RESEARCH

The *BMJ* is an Open Access journal. We set no word limits on *BMJ* research articles, but they are abridged for print. The full text of each *BMJ* research article is freely available on bmj.com

- 14 RESEARCH NEWS All you need to read in the other general medical journals THIS WEEK'S RESEARCH QUESTIONS
- **16** How much does opiate substitution reduce the risk of HIV transmission among people who inject drugs?
- 17 What is the most effective tocolytic agent at delaying delivery?
- **18** How accurate is a single progesterone measurement in discriminating a viable from a non-viable pregnancy in women with pain or bleeding in early pregnancy?
- **19** Does using NSAIDs postoperatively increase the risk of anastomotic leakage after colorectal resection?
- **20** Do trial analyses need to be adjusted for stratification factors used in the randomisation process, how often are these factors adjusted for in practice, and is their use adequately reported?



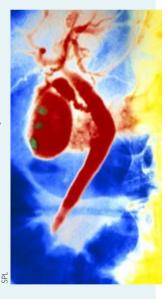




RESEARCH ONLINE: For these and other new research articles see www.bmj.com/research

Effect of intended intraoperative cholangiography and early detection of bile duct injury on survival after cholecystectomy

According to this study using data from the national Swedish registry for gallstone surgery and endoscopic retrograde cholangiopancreatography, survival among patients with bile duct injury during cholecystectomy is significantly impaired, compared with patients without bile duct injury. Furthermore, survival after bile duct injury is impaired by the failure to detect injury intraoperatively, and the intention to use intraoperative cholangiography during cholecystectomy improves survival significantly, say the authors.



Effect of hormone replacement therapy on cardiovascular events in recently postmenopausal women

After 10 years of randomised treatment, women receiving hormone replacement therapy early after menopause had a significantly reduced risk of death, heart failure, or myocardial infarction, without any apparent increase in risk of cancer, venous thromboembolism, or stroke, say the authors of this trial.

Safe exclusion of pulmonary embolism using the Wells rule and qualitative D-dimer testing in primary care

According to this prospective cohort study of 598 adults with suspected pulmonary embolism seen in primary care, a Wells score of ≤4 combined with a negative qualitative D-dimer test result can safely and efficiently exclude pulmonary embolism in primary care.

WHAT OUR READERS ARE SAYING

Benzodiazepine use and risk of dementia

In this prospective population based study, published on bmj.com on 27 September, new use of benzodiazepines was associated with an increased risk of dementia. The result was robust in pooled analyses across cohorts of new users of benzodiazepines throughout the study and in a complementary case-control study. Considering the extent to which benzodiazepines are prescribed and the number of potential adverse effects of this drug class in the general population, indiscriminate widespread use should be cautioned against, say the authors. Our rapid respondents said:

"Although the authors claim the study has a follow-up period long enough to overcome the hypothesis of reverse causation, we believe this is not the case. In the main analysis the delay between exposure and outcome (median 6.2 years) is too short to support the conclusion of an increased risk attributable to the use of benzodiazepines per se."

"It seems that at assessments, users of benzodiazepines were still using them. If this is correct, how were short term effects on cognition from ongoing use separated from long term effects?"

"It seems a shame that, in such a well planned and presented study, the 735 potential participants who had a preceding or prevalent history of benzodiazepine use were excluded. Inclusion of this group as a third arm might have allowed for greater powering of the association found—even with the trade-off of greater potential for confounding."



