ELSEVIER

Journal of Clinical Epidemiology 53 (2000) 964–972

Journal of Clinical Epidemiology

What contributions do languages other than English make on the results of meta-analyses?

David Moher[a,b,*], Ba' Pham[a], Terry P. Klassen[a,b], Kenneth F. Schulz[c], Jesse A. Berlin[d], Alejandro R. Jadad[e], Alessandro Liberati[f],

[a]*Thomas C. Chalmers Center for Systematic Reviews, Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada*
[b]*Departments of Pediatrics, Epidemiology, and Community Medicine, University of Ottawa, Ottawa, Ontario, Canada*
[c]*Family Health International, and Department of Obstetrics and Gynecology, University of North Carolina, Chapel Hill, North Carolina, USA*
[d]*Center for Clinical Epidemiology and Biostatistics, and Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania, USA*
[e]*Health Information Research Unit, Department of Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada*
[f]*Clinical Epidemiology, Mario Negri Institute, Milan, Italy*

Received 22 June, 1999; accepted 5 January 2000

Abstract

Including only a portion of all available evidence may introduce systematic errors into the meta-analytic process and threaten its validity. We set out to examine whether language restricted meta-analyses, compared to language inclusive meta-analyses, provide different estimates of the effectiveness of interventions evaluated in randomized trials. We identified and retrieved all 79 meta-analyses from several disease areas in which explicit eligibility criteria regarding trial selection were reported. General characteristics and quality of reporting of the meta-analyses were assessed using a validated instrument. We explored the effects of language of publication of the randomized trials on the quantitative results using logistic regression analyses. Language restricted meta-analyses, compared to language inclusive meta-analyses, did not differ with respect to the estimate of benefit of the effectiveness of an intervention (ROR = 0.98; 95% CI: 0.81–1.17). These results were also robust after a series of sensitivity analyses. This study provides no evidence that language restricted meta-analyses lead to biased estimates of intervention effectiveness. We encourage others to replicate this study using different sampling frames, clinical topics and interventions.

Keywords: Meta-analysis; Randomized controlled trials; Methodology; Bias; Language of publication

1. Introduction

Meta-analysts have little control over random errors but can exert at least some control over systematic ones. Including only a portion of all available evidence may introduce systematic errors into the review process and threaten its validity. Grégoire and colleagues reported that 78% of identified meta-analyses of randomized trials had language of publication restrictions [1]. The majority (93%) of these restrictions were at the expense of excluding trials published in languages other than English (OEL).

One way to evaluate whether language restrictions are a sensible policy for meta-analysts is to assess the quality of reports of randomized trials. Language restrictions might be appropriate if the quality of reports of OEL were different compared to English language (EL) trials. In a previous study, members of our group [2] compared the methodological characteristics and analytical approaches of 133 EL randomized trials published between 1989 and 1994, with reports of 96 randomized trials published in French, German, Italian, and Spanish over the same time period and type of journal. Within the same language the reports were assessed under masked conditions, using a scale developed with appropriate rigorous standards [3].

The differences found in this study, between OEL and EL trials in the quality of reporting, with respect to randomization, allocation concealment, double-blinding, dropouts and withdrawals, or overall total score, were neither statistically nor substantively significant. The mean differences in the quality of reporting between OEL and EL trials was 5% for the total score, and ranged from 0% to 4% for individual items. Similar results have recently been reported elsewhere [4]. However, these studies did not address whether the exclusion of OEL alters the statistical results of a meta-analysis. This study addresses that question.

* Corresponding author. Thomas C. Chalmers Center for Systematic Reviews, Children's Hospital of Eastern Ontario Research Institute, Room R226, 401 Smyth Road, Ottawa, Ontario K1H 8L1, Canada. Tel: (613) 738-3591; fax: (613) 738-4869.

E-mail address: dmoher@uottawa.ca

0895-4356/00/$ – see front matter
PII: S0895-4356(00)00188-8

2. Methods

2.1. Selection of meta-analyses

We selected our meta-analyses from a collection of 251 meta-analyses of randomized trials from a larger database (n = 455) of such studies [5]. The details of the search strategy used to identify and retrieve the meta-analyses included in this study are reported elsewhere [6,7]. Briefly, a refined MEDLINE search strategy identified the meta-analyses. This search was supplemented with a search of the Cochrane Database of Systematic Reviews (1996, Issue 1). Meta-analyses were eligible if they included between 2 and 99 randomized trials and reported binary outcomes. Three types of meta-analyses were included: those in which OEL were explicitly excluded (language restricted meta-analysis), those that explicitly permitted the inclusion of OEL but in which no OEL contributed to the quantitative analysis (language inclusive meta-analysis/EL), and those that actually included OEL in the quantitative summary. Identified OEL were translated, as required, within our group (French, Italian, and Spanish), through our network of personal contacts, or using a professional translation service.

2.2. Quality assessment of meta-analyses

Once all the meta-analyses were retrieved they were masked to author and any author affiliation, journal, references, and other potential identifiers. The quality of report of each meta-analysis was assessed using a validated scale [8]. This instrument includes nine items pertaining to individual aspects in the reporting of a meta-analysis (e.g., were the search methods used to find evidence on the primary question stated?). Each item is assessed using a three point scale (i.e., no, partially/can't tell, or yes). A final question elicits an overall scientific quality of the meta-analysis. The scoring ranges from 1 to 7 with higher scores indicating superior quality.

We standardized ourselves in using the instrument and pre-tested our methods by completing an inter-observer reliability study [9]. Agreement was assessed with the intraclass correlation coefficient (ICC), using a separate set of 10 meta-analyses. Values above 0.61 were considered as substantial agreement [10], based on an *a priori* decision.

2.3. Data extraction

In addition to quality assessment of each meta-analysis, the following data were extracted using a structured form: the disease category under investigation (using ICD-10 codes); the number of randomized trials; language of publication of these trials; whether the authors reported assessing the presence of publication bias and, if so, the method used; year of publication, funding source, and type of journal (i.e., general and internal medicine, or specialty). We also collected information about the journal in which the meta-analysis was published, including its citation impact factor.

