**Report from the BMJ’s manuscript committee meeting**

These comments are an attempt to summarise the discussions at the manuscript meeting. They are not an exact transcript.

Manuscript meeting 26.03.2015

Elizabeth Loder (chair), Julie Morris (stats), Kristina Fister, Georg Roggla, Alison Tonks, Wim Weber, Tiago Villanueva, Rebecca Burch, Rubin Minhas, 2 guests

Decision: Ask for revision

The committee was interested in the topic of your research. The following concerns were mentioned:

• The committee shared the reviewers concerns.

The distinguished group of reviewers selected by the Journal has provided very helpful feedback on our work. We have tried to address the reviewers’ concerns, and we feel the manuscript is much improved as a result.

• The major issue as with all database studies is the inability to determine causality.

Unfortunately, all retrospective cohort studies are subject to unmeasured confounders, and research using large datasets collected for non-research purposes, such as in this study, is no different. However, we have performed a rigorous analysis of these data, and our findings are novel, biologically plausible, and clinically relevant; given the rarity of perioperative stroke and MI, it is unlikely that any other study design could be used to demonstrate this association. We have tried to present our findings in a light appropriately tempered with restraint because of the work’s inherent methodological limitations.

• What does your paper add to previous publications on this topic?

Reviewer 1 asked this question as well, to which we responded below; our reply is reproduced here:

The novelty of our findings lies in the demonstration of an effect at small (e.g., 1 unit) volumes of transfusion. While this has been shown in the percutaneous coronary intervention population, it has not yet been demonstrated in a perioperative (surgical)
cohort. However, given the prevalence of elective, scheduled surgeries, we as physicians have the opportunity to intervene upon a major modifiable risk factor for perioperative transfusion (i.e., correction of preexisting anemia). If intraoperative transfusions contribute to increased risk of perioperative complications, diligent preoperative treatment of anemia affords a low-risk opportunity to improve perioperative safety and reduce postoperative morbidity. We argue that this is an issue of great importance to any clinician who is positioned to provide preoperative risk stratification or optimization, which includes surgeons, anesthesiologists, primary care physicians, and hospitalists, among others.

• Please provide a clear and robust message for general readers.
• Isn't it just as possible that the need for more units of pRBCs is a marker for underlying surgical or medical complexity and it's that factor increases the risk of these outcomes?

We know of no way to evaluate whether it is “just as possible” that pRBC transfusion is a marker for patient complexity. Although we have attempted to adjust for known risk factors for stroke or MI, unfortunately, our study design does not allow us to state that there is no possibility of residual confounding (as with all retrospective cohort studies). We have attempted to mitigate this limitation with extensive subgroup and sensitivity analyses, and we also attempt to address it as thoroughly as possible in the discussion.

• Please be specific about the types of stroke.

We apologize for being unclear. Under “Outcome Definition” in the Methods section, we list the ICD-9 codes used for coding of our “stroke” variable. We also specify that these codes do not include hemorrhagic stroke or subarachnoid hemorrhage. To provide additional clarity, we have specified at several points in the abstract and manuscript that we refer only to ischemic stroke.

First and foremost, please revise your paper to respond to all of the comments by the reviewers. Their reports are available below. Please also respond to the additional comments by the committee.

IMPORTANT
When you revise and return your manuscript, please take note of all the following points. Even if an item, such as a competing interests statement, was present and correct in the original draft of your paper, please check that it has not slipped out during revision.

a. In your response to the reviewers and committee please provide, point by point, your replies to the comments made by the reviewers and the editors, and please explain how you have dealt with them in the paper. It may not be possible to respond in detail to all these points in the paper itself, so please do so in the box provided.

b. If your article is accepted it will then be edited, proofed, and - after your approval - published on bmj.com with open access. This open access Online First article will not be a pre-print. It will represent the full, citable, publication of that article. The citation will be year, volume, elocator (a unique identifier for that article): eg BMJ 2008;337:a145 — and this is what will appear immediately in Medline, PubMed, and other bibliographical
indexes. We will give this citation in print and online, and you will need to use it when you cite your article.

c. Please write an abridged version of the article for the print and iPad BMJ using the appropriate BMJ pico template for your study's design. Please be reassured that it doesn't take long to complete this. When your BMJ pico is ready please email it to papersadmin@bmjgroup.com. The templates for you to download are at http://resources.bmj.com/bmj/authors/bmj-pico

Please include the items below in the revised manuscript to comply with BMJ style:

* the title of the article should include the study design eg "a retrospective analysis of hospital episode statistics"

We have revised the study design to read “a retrospective population-based analysis”.

