

# Reporting of results from network meta-analyses: methodological systematic review

Aïda Bafeta,<sup>1</sup> Ludovic Trinquart,<sup>1,2,3,4</sup> Raphaële Seror,<sup>1,3</sup> Philippe Ravaud<sup>1,2,3,4</sup>

<sup>1</sup>Assistance Publique-Hôpitaux de Paris, Centre d'Epidémiologie Clinique, INSERM U1153, Hôpital Hôtel-Dieu, 75004 Paris, France

<sup>2</sup>Université Paris Descartes-Sorbonne Paris Cité, Paris

<sup>3</sup>French Cochrane Centre, Paris

<sup>4</sup>Department of Epidemiology, Mailman School of Public Health, Columbia University, New York City, USA

Correspondence to: A Bafeta  
aida.bafeta@htd.aphp.fr

Cite this as: *BMJ* 2014;348:g1741  
doi: 10.1136/bmj.g1741

This is a summary of a paper that was published on *bmj.com* as *BMJ* 2014;348:g1741

## STUDY QUESTION

How are the results of network meta-analyses reported?

## SUMMARY ANSWER

The results of network meta-analyses are heterogeneously reported, possibly reflecting the lack of a general consensus on what should be reported in a network meta-analysis.

## WHAT IS KNOWN AND WHAT THIS PAPER ADDS

Inadequate reporting of results of network meta-analyses may affect their interpretation and mislead clinical researchers. This review identifies a clear need to extend reporting guidelines to network meta-analyses to improve their reporting.

## Selection criteria for studies

We searched the Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effects, Medline, and Embase from inception to 12 July 2012 for network meta-analyses comparing the clinical efficacy of three or more interventions on the basis of randomised controlled trials, excluding those with an open loop network of three interventions.

## Primary outcome(s)

We assessed the reporting of the network and results. A composite outcome included the description of the net-

work (number of interventions, direct comparisons, and randomised controlled trials and patients per comparison) and the reporting of effect sizes derived from direct evidence, indirect evidence, and the network meta-analysis.

## Main results and the role of chance

The network and its geometry (network graph) were not reported in 100 (83%) of the 121 articles included. The effect sizes derived from direct and indirect evidence and the network meta-analysis were not reported in 48 (40%), 108 (89%), and 43 (36%) articles, respectively. In 52 reports that ranked interventions, 43 did not report the uncertainty in ranking. Overall, 119 (98%) reports of network meta-analyses did not give a description of the network or effect sizes from direct evidence, indirect evidence, and the network meta-analysis. This finding did not differ by journal type or funding source.

## Bias, confounding, and other reasons for caution

Our study has some limitations, as we searched for only network meta-analysis reports published in journals and did not search for reports of health technology assessments. Network meta-analyses are increasingly being used to support health technology assessments, and the reporting of results may differ in health technology assessment reports and journal articles. Finally, we built a composite outcome based on the available recommendations from four sources as a possible core set of information that we would like to see in any report of a network meta-analysis. However, we acknowledge that consensus is lacking on which items should be required in reports of network meta-analyses. This is an area of ongoing debate.

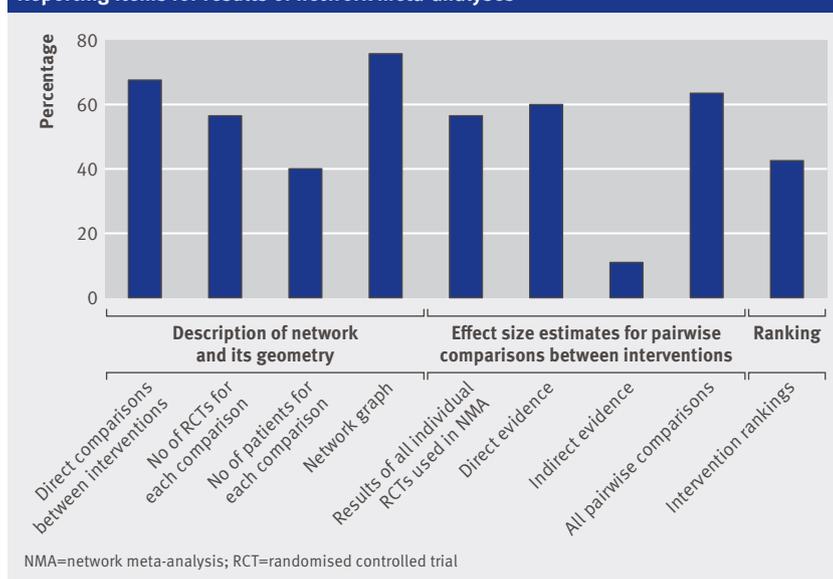
## Study funding/potential competing interests

AB was funded by an academic grant for doctoral students from Pierre et Marie Curie University, France. Our team is supported by an academic grant for the programme Equipe espoir de la Recherche, from the Fondation pour la Recherche Médicale, France.