From each language inclusive meta-analysis/OEL we extracted from each randomized trial the number of events and patients in the control group, and the number of events and patients in the experimental group. We also collected information on the year the trial was published, and the disease category under investigation. These data were extracted independently by two research coordinators. Two investigators (DM, BP) independently reviewed the data extraction and consensus among the four data extractors was achieved for any discrepancies before data entry.

2.4. Data analyses

2.4.1. Between meta-analyses comparisons

We used a Fisher's exact test to compare the three types of meta-analyses with respect to the adequate reporting on each of the first nine items of the Oxman-Guyatt scale. Similarly, the three types of meta-analyses were compared on their overall scientific quality of reporting using the Kruskal-Wallis test.

2.4.2. Within language inclusive/OEL meta-analyses comparisons

We used logistic regression to assess the effect of language of publication on the estimates of intervention effect across the included meta-analyses. The details of this analysis, including model specification, are included in Appendix A. The language effect from the logistic regression is reported as a ratio of odds ratios (ROR) [11]. By our modeling convention, an ROR below one, of language inclusive meta-analyses compared to language restrictive meta-analyses, indicates that OEL report a larger intervention effect.

We also performed sensitivity analyses to further explore whether the results remain robust for meta-analyses that only include one or more OEL, and for varying sample sizes by setting a range of threshold values for the sample size required of a trial for inclusion in the meta-analysis.

For each of the language inclusive/OEL meta-analyses we calculated the width of the 95% confidence interval of the combined odds ratio on a log scale (log upper confidence limit for the OR − log of lower confidence limit for the OR). This calculation was repeated excluding the OEL (i.e., language restricted meta-analyses). We compared the average confidence interval width of the language inclusive to that of the language restrictive meta-analyses using a paired t-test.

All continuous distributions were summarized by the median (inter-quartile range). For all analyses, 2-sided P values $\leqslant 5\%$ were considered statistically significant.

As a quality control check to verify the data elements used in our data analyses, we replicated all language inclusive/OEL meta-analyses using the same analytical procedures reported by the authors of the original publication. This step was necessary to ensure reliable data elements required for our primary analysis.

3. Results

3.1. Between meta-analyses comparisons

Of 251 meta-analyses reviewed, 79 met our inclusion criteria (Table 1 and Appendix B). Only 19 meta-analyses actually

Table 1
General characteristics of language inclusive meta-analyses and language restricted meta-analyses (see text for details)

	Language inclusive meta-analyses/OEL (n = 19) n (%)	Language inclusive meta-analyses/EL (n = 22) n (%)	Language restricted meta-analyses (n = 38) n (%)
Disease area			
Infectious Disease	1 (5.3)	7 (31.8)	2 (5.3)
Circulatory Disease	3 (15.8)	7 (31.8)	16 (42.1)
Complications in Pregnancy and Childbirth	2 (10.5)	4 (18.2)	2 (5.3)
Other	13 (68.4)	4 (18.2)	18 (47.4)
Year of publication			
Median (1st, 3rd Quartiles)	1994 (1992, 1995)	1994 (1992, 1995)	1994 (1992, 1994)
Funding source			
Single pharmaceutical company	0	0	1 (2.6)
Non-pharmaceutical comapany	8 (42.1)	12 (54.5)	16 (42.1)
None listed/ can't tell	11 (57.9)	10 (45.5)	21 (55.3)
General characteristics			
Number of randomized trials - median (1st, 3rd quartiles)	9 (6.5, 18.00)	7.00 (4.75, 10.50)	6.00 (4.0, 9.25)
Journal's citation impact: median (1st, 3rd quartiles)	3.09 (1.83, 4.94)	3.21 (2.40, 8.33)	2.65 (1.54, 5.48)
Evaluated publication bias	6 (27.2)	2 (11.2)	7 (30.5)
Type of journal			
General medical	8 (42.1)	13 (59.1)	11 (28.9)
Specialty	11 (57.9)	9 (40.9)	27 (71.1)

included OEL in their quantitative analysis. One of these meta-analyses was dropped from further analysis because it did not report any binary outcomes, leaving 18 for further analyses. The remaining meta-analyses either had no language restrictions but did not incorporate OEL into a quantitative analysis (n = 22), or explicitly excluded such trials from their study (n = 38). Descriptively the three groups of meta-analyses were similar in terms of their general characteristics, with a couple of exceptions. Meta-analyses that included OEL into the quantitative analysis included more randomized trials

Table 2
Quality of reports of language inclusive meta-analyses and language restricted meta-analyses (see text for details)

Question	Language inclusive meta-analyses/OEL (n = 19) n (%)	Language inclusive meta-analyses/EL (n = 22) n (%)	Language restricted meta-analyses (n = 38) n (%)	2-sided P value
1. Were the search methods used to find evidence reported?	7 (37)	10 (48)	24 (63)	0.15*
2. Was the search for evidence reasonable comprehensive?	7 (37)	5 (24)	17 (45)	0.32
3. Were the criteria for deciding which studies to include in the overview reported?	12 (63)	17 (81)	29 (76)	0.16
4. Was bias in the selection of studies avoided?	4 (21)	9 (43)	9 (24)	0.23
5. Were the criteria used for assessing the validity of the included studies reported?	9 (47)	5 (24)	13 (34)	0.29
6. Was the validity of all of the studies referred to in the text assessed using appropriate criteria?	9 (47)	6 (29)	10 (26)	0.26
7. Were the methods to combine the findings of the relevant studies reported?	13 (68)	15 (71)	30(79)	0.63
8. Were the findings of the relevant studies combined appropriately relative to the priamry question the overview addresses?	13 (68)	14 (67)	32 (84)	0.21
9. Were the conclusions made by the author(s) supported by the data and/or analysis reported in the overview?	13 (68)	16 (76)	29 (76)	0.79
10. How would you rate the scientific quality of this overview**	4 (3.0, 4.0)	3 (2.5, 4.5)	3 (3.0, 5.25)	0.93***

*Fisher's Exact test; **Kruskal-Wallis test? ***Median (Interquartile range).

Table 3
Descriptive characteristics of meta-analyses including one or more randomized trial published in a language other than English.