Opiate substitution treatment and HIV transmission in people who inject drugs: systematic review and meta-analysis

Georgie J MacArthur,¹ Silvia Minozzi,² Natasha Martin,¹³ Peter Vickerman,¹⁴ Sherry Deren,⁵ Julie Bruneau,⁶ Louisa Degenhardt,⁷⁸ Matthew Hickman¹

CEDITORIAL by Gowing

¹School of Social and Community Medicine, University of Bristol, Bristol BS8 2PS, UK

²Lazio Regional Health Service, Department of Epidemiology, Cochrane Drugs and Alcohol Group, Rome. Italy

³Department of Global Health and Development, London School of Hygiene and Tropical Medicine, London, UK

⁴Centre for Research on Drugs and Health Behaviour, London School of Hygiene and Tropical Medicine, London

⁵Center for Drug Use and HIV Research, New York University College of Nursing, New York City, USA

⁶Research Centre (CRCHUM), Centre Hospitalier de l'Université de Montréal, Montreal, Canada ⁷National Drug and Alcohol Research Centre, University of New South Wales, Sydney, Australia ⁸University of Melbourne, Centre for Health Policy, Programs and Economics, Melbourne, Australia Correspondence to: GJ MacArthur georgie.macarthur@bristol.ac.uk

Cite this as: BMJ 2012;345:e5945

doi: 10.1136/bmj.e5945

STUDY QUESTION What is the quantitative association between opiate substitution treatment and risk of HIV transmission among people who inject drugs?

SUMMARY ANSWER Evidence from published and unpublished observational studies shows that opiate substitution treatment is associated with a 54% reduction in risk of HIV infection among people who inject drugs.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS

Opiate substitution treatment is effective for heroin and other opioid dependence, and evidence suggests that associated reductions in injecting risk behaviour could reduce HIV incidence among people who inject drugs. The current study used wider search criteria to expand the evidence base, in part by identifying studies that report relevant data in the full text but not in the title or abstract, and quantified the association between opiate substitution treatment and reduced risk of HIV transmission.

Selection criteria for studies

We searched Medline, Embase, PsycINFO, and the Cochrane Library from inception to 2011 without language restriction to identify studies that directly examined the impact of opiate substitution treatment on HIV transmission. We also identified prospective studies that examined HIV incidence among opiate injectors that might have collected data regarding exposure to opiate substitution treatment but not have reported it. Authors of these studies were contacted.

Primary outcomes

The primary outcome was HIV seroconversion among people who inject drugs.

Main results and role of chance

We identified 12 published and three unpublished studies comprising 1016 incident HIV infections and over 26738 person years of follow-up. All included studies examined the impact of methadone maintenance treatment. Data from nine studies could be pooled in a random effects meta-analysis. Opiate substitution treatment was associated with a 54% reduction in the risk of HIV infection (rate ratio 0.46, 95% confidence interval 0.32 to 0.67; P<0.001). There was heterogeneity between studies (I²=60%, x²=20.12; P=0.010), though there was little evidence that this was explained by geographical region, the provision of incentives, site of recruitment, and the percentage of women or participants from ethnic minority groups. Weak evidence suggested that longer duration of exposure to opiate substitution treatment might be associated with greater benefit.

Bias, confounding, and other reasons for caution

All of the included studies were observational studies subject to bias, and the control of confounders was limited between studies. It is possible that the people who participated in the studies were not representative of everyone who receives opiate substitution treatment-for instance, they could under-represent short term injectors or those who reduce injecting during exposure to treatment. Individuals who enter opiate substitution treatment might also be more motivated to change behaviour and reduce how often they inject, which could overestimate the effect of treatment. From the data reported in published studies we could not examine the impact of interventions or other factors that might affect the impact of opiate substitution treatment-such as needle/syringe programmes or psychosocial interventions-or the extent of practical or social support. Lastly, as all of the included studies examined the impact of methadone, we cannot generalise our findings to other forms of opiate substitution treatment. As HIV incidence rates varied substantially between the sites we reported the rate reduction, rather than an absolute measure of effect (the risk difference), which would not be generalisable to other sites.

