EJA

Eur J Anaesthesiol 2016; **33**:1–5

INVITED COMMENTARY

The quest for the holy volume therapy

Edoardo De Robertis, Arash Afshari and Dan Longrois

European Journal of Anaesthesiology 2016, **33**:000–000

This Invited Commentary accompanies the following article:

Marx G *et al*. Intravascular volume therapy in adults. Guidelines from the German Society of Anaesthesiology and Intensive Care Medicine. *Eur J Anaesthesiol* 2016; **X**:000–000

'Beware of false knowledge; it is more dangerous than ignorance' (G.B. Shaw).

In this issue of the *European Journal of Anaesthesiology*, the German Society of Anaesthesiology and Intensive Care Medicine (DGAI) present their guideline on intravascular volume therapy in adults.[1] It is a systematically developed body of information and recommendation based on a careful analysis of both evidence and established practice that is the basis for decision-making in diagnosis and treatment. The process used to develop this guideline was systematic, independent and transparent, but although the methodology was accurate, it was based on literature that is not without flaws. As a result, for crucial aspects, the strength of recommendation is too often only 'may' (open recommendation) or 'should', and just a few are strong enough for 'must'.

The guidelines tell us that:

(1) Fluid therapy should be based on monitoring. Volume responsiveness is best evaluated by dynamic indices.
(2) The question of whether a volume-responsive patient requires volume expansion and increased cardiac output or not is a difficult issue, and the simple answer is that we do not really know.
(3) It is impossible to recommend one type of fluid over another because the evidence is unsound (see below).

What the guidelines do not address is the timing and extent of volume therapy; must a volume deficit be corrected totally and rapidly? Just to provide food for thought, when thirst, triggered by a tiny 2% increase in plasma osmolality or a 10% decrease in central blood volume, and mediated by an increase in vasopressin, can be corrected by drinking, the sensation of thirst disappears long before correction of the high osmolality or volume deficit.[2]

Much of this guideline is taken up with issues of monitoring, pathophysiological paradigms and methodology. A critical analysis requires that these are dealt with in turn, and an appropriate way to begin is with a brief historical perspective.

A brief historical perspective

In 1832, during the cholera outbreak, a warm hypotonic solution of 'two drachams of muriate, two scruples of carbonate of soda to sixty ounces of water' was first administered intravenously to patients, with '... an immediate return of the pulse, and improvement in the respiration... '.[3] Soon after, *The Lancet* observed in an editorial that '...the mass of the profession is unable to decide; and thus, instead of any uniform mode of treatment, every town and village has its different system or systems ... a suitable clinical investigation is required to resolve between such conflicting authorities...'.[4]

After almost 200 years, those words still sound familiar as an ideal volume expansion regimen (for both intravascular and interstitial spaces) remains a matter of controversy and uncertainty among physicians; one that is frequently discussed in anaesthesia and intensive care medicine.

From the Department of Neurosciences, Reproductive and Odontostomatologic Sciences - Section of Anaesthesia and Intensive Care Medicine, University of Naples Federico II, Naples, Italy (EDR), Department of Anaesthesia, Mother and children section, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark (AA),and Department of Anaesthesia and Intensive Care, Bichat-Claude Bernard Hospital, Assistance Publique Hôpitaux de Paris and Université Paris 7- Denis Diderot, Paris, France (DL)

Correspondence to Professor Edoardo De Robertis, MD, PhD, Department of Neurosciences, Reproductive and Odontostomatologic Sciences - Section of Anaesthesia and Intensive Care Medicine, University of Naples Federico II, via S. Pansini 5, 80131 Naples, Italy
Tel: +39 081 7463550 42; e-mail: ederober@unina.it

DOI:10.1097/EJA.0000000000000441

Monitoring

Over the course of more than one decade, crystalloid and colloid solutions have been greatly improved, and haemodynamic monitoring has significantly evolved. It is clearly recognised that to avoid fluid overload and prevent hypoperfusion, hypoxia and organ failure, accurate haemodynamic monitoring, an assessment of global and regional tissue perfusion and judicious management of volume are of paramount importance, as are maintenance of fluid compartments and function of the vascular barrier.