	Language inclusive meta-analyses/OEL OEL = 1 (n = 13)	Language inclusive meta-analyses/OEL OEL >1 (n = 6)
Language	Dutch = 1 French = 5 German = 5 Italian = 1 Spanish = 1	Chinese = 1 French = 7 German = 7 Italian = 3 Spanish = 1 Danish = 1
Clinical area	Circulatory = 1 Digestive = 2 Genitourinary = 2 Complications in pregnancy = 2 Ill-defined = 2 Infectious Disease = 1 Mental Disease = 1 Neoplasm = 2	Circulatory = 2 Digestive = 1 Genitourinary = 1 Endocrinology = 1 Musculoskeletal = 1
Number of randomized trials per meta-analysis	8 (6,9)*	17 (16,22)*
Number of participants per study	108 (61, 209)*	58 (36,100)*
Number of participants per MA	908 (612, 3836)*	1224 (1141, 1713)*

*Median (1, 3Q).

compared to language restricted meta-analyses. There was also some descriptive differences among the three groups of meta-analyses in terms of their disease areas.

We established substantial agreement among members of the research team with respect to assessing the quality of reports of the meta-analyses (ICC = 0.63). The quality of reports was similar across the three groups of meta-analyses (Table 2). The overall scientific quality of reports of the three groups of meta-analyses was low (median = 3 of a possible 7; inter quartile range: 3, 5) and similar (2P = 0.93).

3.2. Within language inclusive/OEL meta-analyses comparisons

Most (68.4%) of the language inclusive meta-analyses only included one OEL (Table 3). The 33 OEL were published in one of seven languages. There was a similar language distribution across meta-analyses regardless of the number of OEL included. Meta-analyses with more than one OEL, compared to those with only one, included more randomized trials although average sample size per trial in meta-analyses with more than one OEL was smaller (Table 3). The average cumulative sample of a meta-analysis was 908 (inter-quartile range: 612, 3836) for meta-analyses with one OEL, compared to 1224 (inter-quartile range: 1141, 1713) for meta-analyses with more than one OEL.

We were able to replicate closely the results of the published meta-analyses for all 18 language inclusive/OEL meta-analyses. This analysis involved 211 randomized trials. Language restrictions in meta-analyses did not result in a significantly different estimate in the treatment effect compared with the inclusion of all languages (ROR = 0.98; 95% CI: 0.81–1.17; Table 4 and Fig. 1). This result did not change whether the meta-analyses included one (ROR = 0.96; 95% CI: 0.78–1.19) or more (ROR = 1.01; 95% CI: 0.72–1.43) OEL trials (Table 4). Similarly, the results were consistent whether OEL studies included a small number of participants (ROR = 1.00; 95% CI: 0.83–1.21) or larger numbers (ROR = 1.01; 95% CI: 0.80–1.26).

Language inclusive meta-analyses had narrower confidence intervals (average width = 0.79; 95% CI: 0.51–1.07) compared to language restricted meta-analyses (average width = 0.92; 95% CI: 0.53–1.32). This represents a statistically significant relative difference in precision of 16% (2P = 0.045).

To illustrate the effect of language of publication on an individual meta-analysis, we give the example of Poynard and colleagues [12] who investigated the benefits of smooth muscle relaxants for patients with irritable bowel syndrome using 25 trials of which 6 were OEL. The authors report a 27% (95% CI: 18%–36%) global improvement rate in muscle relaxation. Using the same trials we observed similar results in our replication reporting a 26% (95% CI: 16%–36%) global improvement. Language restricted (English only, n = 19) analysis resulted in a corresponding global improvement rate of 27% (95% CI: 14%–40%).

4. Discussion

Our primary result indicated only a 2% difference (ROR = 0.98), on average, between the treatment estimates with versus without explicit restrictions on the language of publication of the trials included. Combined with the narrow confidence intervals (0.81–1.17) and several sensitivity analyses, these results suggest that it is unlikely that any important clinical differences were missed. These results provide empirical evi-

Table 4
The effect of language of publication of randomized trials on the estimates of intervention effectiveness.

Type of analysis[a]	# of meta-analyses/ # randomized trials	Language effect (trials published in languages other than English compared to English language ones only) ROR (95% Confidence Interval)[b]	Estiamted heterogeneity between trials[c]
Language restrictive meta-analyses compared to language inclusive meta-analyses (Overall).	18/211	0.98 (0.81–1.17)	3.16 (χ^2 605.73 with 192 d.f.)
Language restricted meta-analyses compared to language inclusive meta-analyses (> 1 OEL).	5/85	1.01 (0.72–1.43)	2.59 (χ^2 204.81 with 79 d.f.)
Language restricted meta-analyses compared to language inclusive meta-analyses (= 1 OEL)	13/126	0.96 (0.78–1.18)	3.58 (χ^2 400.84 with 112 d.f.)
Analysis limited to randomized trials of sample size > 50.	14/128	1.00 (0.83–1.21)	2.56 (χ^2 289.84 with 113 d.f.)
Analysis limited to randomized trials of sample size > 100.	5/18	1.01 (0.80–1.26)	2.39 (χ^2 28.67 with 12 d.f.)

[a]The basic model: logit of events in a treatment arm = $\alpha + \beta i$ (trial ith indicator) + ε (treatment) + ϕj (treatment × jth meta-analysis) + φ (treatment × language), with × denoting an interaction. The effect of language on treatment effect estimates, the main parameter of interest, was tested using the treatment by language interaction (Appendix A).

[b]A Ratio of Odds Ratios less than 1 implies that language inclusive meta-analyses are associated with a larger treatment effect compared to language restricted meta-analyses.

[c]Mean deviance residual of the fitted models. Values larger than 1 indicates high heterogeneity between trials unexplained by the factors presented in the models.

dence regarding an important issue that meta-analysts and others have attempted to address previously [1,13,14] and add to the growing body of evidence regarding the appropriate conduct of meta-analysis [3,7,11,15–20].