## bmj.com

- Research: The effects of excluding treatments from network meta-analyses: survey (*BMJ* 2013;347:f5195)
- Research: Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review (*BMJ* 2013;347:f3675)
- Research: Can trial quality be reliably assessed from published reports of cancer trials: evaluation of risk of bias assessments in systematic reviews (*BMJ* 2013;346:f1798)

## Reporting items for results of network meta-analyses



# Exclusion of deep vein thrombosis using the Wells rule in clinically important subgroups: individual patient data meta-analysis

G J Geersing,<sup>1</sup> N P A Zuithoff,<sup>1</sup> C Kearon,<sup>2</sup> D R Anderson,<sup>3</sup> A J ten Cate-Hoek,<sup>4</sup> J L Elf,<sup>5</sup> S M Bates,<sup>2</sup> A W Hoes,<sup>1</sup> R A Kraaijenhagen,<sup>6</sup> R Oudega,<sup>1</sup> R E G Schutgens,<sup>7</sup> S M Stevens,<sup>8</sup> S C Woller,<sup>8</sup> P S Wells,<sup>9</sup> K G M Moons<sup>1</sup>

## EDITORIAL by Iorio and Douketis

<sup>1</sup>Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Universiteitsweg 100, 3584 CG Utrecht, Netherlands

<sup>2</sup>Division of Haematology and Thromboembolism, Department of Medicine, McMaster University Hamilton, Hamilton, Canada

<sup>3</sup>Division of Haematology, Department of Medicine, Dalhousie University, Halifax, Canada

<sup>4</sup>Department of Internal Medicine, Maastricht University Medical Center, Maastricht, Netherlands

<sup>5</sup>Vascular Center, Skane University Hospital, Malmö, Sweden

<sup>6</sup>Department of Medicine, Academic Medical Center Amsterdam, Netherlands

<sup>7</sup>Van Creveld Clinic, University Medical Center Utrecht, Utrecht, Netherlands

<sup>8</sup>Thrombosis Clinic, Intermountain Medical Center, Murray, UT, USA

<sup>9</sup>Department of Medicine Ottawa Hospital, University of Ottawa, Ottawa, Canada

Correspondence to: G J Geersing [g.j.geersing@umcutrecht.nl](mailto:g.j.geersing@umcutrecht.nl)

Cite this as: *BMJ* 2014;348:g1340  
doi: 10.1136/bmj.g1340

This is a summary of a paper that was published on [bmj.com](http://bmj.com) as *BMJ* 2014;348:g1340

## [bmj.com/multimedia](http://bmj.com/multimedia)

See the authors talk about their research

## STUDY QUESTION

How accurate is the Wells rule combined with D-dimer testing for excluding deep vein thrombosis, and is this accuracy maintained in different subgroups of patients?

## SUMMARY ANSWER

Deep vein thrombosis can be excluded in most, but not all, subgroups of patients with an unlikely score on the Wells rule combined with a negative D-dimer test result.

## WHAT IS KNOWN AND WHAT THIS PAPER ADDS

The validity of the Wells rule is questioned in various clinically important subgroups. This study found that the Wells rule combined with D-dimer testing is safe and efficient to rule out deep vein thrombosis in most subgroups, but not in patients with cancer, or in patients with suspected recurrent disease. In the second group, one extra point should be added to the rule to enable a safe exclusion of deep vein thrombosis.

## Selection criteria for studies

This study included data at individual patient level from 13 diagnostic accuracy studies performed in consecutive outpatients (n=10 002) with suspected deep vein thrombosis. We created one dataset by pooling the individual patient data, totalling 1864 confirmed cases of deep vein thrombosis (using an a priori defined acceptable reference standard).

## Primary outcome

We used multilevel logistic regression analysis to estimate

the mean predicted probabilities of deep vein thrombosis for different scores on the Wells rule, both with and without D-dimer testing.