Even now, many patients are still not appropriately monitored, the volumes required remain essentially empirical and we still lack a solid pathophysiological foundation on which to base the choice of solution when meeting everyday demands in operating theatres and intensive care units. Questions about efficacy and safety are unresolved, and the quest for the goal of the appropriate volume therapy continues.[5–7] A considerable body of work has been amassed in addressing this topic, but it occasionally overlooks the fact that haemodynamic monitoring *per se* has no favourable impact on outcome, and that when prescribing a fluid solution, like any drug, understanding the indication and the dose are crucial if benefit rather than harm is to result. Haemodynamic monitoring has been too often used to maximise rather than optimise fluid management, and new minimal or non-invasive monitoring devices have distorted its traditional graduation.

At the bedside, assessment of haemodynamic status involves both myocardial performance and the entire spectrum of volume imbalance. Thus, haemodynamic stability and tissue perfusion should be monitored by a combination of clinical examination (often neglected), monitoring devices and laboratory results. Over the years, we have seen many static haemodynamic indices presented, but they have failed to offer a convincing picture of true volume status. Conventional indices, such as heart rate, blood pressure, central venous pressure, pulmonary artery occlusion pressure and urine output, traditionally used to guide fluid management, are influenced by many factors. They can no longer be considered as sensitive and specific, and capable of revealing impaired tissue perfusion and oxygenation, with risk of organ dysfunction. Measuring a pressure and extrapolating a volume is not enough to predict fluid responsiveness; nor is the direct measurement of static volumes, as the slope of the Frank–Starling curve depends on systolic function. Dynamic changes in pulse pressure and stroke volume because of heart–lung interaction have been shown to be sensitive to changes in preload and can better predict the haemodynamic response to volume expansion.[8,9] Accordingly, today several non-invasive or minimally invasive dynamic tests are available to assess volume responsiveness by monitoring the changes in stroke volume induced by modification of preload and analysing the Frank–Starling curve.

The aim of volume therapy and haemodynamic monitoring is to avoid overload and hypoperfusion, both of which interfere with tissue perfusion and oxygenation. Healthy hearts of patients undergoing surgery are fluid responsive. The distressed heart of the critically ill behaves differently. That fluid responsiveness can be demonstrated does not necessarily imply that fluids should be administered, but if coupled with evidence of hypoperfusion, then therapy is needed. Patients undergoing elective surgery should be normovolaemic with a functioning vascular barrier and intact fluid compartments, where an intact tight glycocalyx/vascular endothelial junction can properly retain colloids. In contrast, the critically ill, systemically inflamed patient, with the integrity of the vascular barrier disturbed, risks migration of large molecules outside the circulation. Accordingly, an approach to haemodynamic monitoring and volume therapy should integrate all available information and should be tailored to the need of each individual and the procedure performed. Different patient groups cannot be mixed up.

Pathophysiological paradigms pertinent to the crystalloid versus colloid controversy

Reproducibility in biomedical research has always been a concern. In 2005, John Ioannidis stated that the majority of the published results were 'false' in that they are not reproduced by subsequent work.[10] In his opinion, this may be because of the fact that a *P* value less than 0.05 is much too high, because of inherent bias and because, in the Bayes theorem, the post-test probability that an intervention has a measurable (beneficial) effect depends on the pre-test probability that the intervention has an effect. When applied to the crystalloid versus colloid controversy, the pre-test probability that colloid and crystalloid have different effects on outcome should be based on physiological and pathophysiological paradigms.

The first would be that the use of colloids is associated with higher intravascular oncotic pressure and, because of the Starling equilibrium, this would 'draw' interstitial fluid toward the circulation, preserving intravascular volume and avoiding tissue oedema. Both should contribute to improved oxygen transport and tissue oxygenation at the cellular level.[11] The Starling equilibrium is explained by the barrier function of the capillary endothelial monolayer in which there is a cell–cell and cell–extracellular matrix interaction. For this paradigm to be true, that is validated by observation and experiment, the structure and function of the capillary endothelial barrier, where exchanges occur between the intravascular and the interstitial sub-compartments, must remain undisturbed. For many years this has been the basis of clinical reasoning that argues for the benefit of a colloid over a crystalloid without any serious clinical challenge, at least if we analyse the most recent studies on this subject.

Not so recently, laboratory studies did demonstrate that under certain conditions, there was a discrepancy

between measured Starling forces in the plasma and tissue, and the forces that determine filtration. The work of Michel[12] and Weinbaum[13] suggested that the effective oncotic force across the capillary is determined by the local difference in protein concentration across the surface-matrix layer (the glycocalyx) rather than the global difference in concentration between the plasma and the interstitial fluid in the tissue.