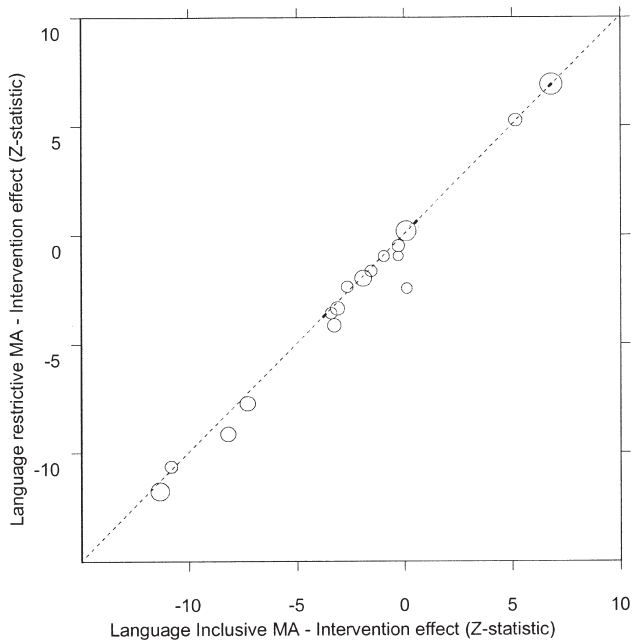

Fig. 1. Test of no intervention effect from language inclusive and language restricted meta-analyses (n = 18). The Z statistic (i.e., log odds ratio divided by its standard error) from language inclusive meta-analyses (X-axis) were plotted against its corresponding score from the same meta-analyses, with the OEL excluded (Y-axis). The sizes of plotting circles were inversely proportional to the variance of the language restricted meta-analyses estimates.

Meta-analysts who have limited their analyses to randomized trials published in English might view these results with some comfort. Whereas, analysts who have presumably expended additional resources to locate, retrieve and include OEL might question their efforts. In our view the decision as to whether or not to include OEL is not a simple one but multi-factorial.

A small trial (<50 participants) reported in any language is unlikely to alter the results or influence the precision of a meta-analysis including 1200 participants. Yet locating and obtaining the study, having it translated, particularly if it is an infrequently used language, will add cost and time to the process, and may influence the generalizability of the results of the meta-analysis. Alternatively, if there are several relevant OEL excluding them might be inappropriate due to their influence on the precision of the result.

We have recently reported the results of a meta-analysis [21] whereby the exclusion of OEL made a difference to the magnitude of the results. We examined the effects of pharmacological interventions, compared to placebo, on increasing maximum walking distance, in patients with intermittent claudication. Restricting the analysis to nine EL reports indicated that therapy extended the maximum walking distance by about 40 meters (WMD = 39.7: 95% CI: 11.3–68.1). However, three OEL trials, each of which had fewer than 35 patients, reported a substantially larger impact of therapy, increasing maximum walking distance to about 75 meters (WMD = 74.9: 95% CI: 10.6–139.1). Adding the OEL studies to the EL ones shows a marginal increase of approximately four meters in maximum walking distance (WMD = 43.2; 95% CI: 17.2–69.1). This finding suggests that OEL had little influence on the overall estimate of effi-

cacy primarily because the trials are small. However, their inclusion provided more precise result, due to the increased sample size. This is a single anecdotal report whereby results from OEL influenced the results. To what extent this result can be generalized is uncertain. This example highlights the point that the decision to include OEL is probably related to the clinical condition under investigation, the intervention, and the number of randomized trials. Additional research is needed to help clarify these issues.

In our sample, only 31% (79/251) of the meta-analyses were explicit about the language of publication eligibility. Clearer reporting of such methodology details will make it easier for replication, minimize bias in the process, and make readers more confident of the results. Perhaps this code of silence is because guides to help improve the quality of reports of meta-analysis have been inconsistent about the merits of reporting this information [22].

We were only able to identify 19 and use 18 meta-analyses out of 79 that actually included OEL. Even in those meta-analyses that do include OEL, the number of such studies included is low. In our sample, the majority (13/19) included only one OEL. This finding may reflect the sampling frame we used. Alternatively, it may indicate that OEL are less of an issue, at least for reports of EL meta-analyses, than previously thought.

Perhaps another way to address the language question is to identify meta-analyses that were language restricted from the outset. A comprehensive search could then be undertaken to identify randomized trials that could have been included in the restricted meta-analysis. Such meta-analyses could be replicated to include any OEL study. Such an investigation has been reported [1]. These authors identified 28 language restricted meta-analyses of which the statistical results of one meta-analysis [23], of selective decontamination of the digestive tract, whereby the inclusion of a German article (Odds Ratio = 0.68, 95% CI: 0.32–1.44) [24] would have changed the results from no statistical effect on mortality (Odds Ratio = 0.70, 95% CI: 0.45–1.09) to a statistically significant reduction in mortality (Odds Ratio = 0.67; 95% CI: 0.47–0.95). However, there has been some debate [25] about whether this German study was a randomized trial and should have been included in the meta-analysis in the first place.

The overall scientific quality of all meta-analyses was low. Similar results, using many of the same journals used here, have been reported elsewhere [26,27]. As such, we believe that our results are a representative sample of meta-analyses with respect to quality. We did not observe any statistically significant differences among the three groups of meta-analyses included here. The scientific quality in two of the groups included in our analysis scored within the "major flaws" category of the validated instrument we used. Perhaps initiatives, similar to those recently developed for randomized trials [28], will help improve the quality of reports of meta-analyses [29].

These results indicate that language inclusive meta-analyses, compared to language restricted meta-analyses, include more randomized trials, and have larger cumulative sample sizes. The net effect is to provide to a more precise result, as seen in narrowing, by approximately 16% on average, of the width of the confidence intervals of meta-analyses.

There are limitations to our study. We did not sample from the largest group of meta-analyses, namely, those in which the language selection criteria were not explicit. It is possible that this group is different from the three groups we did sample. However, the results reported by Jadad and McQuay [26], which included all four groups, are consistent with what we reported here.

Our analysis is based on a relatively small number of meta-analyses. This might be due to the limitations of our sampling frame, clinical topics and interventions. Our study included conventional interventions that might be of more interest in "developed" countries where high citation impact factor journals are published. Typically such journals publish English language reports only. This might explain why we found so few OEL trials. Perhaps the examination of complementary and alternative medicine interventions or a different sampling frame could provide different results than those we observed here [30,31]. We encourage others to replicate our study using different sampling frames, clinical areas and interventions.

Acknowledgments

We thank Alison Jones, Leah Le Page, and Michael Saginur for helping to coordinate the study. Alejandro R. Jadad, MD, is a National Health Research Scholar, Health Canada. This research was funded by the Medical Research Council of Canada, Grant # 3705.