For funding statement see bmj.com

Meta-analysis of studies showing impact of opiate substitution treatment (OST) in relation to HIV transmission in people who inject drugs

Study	Effect estimate	Effect estimate
All pooled studies	(95% CI)	(95% CI)
Williams 1992 ⁴⁶		0.16 (0.02 to 1.30)
Metzger 1993 ⁴⁴		0.13 (0.03 to 0.50)
Chitwood 1995 ³⁹		0.29 (0.09 to 0.94)
Nelson 2002 ⁸	- -	0.56 (0.34 to 0.92)
Kerr 2006 ³⁷	÷ - +	0.67 (0.42 to 1.10)
Van den Berg 2007 ⁴⁵	- 	0.35 (0.23 to 0.54)
Suntharasamai 2009 ¹⁷	-	0.78 (0.58 to 1.05)
Judd 2012 (unpublished)		- 0.79 (0.20 to 3.15)
Bruneau 2012 (unpublished)		0.17 (0.04 to 0.71)
Overall: I ² =60%, P=0.010	🔶 🔶	0.46 (0.32 to 0.67)
Studies reporting adjusted effect estimates	;	
Metzger 1993 ⁴⁴		0.19 (0.06 to 0.61)
Chitwood 1995 ³⁹		- 0.70 (0.18 to 2.81)
Kerr 2006 ³⁷		0.83 (0.51 to 1.35)
Suntharasamai 2009 ¹⁷		0.60 (0.40 to 0.80)
Judd 2012 (unpublished)		- 0.81 (0.20 to 3.23)
Bruneau 2012 (unpublished)		0.30 (0.07 to 1.27)
Overall: I ² =23%, P=0.262	🔶 👘	0.60 (0.42 to 0.85)
C	0.01 0.1 1	10
F	Favours	Favours
(DST	no OST

Tocolytic therapy for preterm delivery: systematic review and network meta-analysis

David M Haas,¹ Deborah M Caldwell,² Page Kirkpatrick,¹ Jennifer J McIntosh,¹ Nicky J Welton²

CEDITORIAL by Alfirevic

¹Department of Obstetrics and Gynecology, Indiana University School of Medicine, Indianapolis, IN, USA

²School of Social and Community Medicine, University of Bristol, Bristol, UK

Correspondence to: D M Haas Wishard Memorial Hospital, 1001 W 10th Street, F5102, Indianapolis, IN 46202, USA

dahaas@iupui.edu

Cite this as: *BMJ* **2012;345:e6226** doi: 10.1136/bmj.e6226

This is a summary of a paper that was published on bmj.com as *BMJ* 2012;345:e6226

bmj.com

 Editorial: Tocolytics and preterm labour (*BMJ* 2009;338:b195)
Research: Adverse drug reactions to tocolytic treatment for preterm labour (*BMJ* 2009;338:b744) **STUDY QUESTION** Given the complexity of studies in the area of tocolysis, what is the most effective tocolytic agent at delaying delivery?

SUMMARY ANSWER Prostaglandin inhibitors and calcium channel blockers had the highest probability of delaying delivery and improving neonatal and maternal outcomes.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS It is known that many different tocolytics are used to delay preterm delivery and that a standard first line drug has not emerged. This novel network meta-analysis of tocolytic therapy showed that when considering all tocolytic trials together, prostaglandin inhibitors and calcium channel blockers had the highest probability of delaying delivery and improving neonatal outcomes.

Selection criteria for studies

Randomized controlled trials of tocolytic therapy in women at risk of preterm delivery.

Primary outcomes

48 hour delay in delivery, neonatal respiratory distress syndrome, neonatal mortality, and maternal side effects.

Main results and role of chance

Utilizing data from 95 randomized controlled trials, the probability of delivery being delayed by 48 hours was highest with prostaglandin inhibitors (odds ratio 5.39, 95% credible interval 2.14 to 12.34) followed by magnesium sulfate (2.76, 1.58 to 4.94), calcium channel blockers

(2.71, 1.17 to 5.91), beta mimetics (2.41, 1.27 to 4.55), and the oxytocin receptor blocker atosiban (2.02, 1.10 to 3.80). No class of tocolytic was significantly superior to placebo in reducing neonatal respiratory distress syndrome. Compared with placebo, side effects requiring a change of medication were significantly higher for beta mimetics (22.68, 7.51 to 73.67), magnesium sulfate (8.15, 2.47 to 27.70), and calcium channel blockers (3.80, 1.02 to 16.92). Prostaglandin inhibitors and calcium channel blockers were the tocolytics with the best probability of being ranked in the top three medication classes for the outcomes of 48 hour delay in delivery, respiratory distress syndrome, neonatal mortality, and maternal side effects (all cause).