## Main results and role of chance

Increasing scores on the Wells rule were associated with a higher probability of deep vein thrombosis being present. An unlikely score on the Wells rule ( $\leq 1$ ) combined with a negative D-dimer test result was associated with an extremely low probability of deep vein thrombosis (the failure rate, 1.2%, 95% confidence interval 0.7% to 1.8%). This combination occurred in 29% (95% confidence interval 20% to 40%) of patients (the efficiency rate). These findings were consistent in various subgroups. For patients with cancer, however, the combination of an unlikely score on the Wells rule and a negative D-dimer test result occurred in only 9% of patients and was associated with a 2.2% probability of deep vein thrombosis. In patients with suspected recurrent events, only the modified Wells rule (adding one point for the previous event) is safe.

## Bias, confounding, and other reasons for caution

The prevalence of deep vein thrombosis varied largely over the 13 included studies (range 5-39%). This variability in prevalence explained the observed heterogeneity around our model estimates.

## Study funding/potential competing interests

This study received no funding. We have no competing interests.

Estimated failure rates and efficiency of using Wells rule score  $\leq 1$  and negative D-dimer test result to exclude deep vein thrombosis in two important subgroups. Values are percentages (95% confidence intervals)

Accuracy measures	Overall	Patients with cancer	Patients with previous deep vein thrombosis	
			Original rule	Adding 1 point for previous event
Failure rate*	1.2 (0.7 to 1.8)	2.2 (0.5 to 8.6)	2.5 (1.2 to 5.4)	1.0 (0.6 to 1.6)
Efficiency†	29 (20 to 40)	9 (6 to 15)	30 (20 to 42)	27 (19 to 37)

\*Mean predicted probability of deep vein thrombosis being present.

†Proportion of patients in whom deep vein thrombosis can possibly be excluded.

# Ibuprofen, paracetamol, and steam for patients with respiratory tract infections in primary care: pragmatic randomised factorial trial

Paul Little, Michael Moore, Joanne Kelly, Ian Williamson, Geraldine Leydon, Lisa McDermott, Mark Mullee, Beth Stuart; on behalf of the PIPS investigators

University of Southampton,  
Aldermoor Health Centre,  
Southampton SO16 5ST, UK  
Correspondence to: P Little  
p.little@soton.ac.uk

Cite this as: *BMJ* 2013;347:f6041  
doi: 10.1136/bmj.f6041

This is a summary of a paper that  
was published on bmj.com as *BMJ*  
2013;347:f6041

## bmj.com

Research: Delayed antibiotic  
prescribing strategies for  
respiratory tract infections  
in primary care: pragmatic,  
factorial, randomised  
controlled trial  
(*BMJ* 2014;348:g1606)

Research: Duration of  
symptoms of respiratory  
tract infections in children:  
systematic review  
(*BMJ* 2013;347:f7027)

Editorial: Prognosis of  
respiratory tract infections in  
primary care  
(*BMJ* 2013;347:f7185)

Clinical Review: Burns caused  
by steam inhalation  
for respiratory tract infections  
in children  
(*BMJ* 2004;328:757)

## STUDY QUESTION

Does symptom management in patients with acute respiratory tract infections improve when advice is given to use ibuprofen alone or ibuprofen and paracetamol compared with paracetamol alone; to take regular doses rather than as required; and to use steam inhalation compared with no inhalation?

## SUMMARY ANSWER

None of these strategies significantly improved symptom control. Advice to use ibuprofen might help more among those with chest infections and in children but is also associated with more consultations with new or unresolved symptoms.

## WHAT IS KNOWN AND WHAT THIS PAPER ADDS

Ibuprofen or the combination of ibuprofen with paracetamol might help fever control marginally more than paracetamol alone, but there is little evidence about the management of symptoms overall and concerns about the applicability of the evidence to primary care

## Design

889 patients were randomised with computer generated random numbers in an open pragmatic parallel group factorial trial by opening prepared sealed numbered envelopes to components of advice or comparator advice: advice on analgesia (take paracetamol, ibuprofen, or both), dosing of analgesia (take as required v regularly), and steam inhalation (no inhalation v steam inhalation).

## Participants

Patients aged  $\geq 3$  with acute respiratory tract infections (mean age 30).

## Primary outcome

Symptom severity on days two to four (primary outcome), temperature, antibiotic use, consultations.