References

[1] Grégoire G, Derderian F, LeLorier J. Selecting the language of the publications included in a meta-analysis: is there a tower of babel bias? J Clin Epidemiol 1995;48(1):159–63.

[2] Moher D, Fortin P, Jadad AR, Juni P, Kiassen T, Le Lorier J, Liberati A, Linde K, Penna A. Completeness of reporting of trials published in languages other than English: Implications for the conduct and reporting of systematic reviews. Lancet 1996;347:363–6.

[3] Jadad AR, Moore A, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, McQuay HJ. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996;17:1–12.

[4] Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. Lancet 1997;350:326–9.

[5] Moher D, Cook DJ, Jadad AR, Tugwell P, Moher M, Jones A, Pham B, Klassen TP. Assessing the quality of randomized controlled trials: implications for the conduct of meta-analyses. Health Technology Assessment 1999;3(12):1–98.

[6] Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, Moher D. Methodology and reports of systematic reviews and meta-analyses: A comparison of Cochrane reviews with articles published in paper-based journals. JAMA 1998;280:278–80.

[7] Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, Tugwell P, Klassen TP. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 1998;352:609–13.

[8] Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. J Clin Epidemiol 1991;44(11):1271–8.
[9] Fleiss JL. The Design and Analysis of Clinical Experiments. New York, NY, USA: John Wiley & Sons; 1986. p. 1–32.
[10] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
[11] Schulz KF, Chalmers I, Hayes R, Altman DG. Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995;273:408–12.
[12] Poynard T, Naveau S, Mory B, Chaput JC. Meta-analysis of smooth muscle relaxants in the treatment of irritable bowel syndrome. Aliment Pharmacol Ther 1994;8:499–510.
[13] Chalmers TC, Berrier J, Sacks HS, Levin H, Reitman D, Nagalingam R. Meta-analysis of clinical trials as a scientific discipline. II: Replicate variability and comparison of studies that agree and disagree. Stat Med 1987;6:733–44.
[14] 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;27:335–71.
[15] Dickersin K, Higgins K, Meinert CL. Identification of meta-analyses: The need for standard terminology. Control Clin Trials 1990;11: 52–66.
[16] Tramer M, Reynolds DJM, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. BMJ 1997;315: 635–40.
[17] Khan KS, Daya S, Collins JA, Walter S. Empirical evidence of bias in infertility research: Overestimation of treatment effect in crossover trials using pregnancy as the outcome measure. Fertil Steril 1996;65:939–945.
[18] Berlin JA. Does blinding of readers affect the results of meta-analyses? Lancet 1997;350:185–6.
[19] Barnes DE, Bero LA. Why review articles on the health effects of passive smoking reach different conclusions. JAMA 1998;279:1566–70.
[20] Simes RJ. Publication bias: the case for an international registry of clinical trials. J Clin Oncol 1986;4:1529–41.
[21] Moher D, Pham B, Ausejo M, Saenz A, Hood SC, Barber G. Pharmacological management of intermittent claudication: a meta-analysis of randomized trials. Drugs, 2000;59:1057–70.
[22] Shea B, Dubé C, Moher D. Assessing the quality of reports of meta-analyses: a systematic review of scales and checklists. *In*: Egger M, Altman D, Smith GD, editors. Systematic Reviews in Health Care, 2nd edition. London, UK: BMJ Publishing Group, In press.
[23] Vandenbroucke-Grauls CMJE, Vandenbroucke JP. Effect of selective decontamination of the digestive tract on respiratory tract infections and mortality in the intensive care unit. Lancet 1991;338:859–62.
[24] Thülig B, Hartenauer U, Diemer W, Lawin P, Fegeler W, Kehrel R, Ritzerfeld W. Selektive Florasuppression zur Infektionskontrolle in der Operativen Intensivmedizin. Anästh Intensivther Notfallmed 1989;24:345–54.
[25] Anonymous. Metaanalysis of randomized controlled trials of selective decontamination of the digestive tract. Selective Decontamination of the Digestive Tract Trialsts' Collaborative Group. BMJ 1993; 307:525–32.
[26] Jadad AR, McQuay HJ. Meta-analyses to evaluate analgesic interventions: a systematic qualitative review of their methodology. J Clin Epidemiol 1996;49:235–43.
[27] Shea B, Moher D, Pham B, Tugwell P. Assessing the quality of reporting of meta-analyses of randomized controlled trials. Poster presentation at the VII Cochrane Colloquium, Rome, Italy, 5–9 October 1999.
[28] Begg CB, Cho MK, Eastwood S, Horton R, Moher D, Olkin I, Rennie D, Schulz KF, Simel DL, Stroup DF. Improving the quality of reporting of randomized controlled trials: The CONSORT statement. JAMA 1996;276:637–9.
[29] Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup D, for the QUOROM group. Improving the quality of reporting of meta-analysis of randomized controlled trials: The QUOROM statement. Lancet 1999; 354:1896–1900.
[30] Vickers A, Goyal N, Harland R, Rees R. Do certain countries produce only positive results? A systematic review of controlled trials. Control Clin Trials 1998;19:159–66.
[31] Tang JL, Zhan SY, Ernst E. Review of randomised controlled trials of traditional Chinese medicine. BMJ 1999;319:160–1.

Appendix A

The estimate of language effect reported here (i.e., the ratio of odds-ratios of treatment effects estimated from EL and OEL trials, respectively) was derived from a logistic regression and verified in a conditional logistic regression to condition out the trial effect (i.e., treatment groups were matched on trial). The binary outcomes from the included trials were re-expressed as unwanted endpoints (e.g. mortality instead of survival outcomes), if necessary. The main model was specified as follows: logit of events in a treatment arm = $\alpha + \beta i$ (trial ith indicator) + ϵ (treatment) + ϕj (treatment $\times$ jth meta-analysis) + φ (treatment $\times$ language), with $\times$ denoting an interaction. The effect of language on treatment effect estimates, the main parameter of interest, was tested using the treatment by language interaction. Testing this parameter in the model is interpretable in terms of simpler analyses. Suppose we looked within a given meta-analysis. If we assessed the treatment effect within that meta-analysis separately from the EL trials and OEL trials, then tested the difference between the treatment effect estimates, that would be equivalent to the interaction test in a logistic regression model. The meta-logistic regression model simply averaged those within-meta-analysis interaction tests across all meta-analyses included in the model.