Disruption of the endothelial barrier is the result of both alteration in the glycocalyx and compromised inter-endothelial cell junctions. Glycocalyx degradation, induced by many stimuli, such as inflammation, cytokines and natriuretic peptides secreted in response to acute dilatation of the heart secondary to iatrogenic hypervolaemia, abolishes the effects of both albumin and colloids such as hydroxyethyl starch (HES) in experimental models.[14] This could explain why, in the critically ill with either chronic alteration of the glycocalyx and/or *de novo* disruption, the effects of colloids (including albumin) and crystalloids result in very similar (if different at all) effects on morbidity and mortality.

The second paradigm is the consequence of the first and concerns the initial status of intravascular volume. As Chappel and Jacob explained,[11] in terms of volume of distribution and intravascular persistence, colloids and crystalloids behave as Starling had predicted when hypovolaemia is treated by volume expansion. But this does not apply when the normovolaemic are given colloids or crystalloids resulting in hypervolaemic haemodilution. This triggers secretion of natriuretic peptides that activate metalloproteases that in turn alter the glycocalyx,[15] with the consequences described above. Hypervolaemia may transform a healthy endothelium into a diseased endothelium.

It is highly probable that when the critically ill are randomised to receive either crystalloids or colloids, their pre-infusion volume status is quite variable. Some might be hypovolaemic whereas others are normovolaemic, especially when randomisation has occurred 24 h after admission to the intensive care unit. It is probable that during this period their volume status has been corrected and by the time they are exposed to the allocated infusion, the vast majority are normovolaemic. When single studies are included in meta-analyses, it is highly probable that the heterogeneity of the patients is such that the beneficial effects of one type of volume expansion solution is counterbalanced by the deleterious effects in other groups of patients. As a result, it is hard to identify a clinically important difference among the volume expanding solutions under study. It seems that when clinical trials are designed, reported and analysed these details are overlooked.

The third paradigm is that changes in the composition of plasma, secondary to volume expansion with nonblood solutions, modify blood rheology. Both plasma viscosity and the behaviour of red blood cells have been shown to be modified by different types of crystalloids and colloids; each solution has a specific signature on plasma and red blood cells rendering the generic terms 'crystalloids' or 'colloids' inappropriate.[16] Plasma viscosity is regulated by the plasma concentration of proteins and fibrinogen and is critical for maintaining the functional capillary density (FCD). Even during extreme anaemia, and decreased oxygen transport, maintenance of plasma viscosity was shown to preserve FCD and tissue perfusion. Colloid solutions when compared with crystalloids, and matched for similar degrees of haemodilution, have been shown *in vitro* to increase plasma viscosity. In-vitro HES 200/0.5 increased plasma viscosity more than HES 130/0.4 but the difference was not statistically significant *in vivo* in patients with traumatic brain injury.[17] In a porcine model of severe haemorrhagic shock, animals were resuscitated at the extreme of the viscosity spectrum with either hyper-HES or Ringer's lactate as was required to achieve similar macro-haemodynamic values. Despite the administration of a six-fold higher volume of Ringer's compared with hyper-HES, plasma viscosity values were similar.[18] The mechanisms that maintain plasma viscosity in the face of acute challenges are not known, but it is highly probable that modern colloids, with lower molecular mass and lower volumes/unit of time, do not change it sufficiently to affect FCD. If there is no detectable difference between crystalloids and colloids with respect to plasma viscosity, then the probability is that their impact on capillary flow and tissue perfusion is low.

To summarise, the three physiological and pathophysiological paradigms that could account for a beneficial effect of colloids compared with crystalloids are such that the pre-test probability that colloids improve outcome compared with crystalloids is low. The strongest argument for this interpretation is that throughout the German guidelines (and elsewhere in the literature), there is little or no evidence for improved efficacy of colloids compared with crystalloids, and most of the debate that would advocate for or against colloids in clinical practice is concerned with serious side-effects. Coupled with methodological problems (see below) these factors together might explain why we have difficulty formulating clear guidelines for crystalloids and colloids in routine clinical practice.