## Main results and the role of chance

Advice to use regular analgesics or to inhale with steam did not significantly change outcomes. Compared with par-

acetamol, symptom severity was also little different with ibuprofen alone (adjusted difference 0.04, 95% confidence interval  $-0.11$  to  $0.19$ ) or the combination ( $0.11$ ,  $-0.04$  to  $0.26$ ). There was no evidence for selective benefit for ibuprofen among most subgroups defined before analysis (presence of otalgia; previous duration of symptoms; temperature  $>37.5^{\circ}\text{C}$ ; severe symptoms), but evidence of benefit for chest infections (ibuprofen  $-0.40$ ,  $-0.78$  to  $-0.01$ ; combination  $-0.47$ ,  $-0.84$  to  $-0.10$ ), equivalent to almost one in two symptoms rated a slight rather than moderately bad problem. Children might also benefit more (ibuprofen  $-0.47$ ,  $-0.76$  to  $-0.18$ ; combination  $-0.04$ ,  $-0.31$  to  $0.23$ ).

## Harms

Reconsultations with new/unresolved symptoms or complications were documented in 12% of those advised to take paracetamol, 20% of those advised to take ibuprofen (adjusted risk ratio 1.67, 1.12 to 2.38), and 17% of those advised to take the combination (1.49, 0.98 to 2.18). Mild thermal injury with steam was documented in four patients (2%) who returned full diaries, but no consultations with scalding were documented.

## Bias, confounding, and other reasons for caution

To avoid errors in intervention delivery in a complex factorial trial we used sealed numbered envelopes to execute randomisation. Potential confounders were mostly well distributed between groups, there was high retention (84%), and an intention to treat analysis was done without imputation. The trial was not powered for potentially important outcomes such as complications.

## Generalisability

The limited invitation to recruit potential participants reflects the realities of doing trials in busy acute clinics at the busiest times of year. However, patients came from a wide range of settings, and the sample, if anything, had slightly more patients with severe symptoms compared with large observational cohorts from primary care.

**Trial registration number** ISRCTN 38551726.

Effectiveness of advice strategies for analgesia on symptom control in patients with acute respiratory tract infections. Figures are crude mean symptom severity scores two to four days after seeing doctor

	Ibuprofen			Paracetamol and ibuprofen	
	Paracetamol crude mean (SD)	Crude mean (SD)	Adjusted* difference (95% CI), P value	Crude mean (SD)	Adjusted difference (95% CI), P value
Whole cohort (n=743)	1.67 (0.82)	1.71 (0.96)	0.04 ( $-0.11$ to $0.19$ ), 0.59	1.78 (0.94)	0.11 ( $-0.04$ to $0.26$ ), 0.14
LRTI (n=113)	2.14 (0.99)	1.70 (0.94)	$-0.40$ ( $-0.78$ to $-0.01$ ), 0.04	1.74 (0.84)	$-0.47$ ( $-0.84$ to $-0.10$ ), 0.01
Non-LRTI (n=630)	1.60 (0.77)	1.73 (0.97)	0.11 ( $-0.05$ to $0.27$ ), 0.19	1.77 (0.96)	0.20 ( $0.03$ to $0.36$ ), 0.03
Age $\leq 16$ (n=200)	1.74 (0.80)	1.20 (0.79)	$-0.47$ ( $-0.76$ to $-0.18$ ), $<0.01$	1.61 (0.90)	$-0.04$ ( $-0.31$ to $0.23$ ), 0.77
Age $\geq 17$ (n=543)	1.65 (0.83)	1.88 (0.96)	0.20 ( $0.03$ to $0.38$ ), 0.02	1.85 (0.95)	0.16 ( $-0.02$ to $0.34$ ), 0.08

\*Adjusted for baseline symptom severity, dosing, steam, antibiotic prescribing, and smoking (as smoking significantly predicted symptom severity). Outcome is severity of symptoms so lower symptom severity (or negative change) is better. Denominators vary because of missing data.

bmj.com Read the latest on respiratory medicine on our dedicated portal

## Delayed antibiotic prescribing strategies for respiratory tract infections in primary care: pragmatic, factorial, randomised controlled trial

Paul Little, Michael Moore, Jo Kelly, Ian Williamson, Geraldine Leydon, Lisa McDermott, Mark Mullee, Beth Stuart; on behalf of the PIPS Investigators

University of Southampton,  
Aldermoor Health Centre,  
Southampton SO16 5ST, UK  
Correspondence to: P Little  
p.little@soton.ac.uk

Cite this as: *BMJ* 2014;348:g1606  
doi: 10.1136/bmj.g1606

This is a summary of a paper that  
was published on bmj.com as *BMJ*  
2014;348:g1606

**Trial registration number**  
ISRCTN38551726.

### STUDY QUESTION

How effective are different strategies of delayed antibiotic prescription in symptom control for acute respiratory tract infections?