The language effect estimate was unchanged regardless of the variations in the above model specification. In particular, the main effect of language or meta-analysis did not explain anything beyond the trial main factor and did not change the language effect estimate.

The inclusion of many indicator variables for the trial factor in this logistic regression model was a cause for concern. As a sensitivity analysis, we performed a conditional logistic regression matching treatment by trial. Note that when conditioning on trial, the main effect of language was not meaningful, or at least its interpretation was problematic. This was however parallel to similar considerations in matched case-control studies. For example, one can match on a factor such as age, and still estimate the age and exposure interactions, even though the main effect of age in such a study would not have a meaningful clinical interpretation. For simplicity, we elected to report our findings from the logistic regression as the inclusion of indicators for trials did not affect the language effect estimates.

Appendix B

Language Inclusive - OEL

1. a'Rogvi-Hansen B. Glycerol treatment for acute ischaemic stroke. Warlow C, Van Gijn J, Sandercock P (eds.), Cochrane Database of Systematic Reviews Issue 3, 1996. The Cochrane Library.
2. Anonymous. Meta-analysis of randomised controlled trials of selective

decontamination of the digestive tract. Selective Decontamination of the Digestive Tract Trialist's Collaborative Group. BMJ 1993;307(6903): 525–32.
3. Covey LS, Glassman AH. A meta-analysis of double-blind placebo-controlled trials of clonidine for smoking cessation. Br J Addict 1991; 86(8):991–8.
4. Fine MJ, Smith MA, Carson CA, et al. Efficacy of pneumococcal vaccination in adults. A meta-analysis of randomized controlled trials. Arch Intern Med 1994;154(23):2666–77.
5. Glowacki LS, Smaill FM. Use of immune globulin to prevent symptomatic cytomegalovirus disease in transplant recipients–a meta-analysis. Clin Transplant 1994;8(1):10–18.
6. Gotzsche PC. Patient's preference in indomethacin trials: An overview. Lancet 1989;1(8629):88–91.
7. Gregory WM, Richards MA, Malpas JS. Combination chemotherapy versus melphalan and prednisolone in the treatment of multiple myeloma: An overview of published trials. J Clin Oncol 1992;10(2):344–52.
8. Halpern S, Preston R. Postdural puncture headache and spinal needle design. Metaanalyses. Anesthesiology 1994;81(6):1376–83.
9. Hofmeyr GJ. External cephalic version at term. Enkin MW, Keirse MJ, Renfrew MJ, Neilson JP (eds.), Cochrane Database of Systematic Reviews, Issue 3, 1996. The Cochrane Library.
10. Hofmeyr GJ. Cephalic version by postural management. Enkin MW, Keirse MJ, Renfrew MJ, Neilson JP (eds.), Cochrane Database of Systematic Reviews, Issue 3, 1996. The Cochrane Collaboration.
11. Leizorovcz A, Simonneau G, Decousus H, Boissel JP. Comparison of efficacy and safety of low molecular weight heparins and unfractionated heparin in initial treatment of deep venous thrombosis: A meta-analysis. BMJ 1994;309(6950):299–304.
12. Marino P, Pampallona S, Preatoni A, Cantoni A, Invernizzi F. Chemotherapy vs supportive care in advanced non-small-cell lung cancer. Results of a meta-analysis of the literature. Chest 1994;106(3):861–5.
13. Meijer WS, Schmitz PI, Jeekel J. Meta-analysis of randomized, controlled clinical trials of antibiotic prophylaxis in biliary tract surgery. Br J Surg 1990;77(3):283–90.
14. Pace F, Maconi G, Molteni P, Minguzzi M, Bianchi Porro G. Meta-analysis of the effect of placebo on the outcome of medically treated reflux esophagitis. Scand J Gastroenterol 1995;30(2):101–5
15. Pouleur H, Buyse M. Effects of dipyridamole in combination with anticoagulant therapy on survival and thromboembolic events in patients with prosthetic heart valves. A meta-analysis of the randomized trials. J Thorac Cardiovasc Surg 1995;110(2):463–72.
16. Poynard T, Naveau S, Mory B, Chaput JC. Meta-analysis of smooth muscle relaxants in the treatment of irritable bowel syndrome. Aliment Pharmacol Ther 1994;8(5):499–510.
17. Vandekerckhove P, Lilford R, Vail A, Hughes E. The medical treatment of idiopathic oligo/asthenospermia: Bromocriptine versus placebo or no treatment. Lilford R, Hughes E, Vandekerckhove P (eds.), Cochrane Database of Systematic Reviews, Issue 2, 1997. The Cochrane Collaboration.
18. Vanderkerckhove P, Lilford R, Vail A, Hughes E. Androgens versus placebo or no treatment for idiopathic oligo/asthenospermia. Cochrane Database of Systematic Reviews, Issue 2, 1999. The Cochrane Collaboration.
19. Wilson AP, Shrimpton S, Jaderberg M. A meta-analysis of the use of amoxycillin-clavulanic acid in surgical prophylaxis. J Hosp Infect 1992;22(Suppl A):9–21.