Methodological aspects of creating guidelines

In essence, the validity of the recommendations made by guidelines in general depends on the stringency with which the authors have assessed the existing evidence. This is essential for the German guidelines if they are to be trusted by clinicians and implemented in clinical practice. In this regard, the authors of guidelines have to scrutinise not only the literature to find relevant papers

but should look beyond the classical hierarchical pyramid of evidence, which often fails to search further than randomised controlled trials (RCTs) and systematic reviews superior to other types of studies and evidence.

Assessment of bias is the core of the guideline creation process but to achieve this there must be a systematic approach with incorporation of tools to grade the recommendations. Tools such as Preferred Reporting Items of Systematic reviews and Meta-Analyses (PRISMA), Consolidated Standards of Reporting Trials (CONSORT) statements, the Cochrane risk of bias tool and the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system should not only be applied by the authors of guidelines but ideally should also be incorporated in the peer-review process of the original papers that are identified in the search.[19–22] These tools should be mandatory and vigorously promoted by the editors of journals but unfortunately this is still not the case.[23,24]

One other major issue that may undermine the validity of guidelines is our underlying trust in the published literature and its peer-review process. Unfortunately, our discipline has experienced serious setbacks recently because of fraudulent science and peer review.[25] The widespread implementation of tools to assess the impact of bias may help restore our trust.

But, if guidelines naively take for granted that the peer-review process is a benchmark for quality assurance, and merely grade the evidence based on a publication label (the superiority of systematic reviews over RCTs, and RCTs over observational studies), then despite implementation of tools such as GRADE and the Appraisal of Guidelines for Research and Evaluation (AGREE), there is still a great risk that the underlying evidence itself is biased and based on flawed methodology. Poor quality studies often overestimate the intervention effects,[26,27] and their outcomes are that flawed and false recommendations appear in guidelines.

Systematic reviews serve an increasingly important role in the creation of guidelines and are used to identify benefits or harms of interventions which can justify funding for further research.[28] But where the fluid therapy debate is concerned, different systematic reviews have come to different conclusions. Why is this so? The cause of discrepancy often lies in the different ways of assessing the limitations of the included trials with regard to their design, conduct, analysis and presentation of data. And this brings us back to the issue of bias. If the results of the included studies are biased, then the systematic reviews will suffer equally from high risk of bias unless the authors recognise this and adjust the evidence accordingly. If it is not dealt with properly, the underlying bias becomes a systemic bias and the recommendations of the guidelines will be consequently mired. There is solid evidence to support a stringent assessment of systematic reviews as the introduction of PRISMA statements has improved their quality.[23,29]

Another challenging issue is the interpretation of scarce evidence (small underpowered trials) and how to formulate recommendations for clinical practice. Often, readers and authors of guidelines have a misperception that systematic reviews and meta-analyses can resolve the issue of lack of power. But this is far from the truth.[30] Indeed, the inclusion of trials with inadequate sample size in systematic reviews leads to increased risk of random error and may ultimately lead to false conclusions and uncertainty.[31] This problem grows larger as systematic reviews are updated with newer studies, increasing the risk of type I error, which may be as high as 30% (also known as multiplicity because of multiple and repeated significance testing).[32–35]

Systematic reviews with inadequate power and small event rates examining the outcomes of interest may not only blur the overall picture by overestimating the intervention effect, but may equally fail to provide reliable conclusions, or at worst delay decisions, on the harms or benefits of an intervention because of bias, random error and lack of precision.[30,36,37]

As previously advocated in this journal, sequential methods such as trial sequential analysis can provide information on when firm evidence is available in meta-analysis.[38] The need for the application of sequential methods to examine the robustness of findings in systematic reviews with meta-analysis is not to be underestimated, as the number of false positive meta-analyses is thought to be between 16 and 37%. Their inclusion in guidelines inevitably increases the risk of false recommendations with severe consequences for our clinical decision-making and for treatment of our patients.[39]

This is indeed the case when it comes to the safety of synthetic colloids in the perioperative setting. One may not only criticise the trials included in the various meta-analysis as being too small with high risk of bias, suffering from design shortcomings and inadequate or short follow-up, but, more importantly, the same criticism can be applied to almost all of the published systematic reviews on this topic; when trial sequential analysis is carried out, it becomes clear that they suffer from inadequate power. Any assumption as to whether starches are well tolerated or harmful in the perioperative setting does not appear to be supported by solid evidence at this stage, as the power that would justify support or rejection is inadequate. But it would be possible to argue against their use as long as there is continuing uncertainty about their safety, at least until enough evidence is provided.