* ID of ethics committee approval and name of the ethics committee/IRB; or a statement that approval was not required (see ttp://resources.bmj.com/bmj/authors/editorial-policies/guidelines)

Checked and present.

* Please complete the following statement and add it to your manuscript:
"All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work OR [author initials] had support from [name of organisation] for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years OR [author initials] [had specified relationship] with [name of organisation] in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work OR [initials of relevant authors] [had specified relationships or activities of this type]"

The statement was revised to include the parenthetical phrase regarding availability of ICMJE forms from the corresponding author.

Guidance and examples can be found at http://resources.bmj.com/bmj/authors/editorial-policies/competing-interests

Each author should download and complete a copy of the disclosure form www.icmje.org/coi_disclosure.pdf.
They should keep a copy of the form and send a copy to the corresponding author to be filed and, if necessary, retrieved later. They do NOT need to send a copy to the BMJ.

Please see http://resources.bmj.com/bmj/authors/editorial-policies/competing-interests.

* contributorship statement+ guarantor
(see http://resources.bmj.com/bmj/authors/article-submission/authorship-contributorship)

Checked and present.
* copyright statement/ licence for publication  (see  http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse)

Checked and present.

* signed patient consent form(s), if the article gives enough personal information about any patient(s): this sometimes occurs even in research papers - for example in a table giving demographic and clinical information about a small subgroup in a trial or observational study, or in quotes/tables in a qualitative study - (see http://resources.bmj.com/bmj/authors/editorial-policies/copy_of_patient-confidentiality)

N/A

* for a clinical trial, the trial registration number and name of register – in the last line of the structured abstract

N/A

* for any other registered study (eg a systematic review), the registration number and name of register – in the last line of the structured abstract

N/A

* a data sharing statement declaring what further information and data your are willing to make available. Suggested wording: "Data sharing: technical appendix, statistical code, and dataset are available at this repository or website <url> OR from the corresponding author at <email address or url>". If there are no such further data available, please use this wording: "Data sharing: no additional data available"

Checked and present.

* please write the discussion section of your paper in a structured way, to minimise the risk of careful explanation giving way to polemic.
Please follow this structure:
* statement of principal findings of the study
* strengths and weaknesses of the study
* strengths and weaknesses in relation to other studies, discussing important differences in results and what your study adds. Whenever possible please discuss your study in the light of relevant systematic reviews and meta-analyses (eg Cochrane reviews)
* meaning of the study: possible explanations and implications for clinicians and policymakers and other researchers; how your study could promote better decisions
* unanswered questions and future research

We have structured our discussion in that way.

* please note, too, that the article’s introduction should cover no more than three paragraphs, focusing on the research question and your reasons for asking it now

Checked and in compliance.
* What this paper adds/what is already known box (as described at [http://resources.bmj.com/bmj/authors/types-of-article/research](http://resources.bmj.com/bmj/authors/types-of-article/research))

Checked and present.

* funding statement (see [http://resources.bmj.com/bmj/authors/article-submission/article-requirements](http://resources.bmj.com/bmj/authors/article-submission/article-requirements))

Checked and present.

* statement of the independence of researchers from funders (see [http://resources.bmj.com/bmj/authors/article-submission/article-requirements](http://resources.bmj.com/bmj/authors/article-submission/article-requirements))

Checked and present.