Language Inclusive - EL

1. Andrews TC, Reimold SC, Berlin JA, Antman EM. Prevention of supraventricular arrhythmias after coronary artery bypass surgery. A meta-analysis of randomized control trials. Circulation 1991;84(5 Suppl):III 236–44.
2. Barker RG II. Efficacy of prophylactic antibiotics for craniotomy: A meta-analysis. Neurosurgery 1994;35(3):484–90; Discussion 491–2.
3. Cohard M, Poynard T, Mathurin P, Zarski JP. Prednisone-interferon combination in the treatment of chronic hepatitis B: direct and indirect metanalysis. Hepatology 1994;20(6):1390–8.
4. Colditz GA, Brewer TF, Berkey CS, Wilson ME, Burdick E, Fineberg HV, Mosteller F. Efficacy of BCG vaccine in the prevention of tuberculosis. Meta-analysis of the published literature. JAMA 1994;271(9): 698–702.
5. Colditz GA, Berkey CS, Mosteller F, Brewer TF, Wilson ME, Burdick E, Fineberg HV. The efficacy of bacillus Calmette-Guerin vaccination of newborns and infants in the prevention of tuberculosis: meta-analyses of the published literature. Pediatrics 1995;96(1Pt1):29–35.
6. Coplen SE, Antman EM, Berlin JA, Hewitt P, Chalmers TC. Efficacy and safety of quinidine therapy for maintenance of sinus rhythm after cardioversion. A meta-analysis of randomized control trials. Circulation 1990;82(4):1106–16.
7. Counsell C, Salinas R, Warlow C, Naylor R. The role of patch angioplasty in carotid endartechtomy: A systematic review of the randomized trials comparing patching with primary closure. In: Warlow C, Van Gijn J, Sandercock P (editors), Cochrane Database of Systematic Reviews, Issue 3, 1996. The Cochrane Collaboration.
8. Cummings P, Del Beccaro MA. Antibiotics to prevent infection of simple wounds: A meta-analysis of randomized studies. Am J Emerg Med 1995;13(4):396–400.
9. Fardy JM, Laupacis A. A meta-analysis of prophylactic endoscopic sclerotherapy for esophageal varices. Am J Gastroenterol 1994; 89(11):1938–48.
10. Graves P. Human malaria vaccines. In: Feng C, Garner P, Gelband H, Salinas R (editors), Cochrane Database of Systematic Reviews, Issue 3, 1996. The Cochrane Collaboration.
11. Hofmeyr GJ. Abdominal decompression for suspected fetal compromise/pre-eclampsia. In: Enkin MW, Keirse MJ, Renfrew MJ, Neilson JP (editors), Cochrane Database of Systematic Reviews, Issue 3, 1996. The Cochrane Collaboration.
12. Kramer MS. Nutritional advice in pregnancy. In: Enkin MW, Keirse MJ, Renfrew MJ, Neilson JP (editors), Cochrane Database of Systematic Reviews, Issue 3, 1996. The Cochrane Collaboration.
13. Kramer MS. Maternal antigen avoidance during lactation in women at high risk for atopic offspring. In: Enkin MW, Keirse MJ, Renfrew MJ, Neilson JP (editors), Cochrane Database of Systematic Reviews, Issue 3, 1996. The Cochrane Collaboration.
14. Lancaster T, Silagy C, Gray S. Primary care management of acute herpes zoster: systematic review of evidence from randomized controlled trials. Br J Gen Pract 1995;45(390):39–45.
15. Langhorne P, Williams BO, Gilchrist W, Howie K. Do stroke units save lives? Lancet 1993;342(8868):395–8.
16. Leizorovicz A, Haugh MC, Chapuis FR, Samama MM, Boissel JP. Low molecular weight heparin in prevention of perioperative thrombosis. BMJ 1992;305(6859):913–20.
17. Lugo-Miro VI, Green M, Mazur L. Comparison of different metronidazole therapeutic regimes for bacterial vaginosis. A meta-analysis. JAMA 1992;268(1):92–5.
18. Mari JJ, Streiner DL. An overview of family interventions and relapse on schizophrenia: Meta-analysis of research findings. Psychol Med 1994;14(3):565–78.
19. Pharoah FM, Mari JJ, Streiner D. Family intervention for schizophrenia. Cochrane Database of Systematic Reviews, Issue 3, 1996. The Cochrane Collaboration.
20. Renfrew MJ, Langhorne P. Early initiation of breastfeeding and its effect on duration. In: Enkin MW, Keirse MJ, Renfrew MJ, Neilson JP (editors), Cochrane Database of Systematic Reviews, Issue 3, 1996. The Cochrane Collaboration.
21. Spina GP, Henderson JM, Rikkers LF, *et al.* Distal spleno-renal shunt versus endoscopic sclerotherapy in the prevention of variceal rebleeding. A meta-analysis of 4 randomized clinical trials. J Hepatol 1992; 16(3):338–45.
22. Tine F, Magrin S, Craxi A, Pagliaro L. Interferon for non-A, non-B chronic hepatitis. A meta-analysis of randomised clinical trials. J Hepatol 1991;13(2):192–9.

