and MH wrote the protocol. MKB searched for studies. MKB, EOH, MH and PZ screened studies for inclusion. MKB, EOH, BMW, MH and PZ extracted data. MKB conducted the analyses. PZ cross-checked the analyses. All authors interpreted the analyses. MKB wrote the first draft of the manuscript. All authors contributed to the manuscript and read and approved the final version. MKB is guarantor for this work. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

What is New?

What is already known:

- Systematic-reviews and meta-analyses that include unpublished data from trial registries may reach different conclusions to those that exclude unpublished data.

What are the new findings:

- Including unpublished data from trial registries changes the effect sizes reported in recently published meta-analyses of common analgesic medicines for low back pain

- Including unpublished data reduced the estimate of treatment effect in all comparisons for which new data were available.

- Current systematic-reviews of medicines for low back pain likely overestimate the benefits of these interventions