Bias, confounding, and other reasons for caution

Meta-analyses of rare events can be problematic and few neonatal deaths were reported in the trials. Not all trials reported on all outcomes. Also, many trials used different treatment protocols, leading to some clinical heterogeneity.

Study funding/potential competing interests

This study was supported by grants: NIH-NICHD K23HD055305 (DMH) and the Indiana University-Purdue University-Indianapolis Signature Center grant to the Indiana University Center for Pharmacogenetics and Therapeutics Research in Maternal and Child Health (PREGMED). The funding agencies had no role in the study design, implementation, or preparation of results. No author activities directly conflict with the network meta-analysis presented.

Efficacy, neonatal respiratory distress syndrome, and tolerability of tocolytics, using placebo as reference class. Values are posterior median odds ratios (95% credible intervals) unless stated otherwise

Drug class	48 hour delay in delivery*	Probability of being best†	Neonatal mortality‡	Probability of being best	Neonatal respiratory distress syndrome‡	Probability of being best	Maternal side effects§	Probability of being best
Placebo or control	_	<0.01	-	0.01	-	0.02	-	0.61
Beta mimetic	2.41 (1.27 to 4.55)	0.01	0.62 (0.14 to 2.48)	0.12	0.85 (0.50 to 1.45)	0.14	22.68 (7.51 to 73.67)	<0.01
Prostaglandin inhibitor	5.39 (2.14 to 12.34)	0.83	0.62 (0.04 to 4.63)	0.28	0.87 (0.40 to 1.75)	0.20	1.63 (0.40 to 6.85)	0.21
Calcium channel blocker	2.71 (1.17 to 5.91)	0.06	0.39 (0.09 to 1.49)	0.41	0.71 (0.37 to 1.43)	0.47	3.80 (1.02 to 16.92)	0.01
Other	0.93 (0.13 to 6.14)	0.04	2.79 (0.28 to 31.75)	0.02	1.54 (0.55 to 4.71)	0.04	-	_
Magnesium sulfate	2.76 (1.58 to 4.94)	0.02	0.97 (0.29 to 3.29)	0.03	0.99 (0.58 to 1.71)	0.03	8.15 (2.47 to 27.70)	<0.01
Oxytocin receptor blocker (atosiban)	2.02 (1.10 to 3.80)	0.01	0.62 (0.16 to 2.35)	0.13	0.89 (0.55 to 1.37)	0.10	1.99 (0.61 to 6.94)	0.08
Nitrates	1.91 (0.64 to 5.33)	0.04	-	_			3.19 (0.41 to 20.84)	0.10

*Odds ratios >1 favor active class.

†Probability that given drug class is best agent to use for given outcome based on rankings over all eight drug classes.

‡Odds ratios <1 favor active class.

§Odds ratios >1 favor placebo

Accuracy of single progesterone test to predict early pregnancy outcome in women with pain or bleeding: meta-analysis of cohort studies

Jorine Verhaegen,¹ Ioannis D Gallos,² Norah M van Mello,³ Mohamed Abdel-Aziz,² Yemisi Takwoingi,³ Hoda Harb,² Jonathan J Deeks,³ Ben W J Mol,¹ Arri Coomarasamy²

¹Department of Obstetrics and Gynaecology, Academic Medical Centre University of Amsterdam. Amsterdam, Netherlands School of Clinical and Experimental Medicine, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK ³Public Health, Epidemiology and Biostatistics, School of Health and Population Sciences, College of Medical and Dental Sciences. University of Birmingham Correspondence to: I D Gallos, Nuffield Department of Obstetrics and Gynaecology, University of Oxford, Oxford Radcliffe Hospitals NHS Trust, The Women's Centre, Oxford OX3 9DU, UK

ioannis.gallos@nhs.net

Cite this as: *BMJ* **2012;345:e6077** doi: 10.1136/bmj.e6077

This is a summary of a paper that was published on bmj.com as *BMJ* 2012;345:e6077

STUDY QUESTION How accurate is a single progesterone measurement in discriminating a viable from a non-viable pregnancy in women with pain or bleeding in early pregnancy?