Language Restricted

1. Abramson MJ, Puy RM, Weiner JM. Is allergen immunotherapy effective in asthma? A meta-analysis of randomized controlled trials. Am J Respir Crit Care Med 1995;151(4):969–74.
2. Appel LJ, Miller ER III, Seidler AJ, Whelton PK. Does supplementation of diet with 'fish oil' reduce blood pressure? A meta-analysis of controlled clinical trials. Arch Intern Med 1993;153(12):1429–38.
3. Avgerinos A, Armonis A, Raptis S. Somatostatin or octreotide versus endoscopic sclerotherapy in acute variceal haemorrhage: A meta-analysis study. J Hepatol 1995;22(2):247–8.
4. Browman GP. Evidence-based recommendations against neoadjuvant chemotherapy for routine management of patients with squamous cell head and neck cancer. Cancer Invest 1994;12(6):662–70.
5. Cappelleri JC, Fiore LD, Brophy MT, Deykin D, Lau J. Efficacy and safety of combined anticoagulant and antiplatelet therapy versus anticoagulant monotherapy after mechanical heart-valve replacement: A metaanalysis. Am Heart J 1995;130(3Pt1):547–52.
6. Cummings P, Psaty BM. The association between cholesterol and death from injury. Ann Intern Med 1994;120(10):848–55.
7. Eisenberg E, Berkey CS, Carr DB, Mosteller F, Chalmers TC. Efficacy and safety of nonsteroidal antiinflammatory drugs for cancer pain: A meta-analysis. J Clin Oncol 1994;12(12):2756–65.
8. Fraser EJ, Grimes DA, Schulz KF. Immunization as therapy for recurrent spontaneous abortion: A review and meta-analysis. Obstet Gynecol 1993;82(5):854–9.
9. Fremes SE, Wong BI, Lee E, *et al.* Metaanalysis of prophylactic drug treatment in the prevention of postoperative bleeding. Ann Thorac Surg 1994;58(6):1580–8.
10. Hansen JF. Review of postinfarct treatment with verapamil: Combined experience of early and late intervention studies with verapamil in patients with acute myocardial infarction. Danish Study Group on Verapamil in Myocardial Infarction. Cardiovasc Drugs Ther 1994;8 Suppl 3:543–7.
11. Hauth JC, Goldenberg RL, Parker CR, Jr., Cutter GR, Cliver SP. Low-dose aspirin: Lack of association with an increase in abruptio placentae or perinatal mortality. Obstet Gynecol 1995;85(6):1055–8.
12. Hazell P, O'Connell D, Heathcote D, Robertson J, Henry D. Efficacy of tricyclic drugs in treating child and adolescent depression: A meta-analysis. BMJ 1995;310(6984):897–901.
13. Hillegass WB, Ohman EM, Leimberger JD, Califf RM. A meta-analysis of randomized trials of calcium antagonists to reduce restenosis after coronary angioplasty. Am J Cardiol 1994;73(12):835–9.
14. Hooker KD, DiPiro JT, Wynn JJ. Aminoglycoside combinations versus beta-lactams alone for penetrating abdominal trauma: A meta-analysis. J Trauma 1991;31(8):1155–60.
15. Hricik DE, O'Toole MA, Schulak JA, Herson J. Steroid-free immunosuppression in cyclosporine-treated renal transplant recipients: A meta-analysis. J Am Soc Nephrol 1993;4(6):1300–5.
16. Imperiale TF, Goldfarb S, Berns JS. Are cytotoxic agents beneficial in idiopathic membranous nephropathy? A meta-analysis of the controlled trials. J Am Socl Nephrol 1995;5(8):1553–8.
17. Kreter B, Woods M. Antibiotic prophylaxis for cardiothoracic operations. Meta-analysis of thirty years of clinical trials. J Thorac Cardiovasc Surg 1992;104(3):590–9.
18. Labrecque M, Dostaler LP, Rousselle R, Nguyen T, Poirier S. Efficacy of nonsteroidal anti-inflammatory drugs in the treatment of acute renal colic. A meta-analysis. Arch Intern Med 1994;154(12):1381–7.
19. Lensing AW, Prins MH, Davidson BL, Hirsh J. Treatment of deep venous thrombosis with low-molecular-weight heparins. A meta-analysis. Arch Intern Med 1995;155(6):601–7.
20. Macharia WM, Leon G, Rowe BH, Stephenson BJ, Haynes RB. An overview of interventions to improve compliance with appointment keeping for medical services. JAMA 1992;267(13):1813–17.
21. May GR, Sutherland LR, Shaffer EA. Efficacy of bile acid therapy for gallstone dissolution: A meta-analysis of randomized trials. Aliment Pharmacol Ther 1993;7(2):139–48.
22. Meunier F, Paesmans M, Autier P. Value of antifungal prophylaxis with antifungal drugs against oropharyngeal candidiasis in cancer patients. Eur J Cancer Part.B, Oral Oncology 1994;30B(3):196–9.
23. Midgette AS, O'Connor GT, Baron JA, Bell J. Effect of intravenous streptokinase on early mortality in patients with suspected acute myocardial infarction. A meta-analysis by anatomic location of infarction. Ann Intern Med 1990;113(12):961–8.
24. Midgette AS, Wong JB, Beshansky JR, *et al.* Cost-effectiveness of streptokinase for acute myocardial infraction: A combined meta-analysis and decision analysis of the effects of infarct location and of likelihood of infarction. Med Decis Making 1994;14(2):108–17.
25. Morganroth J, Goin JE. Quinidine-related mortality in the short-to-medium-term treatment of ventricular arrhythmias. A meta-analysis. Circulation 1991;84(5):1977–83.
26. O'Brien BJ, Anderson DR, Goeree R. Cost-effectiveness of enoxaparin versus warfarin prophylaxis against deep-vein thrombosis after total hip replacement. Can Med Assoc J 1994;150(7):1083–90.
27. O'Connor GT, Malenka DJ, Olmstead EM, Johnson PS, Hennekens CH. A meta-analysis of randomized trials of fish oil in prevention of restenosis following coronary angioplasty. Am J Prev Med 1992;8(3): 186–92.
28. Pichichero ME, Margolis PA. A comparison of cephalosporins and penicillins in the treatment of group A beta-hemolytic streptococcal pharyngitis: A meta-analysis supporting the concept of microbial copathogenicity. Pediatr Infect Dis J 1991;10(4):275–81.
29. Pyorala S, Huttunen NP, Uhari M. A review and meta-analysis of hormonal treatment of cryptorchidism. J Clin Endocrinol Metab 1995; 80(9):2795–9.
30. Rossetti L, Marchetti I, Orzalesi N, *et al.* Randomized clinical trials on medical treatment of glaucoma. Are they appropriate to guide clinical practice? Arch Ophthalmol 1993;111(1):96–103.
31. Rowe BH, Keller JL, Oxman AD. Effectiveness of steroid therapy in acute exacerbations of asthma: A meta-analysis. Am J Emerg Med 1992;10(4):301–10.
32. Sacks HS, Chalmers TC, Blum AL, Berrier J, Pagano D. Endoscopic hemostatis. An effective therapy for bleeding peptic ulcers. JAMA 1990;264(4):494–9.
33. Thomas JA, McIntosh JM. Are incentive spirometry, intermittent positive pressure breathing, and deep breathing exercises effective in the prevention of postoperative pulmonary complications after upper abdominal surgery? A systematic overview and meta-analysis. Phys Ther 1994;74(1):3–10; Discussion 10–16.
34. Van Ruiswyk J, Byrd JC. Efficacy of prophylactic sclerotherapy for prevention of a first variceal hemorrhage. Gastroenterology 1992; 102(2):587–97.
35. Wang PH, Lau J, Chalmers TC. Meta-analysis of effects of intensive blood-glucose control on late complications of type I diabetes. Lancet 1993;341(8856):1306–9.
36. Yurkowski PJ, Plaisance KI. Prevention of auditory sequelae in pediatric bacterial meningitis: a meta-analysis. Pharmacotherapy 1993; 13(5):494–9.
37. Zhang WY, Li Wan Po A. The effectiveness of topically applied capsaicin. A meta-analysis. Eur J Clin Pharmacol 1994;46(6):517–22.

Language Restricted - Excluded from the analysis

38. Moreland J, Thomson MA. Efficacy of electromyographic biofeedback compared with conventional physical therapy for upper-extremity function in patients following stroke: A research overview and meta-analysis. Phys Ther 1994;74(6):534–543; Discussion 544–547.