Having read the German guidelines, and this editorial, anaesthesiologists and intensivists should adhere to routine practice but should accept the following principles:

(1) Building the 'truth' (that would allow solid recommendations) in clinical practice is based on uncertainty, and that will continue;
(2) Guidelines, despite their limitations, are a necessary step to conformity in clinical practice and should be followed critically (and this editorial is intended to build constructive criticism in the readers' minds);
(3) Changes in direction in clinical practice, at times turning 180°, are to be expected;
(4) Clinical and basic scientists will collaborate to design new studies on new and updated paradigms.

This is, in our opinion, one of the most credible ways to evolve from 40 years of controversy over volume therapy that has taken us only to 'Uncertainty-land', something that clinicians (and this is highly understandable) profoundly dislike.

Acknowledgements relating to this commentary

Assistance with the commentary: none.

Financial support and sponsorship: none.

Conflicts of interest: none.

Comment from the Editor: EDR is the chairperson, and AA and DL are members of the European Society of Anaesthesiology (ESA) Guideline Committee. DL is the chairperson of the National Anaesthesiologists' Societies Committee (NASC) of the ESA.

This article was checked and accepted by the Editors, but was not sent for external peer-review.

References

1 Marx G, *et al.*, Intravascular volume therapy in adults. Guidelines from the German society of anaesthesiology and intensive care medicine. *Eur J Anaesthesiol* 2016; **00**:000–000.
2 Cheuvront SN, Kenefick RW, Charkoudian N, *et al.* Physiologic basis for understanding quantitative dehydration assessment. *Am J Clin Nutr* 2013; **97**:455–462.
3 Lewins R, Lewins R. Injection of saline solutions in extraordinary quantities into the veins in cases of malignant cholera. *Lancet* 1832; **18**:243–244.
4 The cases of cholera successfully treated. *The Lancet* 1832; **18**:284–286.
5 Kozek-Langenecker SA, Afshari A, Albaladejo P, *et al.* Management of severe perioperative bleeding: guidelines from the European Society of Anaesthesiology. *Eur J Anaesthesiol* 2013; **30**:270–382.
6 Palumbo D, Servillo G, D'Amato L, *et al.* The effects of hydroxyethyl starch solution in critically ill patients. *Minerva Anestesiol* 2006; **72**:655–664.
7 Aldecoa C, Kozek-Langenecker S, Rico-Feijoo J. Colloids in surgery: bad drugs, bad protocol, or bad data analysis? *Minerva Anestesiol* 2014; **80**:856–857.
8 Cavallaro F, Sandroni C, Antonelli M. Functional hemodynamic monitoring and dynamic indices of fluid responsiveness. *Minerva Anestesiol* 2008; **74**:123–135.
9 Kubitz JC, Richter HP, Petersen C, *et al.* Right ventricular stroke volume variation: a tool to assess right ventricular volume responsiveness. *Minerva Anestesiol* 2014; **80**:992–995.
10 Ioannidis JPA. Why most published research findings are false. *PLoS Med* 2005; **2**:e124.
11 Chappell D, Jacob M. Role of the glycocalyx in fluid management: small things matter. *Best Pract Res Clin Anaesthesiol* 2014; **28**:227–234.
12 Michel CC, Phillips ME. Steady-state fluid filtration at different capillary pressures in perfused frog mesenteric capillaries. *J Physiol* 1987; **388**:421–435.
13 Weinbaum S. 1997 Whitaker Distinguished Lecture: models to solve mysteries in biomechanics at the cellular level; a new view of fiber matrix layers. *Ann Biomed Eng* 1998; **26**:627–643.
14 Mehta D, Ravindran K, Kuebler WM. Novel regulators of endothelial barrier function. *Am J Physiol Lung Cell Mol Physiol* 2014; **307**:L924–935.
15 Jacob M, Saller T, Chappell D, *et al.* Physiological levels of A-, B- and C-type natriuretic peptide shed the endothelial glycocalyx and enhance vascular permeability. *Basic Res Cardiol* 2013; **108**:347.