SUMMARY ANSWER A single progesterone measurement for women in early pregnancy presenting with bleeding or pain and inconclusive ultrasound assessments can rule out a viable pregnancy.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS A single progesterone measurement in early pregnancy has been suggested to be a useful test for discriminating between viable and non-viable pregnancies, but its accuracy remains controversial. A single progesterone measurement can discriminate between a viable and a non-viable pregnancy for women in early pregnancy with pain or bleeding and can rule out a viable pregnancy in women with inconclusive ultrasound examinations.

Selection criteria for studies

We searched Medline, Embase, CINAHL, Web of Science, ProQuest, Conference Proceedings Citation Index, and the Cochrane Library from inception until April 2012, as well as reference lists of relevant studies. We selected studies on the basis of the participants (women with spontaneous pregnancy of less than 14 weeks of gestation), test (single serum progesterone measurement), outcome (viable intrauterine pregnancy, miscarriage, or ectopic pregnancy), and design (cohort studies of test accuracy).

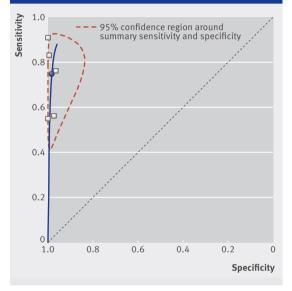
Primary outcome(s)

The main outcome was the accuracy of a single progesterone measurement in discriminating a viable from a nonviable pregnancy for women in early pregnancy with pain or bleeding.

Main results and role of chance

We included 26 cohort studies (including 9436 pregnant women), consisting of seven studies in women with symptoms and inconclusive ultrasound assessment and 19 studies in women with symptoms alone. Among women with symptoms and inconclusive ultrasound assessments, the progesterone test (five studies with 1998 participants and cut-off values from 3.2 to 6 ng/mL) predicted a non-viable pregnancy with pooled sensitivity of 74.6% (95% confidence interval 50.6% to 89.4%), specificity of 98.4% (90.9% to 99.7%), positive likelihood ratio of 45 (7.1 to

Summary receiver operating characteristics plot of progesterone test used to identify non-viable pregnancies in women with pain or bleeding and inconclusive ultrasound assessment



289), and negative likelihood ratio of 0.26 (0.12 to 0.57). The overall median prevalence of a non-viable pregnancy was 73.2%, and the probability of a non-viable pregnancy was raised to 99.2% if the progesterone was low. For women with symptoms alone, the progesterone test had a higher specificity using a 10 ng/mL threshold (9 studies with 4689 participants) and predicted a non-viable pregnancy with pooled sensitivity of 66.5% (53.6% to 77.4%), specificity of 96.3% (91.1% to 98.5%), positive likelihood ratio of 18 (7.2 to 45), and negative likelihood ratio of 0.35 (0.24 to 0.50). The probability was raised from 62.9% to 96.8%.

Bias, confounding, and other reasons for caution

Many studies had very different prevalences of pregnancy outcomes, reflecting the different clinical settings and populations that the test was evaluating. We report the median prevalence with its respective range for non-viable pregnancies, from which we generated post-test probabilities to make our results generalisable and applicable in a variety of settings in clinical practice.

Study funding/potential competing interests

No funding was sought for this study.

Postoperative use of non-steroidal anti-inflammatory drugs in patients with anastomotic leakage requiring reoperation after colorectal resection: cohort study based on prospective data

Mads Klein, Ismail Gögenur, Jacob Rosenberg

Department of Surgery, Herlev Hospital, University of Copenhagen, 2730 Herlev, Denmark Correspondence to: M Klein madsklein1@gmail.com

Cite this as: *BMJ* **2012;345:e6166** doi: 10.1136/bmi.e6166

This is a summary of a paper that was published on bmj.com as *BMJ* 2012;345:e6166

STUDY QUESTION Does the postoperative use of non-steroidal anti-inflammatory drugs (NSAIDs) increase the risk of anastomotic leakage after colorectal resection?