### SUMMARY ANSWER

Symptom control did not differ by strategy; either no or delayed prescription resulted in fewer than 40% of patients using antibiotics compared with over 90% with an immediate prescription.

### WHAT IS KNOWN AND WHAT THIS PAPER ADDS

Systematic reviews of delayed antibiotics suggest worse symptom control of respiratory tract infections than immediate antibiotics and higher antibiotic use than no prescription. This study demonstrates good symptom control and similar outcomes for the different delayed prescription methods.

### Design

Of 889 participants, 333 were prescribed immediate antibiotics, and the remainder were randomised in an open, pragmatic, parallel group, factorial trial to different approaches of delayed prescribing. Approaches were recontact for a prescription, post-dated prescription, collection of the prescription, and a patient led approach (the patient was given the prescription). During the trial, a strategy of no antibiotic prescription was added as another randomised comparison. Advice to use analgesia or steam was controlled in a factorial design.

### Participants

Patients aged 3 years or over with acute respiratory tract infections.

### Main outcomes

Symptom severity on days 2-4 (primary outcome), and antibiotic use.

### Main results and the role of chance

On a scale of 0 to 6 with increasing severity, symptom severity showed minimal differences between groups (no prescription 1.62, recontact 1.60, post-date 1.82, collection 1.68, patient led 1.75; likelihood ratio test  $\chi^2$  2.61,  $P=0.625$ ), as did median duration of symptoms (no prescription, three days v all delayed groups, four days; 4.29,  $P=0.368$ ). The groups showed modest and non-significant differences in antibiotic use. By contrast, significantly more patients given immediate antibiotics used them (97%), but with no benefit for symptom severity (score 1.76) or duration (median four days).

### Harms

Complications occurred in 2.5% (3/122) of patients in the no prescription group. Fewer complications were seen in the delayed and immediate prescription groups but this difference was not significant (adjusted risk ratio 0.56 (95% confidence interval 0.13 to 2.37); 0.66 (0.15 to 2.88)).

### Bias, confounding, and other reasons for caution

We used sealed numbered envelopes to execute randomisation, paying particular attention to practitioner equipoise, and found no evidence of subversion. Potential confounders were mostly balanced between groups, there was high retention (84%), and an intention to treat analysis was done without imputation. The trial was not powered for complications, nor to detect interaction between antibiotic strategies and other advice strategies; however, the other strategies did not modify outcomes so negative interactions due to ceiling effects are unlikely.

### Generalisability

The limited invitation to recruit participants reflects the realities of conducting trials in acute clinics at the busiest times of year. However, patients came from a wide range of practice settings and the sample had slightly more severe symptoms than large primary care observational cohorts.

### Effectiveness of antibiotic strategies in randomised groups

	No antibiotics	Delayed antibiotic prescription				Likelihood ratio test $\chi^2$ (P)
		Recontact	Post-date	Collection	Patient led	
<b>Mean symptom severity, days 2-4</b>						
Crude mean (standard deviation)	1.62 (0.88)	1.60 (0.91)	1.82 (0.94)	1.68 (0.88)	1.75 (0.88)	2.61 (0.625)
Adjusted mean difference* (95% CI; P; n=465)	—	-0.01 (-0.24 to 0.23; 0.964)	0.14 (-0.10 to 0.37; 0.249)	-0.02 (-0.27 to 0.22; 0.850)	0.08 (-0.16 to 0.33; 0.499)	
<b>Symptoms rated as moderately bad</b>						
Median duration (interquartile range)	3 (2-6.5)	4 (3-7)	4 (3-7)	4 (3-7)	4 (3-7)	4.29 (0.368)
Adjusted hazard ratio* (95% CI; P; n=455)	—	0.91 (0.66 to 1.25; 0.561)	0.86 (0.63 to 1.17; 0.338)	0.86 (0.62 to 1.20; 0.380)	0.71 (0.50 to 0.99; 0.045)	
<b>Antibiotic use</b>						
Crude no (%) of patients	26/99 (26)	34/92 (37)	37/101 (37)	28/85 (33)	35/89 (39)	4.96 (0.292)
Adjusted risk ratio* (95% CI; P; n=460)	—	1.45 (0.95 to 2.03; 0.083)	1.41 (0.92 to 1.98; 0.108)	1.28 (0.80 to 1.87; 0.275)	1.52 (1.00 to 2.10; 0.050)	

\*All models controlled for baseline symptom severity, dosing, steam, and smoking.