16 Jouan-Hureaux V, Audonnet-Blaise S, Lacatusu D, *et al.* Effects of a new perfluorocarbon emulsion on human plasma and whole-blood viscosity in the presence of albumin, hydroxyethyl starch, or modified fluid gelatin: an in vitro rheologic approach. *Transfusion* 2006; **46**:1892–1898.
17 Neff TA, Fischler L, Mark M, *et al.* The influence of two different hydroxyethyl starch solutions (6% HES 130/0.4 and 200/0.5) on blood viscosity. *Anesth Analg* 2005; **100**:1773–1780.
18 Guerci P, Tran N, Menu P, *et al.* Impact of fluid resuscitation with hypertonic-hydroxyethyl starch versus lactated ringer on hemorheology and microcirculation in hemorrhagic shock. *Clin Hemorheol Microcirc* 2014; **56**:301–317.
19 Higgins JPT, Altman DG, Gøtzsche PC, *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011; **343**:d5928.
20 Hróbjartsson A, Boutron I, Turner L, *et al.* Assessing risk of bias in randomised clinical trials included in Cochrane Reviews: the why is easy, the how is a challenge. *Cochrane Database Syst Rev* 2013; **4**:ED000058.
21 Moher D, Cook DJ, Eastwood S, *et al.* Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999; **354**:1896–1900.
22 Moher D, Liberati A, Tetzlaff J, *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 2010; **8**:336–341.
23 Panic N, Leoncini E, de Belvis G, *et al.* Evaluation of the endorsement of the preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. *PLoS ONE* 2013; **8**:e83138.
24 Tao K-m, Li X-q, Zhou Q-h, *et al.* From QUOROM to PRISMA: a survey of high-impact medical journals' instructions to authors and a review of systematic reviews in anesthesia literature. *PLoS ONE* 2011; **6**:e27611.
25 Haug CJ. Peer-Review Fraud: hacking the scientific publication process. *N Engl J Med* 2015; **373**:2393–2395.
26 Brouwers MC, Kho ME, Browman GP, *et al.* AGREE II: advancing guideline development, reporting and evaluation in healthcare. *CMAJ* 2010; **182**:E839–E842.
27 Guyatt GH, Oxman AD, Vist GE, *et al.* GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; **336**:924–926.
28 Swingler GH, Volmink J, Ioannidis JPA. Number of published systematic reviews and global burden of disease: database analysis. *BMJ* 2003; **327**:1083–1084.
29 Tunis AS, McInnes MDF, Hanna R, *et al.* Association of study quality with completeness of reporting: have completeness of reporting and quality of systematic reviews and meta-analyses in major radiology journals changed since publication of the PRISMA statement? *Radiology* 2013; **269**:413–426.
30 Hedges LV, Pigott TD. The power of statistical tests in meta-analysis. *Psychol Methods* 2001; **6**:203–217.
31 Thorlund K, Devereaux PJ, Wetterslev J, *et al.* Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? *Int J Epidemiol* 2009; **38**:276–286.
32 Imberger G, Vejlby AD, Hansen SB, *et al.* Statistical multiplicity in systematic reviews of anaesthesia interventions: a quantification and comparison between Cochrane and non-Cochrane reviews. *PLoS ONE* 2011; **6**:e28422.
33 Higgins JPT, Whitehead A, Simmonds M. Sequential methods for random-effects meta-analysis. *Stat Med* 2011; **30**:903–921.
34 Brok J, Thorlund K, Gluud C, Wetterslev J. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *J Clin Epidemiol* 2008; **61**:763–769.
35 Borm GF, Donders AR. Updating meta-analyses leads to larger type I errors than publication bias. *J Clin Epidemiol* 2009; **62**:825–830.
36 Thorlund K, Imberger G, Walsh M, *et al.* The number of patients and events required to limit the risk of overestimation of intervention effects in meta-analysis: a simulation study. *PLoS ONE* 2011; **6**:e25491.
37 Turner RM, Bird SM, Higgins JPT. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. *PLoS ONE* 2013; **8**:e59202.
38 Afshari A, Wetterslev J. When may systematic reviews and meta-analyses be considered reliable? *Eur J Anaesthesiol* 2015; **32**:85–87.
39 Pereira TV, Ioannidis JPA. Statistically significant meta-analyses of clinical trials have modest credibility and inflated effects. *J Clin Epidemiol* 2011; **64**:1060–1069.