SUMMARY ANSWER Postoperative diclofenac treatment resulted in an increased proportion of patients having anastomotic leakage evaluated at reoperation after colorectal resections with primary anastomosis for colorectal cancer.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS

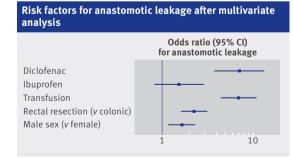
Retrospective studies have indicated an increased risk of anastomotic leakage with postoperative NSAID treatment. Using prospective data and precise information on postoperative NSAID consumption, we found that treatment with NSAIDs that are predominantly cyclo-oxygenase-2 selective (such as diclofenac) should be avoided after colorectal surgery owing to an increased risk of anastomotic leakage.

Participants and setting

We included 2756 patients (1441 (52%) men) with available data who had undergone an elective operation for colorectal cancer with colonic or rectal resection and primary anastomosis at six specialised centres in eastern Denmark. Of these patients, 1871 did not receive regular postoperative treatment of NSAIDs (defined as at least two days' treatment in the first seven postoperative days), and 885 did.

Design, size, and duration

Cohort study based on information about NSAID consumption, from a prospective clinical database and elec-



tronic medical records. Analysis focused on patients who had undergone operations between 1 January 2006 and 31 December 2009. Univariate and multivariate logistic regression analyses were performed to identity risk factors for anastomotic leakage.

Main results and the role of chance

The rate of anastomotic leakage requiring reoperation was significantly increased among patients receiving diclofenac and ibuprofen treatment, compared with controls receiving no NSAID treatment (12.8% and 8.2% v 5.1%; P<0.001). In effect, after unadjusted analyses and when compared with controls, the absolute risk increase for anastomotic leakage was 7.8% (95% confidence interval 3.9% to 12.8%) and 3.2% (1.0% to 5.7%) after treatment with diclofenac and ibuprofen, respectively. In a multivariate logistic regression analysis, only diclofenac treatment was a risk factor for anastomotic leakage (odds ratio 7.2 (95% confidence interval 3.8 to 13.4), P<0.001; ibuprofen 1.5 (0.8 to 2.9), P=0.18).

Bias, confounding, and other reasons for caution

Variables (for example, obesity, diabetes, or ischaemic heart disease) not registered in the database could not be adjusted for. For several reasons, confounding by indication was not a factor in this study. Firstly, surgical centres use standard analgesic regimens that are already prescribed before elective colorectal resections. Secondly, our definition of NSAID treatment ruled out patients receiving a single dose administered in the acute setting owing to increased abdominal pain or other reasons for pain. Thirdly, we did multivariate logistic regression and adjusted for interactions between variables to reduce confounding. Finally, for breakthrough pain, strong opioids are normally used in Denmark.

Generalisability to other populations

Results can be extrapolated to other populations undergoing colorectal resection with primary anastomosis for colorectal cancer.

Study funding/potential competing interests

No external funding was received, and the authors have no competing interests.

Reporting and analysis of trials using stratified randomisation in leading medical journals: review and reanalysis

Brennan C Kahan, Tim P Morris

MRC Clinical Trials Unit, Aviation House, 125 Kingsway, London WC2B 6NH, UK Correspondence to: B C Kahan brk@ctu.mrc.ac.uk

Cite this as: *BMJ* 2012;345:e5840 doi: 10.1136/bmi.e5840

This is a summary of a paper that was published on bmj.com as *BMJ* 2012;345:e5840

bmj.com

 Research: The influence of study characteristics on reporting of subgroup analyses in randomised controlled trials (*BM*/ 2011;342:d1569)
Research: Reporting of sample size calculation in randomised controlled trials (*BM*/ 2009;338:b1732) **STUDY QUESTION** Do trial analyses need to be adjusted for stratification factors used in the randomisation process, how often are these factors adjusted for in practice, and is their use adequately reported?

SUMMARY ANSWER Adjustment for stratification is needed, but in about three quarters of trials published in four general medical journals in 2010, adjustment was inadequate, and in over a third of them, reporting was poor.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS

Ignoring stratification factors in the analysis leads to P values that are too large and a reduction in power. Stratified randomisation is extremely common but is rarely accounted for in the analysis, potentially leading to incorrect P values and a loss of power.

Selection criteria for studies

We identified trials from four general medical journals (*BMJ*, *Journal of the American Medical Association, Lancet*, and *New England Journal of Medicine*) published in 2010 using their electronic table of contents. Cluster randomised, crossover, non-randomised, single arm, and phase I or II trials were excluded, as were trials reporting secondary analyses, interim analyses, or results that had been previously published in 2010.

Reporting and analysis of trials						
Variables	No (%) of trials (n=258)					
Balanced on centre	120 (47)					
Adjusted primary analysis for centre (n=120)*	31 (26)					
Balanced on prognostic factors	111 (43)					
Adjusted primary analysis for prognostic factors (n=111)*	40 (36)					
Balanced on centre or prognostic factors	163 (63)					
Adjusted primary analysis for centre or prognostic factors $(n=163)^*$	42 (26)					

*Trials were only assessed for whether they adjusted for centre or prognostic factors if they had balanced on these factors.

Primary outcomes

Whether the method of randomisation was adequately reported, how often stratified randomisation was used, and whether stratification factors were appropriately adjusted for in the analysis. To assess the impact of an unadjusted analysis after balanced randomisation has been used, we reanalysed data from a published randomised trial, the second Multicenter Intrapleural Sepsis Trial (MIST2).

Main results and role of chance

Reanalysis of MIST2 found that not adjusting for stratification factors in the analysis led to larger P values for need for surgery (0.095 v 0.175 for adjusted and unadjusted, respectively) and time to hospital discharge (0.011 v 0.044), which could lead to a reduction in power. Our review identified 258 eligible trials, 163 (63%) of which used at least one stratification factor in the randomisation process. The most common methods of balancing on baseline covariates were stratified permuted blocks (n=85) and minimisation (n=27). The method of randomisation was unclear in 37% of trials. Most trials that used centre or prognostic variables as stratification factors were not adequately analysed; only 26% of trials adjusted for all stratification factors in their primary analysis. Trials that did not adjust for stratification factors in their analysis were less likely to show a significant result (unadjusted 57% v adjusted 78%, P=0.02).

Bias, confounding, and other reasons for caution

Our review was limited to articles published in four major medical journals, which is unlikely to be a representative sample as articles published in other medical journals may have different reporting standards.

Study funding/potential competing interests

Both authors are employed by the MRC Clinical Trials Unit. TPM is funded by an MRC studentship (MC-US-A737-0012). We have no competing interests.

BMJ pico: advice to authors

The full text of all accepted *BMJ* research articles is published online in full, with open access and no word limit, on bmj. com as soon as it is ready. In the print *BMJ* each research article is abridged, as a one page BMJ pico, with the aim of making research more inviting and useful to readers. Since August 2009, authors have written their own BMJ picos.

We have designed BMJ pico with evidence based medicine experts to succinctly present the key evidence from each study, to help minimise delay between online and print publication, and to enable us to publish more research in each week's print *BMJ*. For more details, see http://tinyurl. com/kp5c7o/.

There is no need for authors to prepare a BMJ pico to submit along with the full research article. Authors produce their own BMJ pico, using a template from us, only after the full article has been accepted.

Because publication of research on bmj.com is definitive, rather than interim "epublication ahead of print," authors who do not wish to abridge their articles using BMJ pico will be able to opt for online only publication.

BMJ | 20 OCTOBER 2012 | VOLUME 345