* for studies funded or sponsored by industry (see [http://resources.bmj.com/bmj/authors/article-submission/article-requirements](http://resources.bmj.com/bmj/authors/article-submission/article-requirements)):
  * a statement describing the role of the study sponsor(s), if any, in study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication
  * assurance, in the cover letter, that a clinical trial funded by a pharmaceutical or other commercial company follows the guidelines on good publication practice (see [http://resources.bmj.com/bmj/authors/article-submission/article-requirements](http://resources.bmj.com/bmj/authors/article-submission/article-requirements))
  * inclusion in the list of contributors the name(s) any professional medical writer(s), specifying in the formal funding statement for the article who paid the writer. Writers and authors must have access to relevant data while writing articles.

N/A

* structured abstract (see [http://resources.bmj.com/bmj/authors/types-of-article/research](http://resources.bmj.com/bmj/authors/types-of-article/research))

Checked and present.

* summary statistics to clarify your message
We do want your piece to be easy to read, but also want it to be as scientifically accurate as possible. Please include in the results section of your structured abstract (and, of course, in the article's results section) the following terms, as appropriate:
   For a clinical trial:
   • Absolute event rates among experimental and control groups
   • RRR (relative risk reduction)
   • NNT or NNH (number needed to treat or harm) and its 95% confidence interval (or, if the trial is of a public health intervention, number helped per 1000 or 100,000)
   For a cohort study:
   • Absolute event rates over time (eg 10 years) among exposed and non-exposed groups
   • RRR (relative risk reduction)
   For a case control study:
   • OR (odds ratio) for strength of association between exposure and outcome
   For a study of a diagnostic test:
   • Sensitivity and specificity
   • PPV and NPV (positive and negative predictive values)
Checked and present.

For research articles
As well as submitting your revised manuscript, we also require a copy of the manuscript with changes highlighted. Please upload this file with file designation ‘Revised Manuscript Marked copy’.

REFEREES COMMENTS

Reviewer: 1

Recommendation:

Comments:
The authors of this study performed a retrospective analysis of administrative data collected by Premier Inc. with the goal to determine if the perioperative transfusion of blood was associated with adverse outcome, specifically MI/stroke. They hypothesized that “transfusion of as little as 1 unit of pRBCs in the operating room or the day following would be associated with measurably increased odds for perioperative stroke and myocardial infarction (MI).” They concluded that indeed the transfusion of blood in the perioperative period was associated with increased odds for stroke and/or MI.

The topic is interesting and clinically relevant. The manuscript is well written and the methods are well described. The abstract conveys the main message.

However, there are a number of issues that I would like to see addressed by the authors:

The major issue, as with all database studies is the inability to determine causality. The authors rightfully speak of associations (and not causality), but nevertheless in the case of blood transfusions the question has to remain if the transfusion of blood is just a surrogate marker for other variables determining adverse outcome. In this context administrative databases such as Premier lack crucial clinical detail such as the amount of blood loss, hematocrits, transfusion triggers used, invasiveness of the procedure, the occurrence of interoperative surgical complications, length of operation times, hemodynamic details and level of patient comorbidity burden. Even if some information is available, it is often presented in binary form (i.e. present or not via the ICD9 coding system), thus not allowing for determination of the severity of the condition. This is not to say that because of this, all studies utilizing databases are “not useful”. Quite the opposite is the case, especially when crucial observations have not yet been reported in the literature and/or cannot be obtained from clinical trials, databases can present a first look into a subject matter that can then lead to the formation of further hypotheses. In the case of blood transfusions, however, much evidence has been published on the ill-effects of liberal transfusion practices (although as the authors suggest data specifically for MI and stroke outcomes may be less common). Thus, the ability of the results from this study to add significant knowledge is limited. For example, without knowing which of these transfusions could be considered appropriate and which not (which is impossible without knowing any laboratory values, hemodynamic parameters and other clinical detail) conclusion that can be drawn are limited. The issue of using databases such as Premier for the topic of perioperative blood transfusions is particularly challenging (see Blood transfusions in total hip and knee arthroplasty: an analysis of outcomes.Danninger
T, Rasul R, Poeran J, Stundner O, Mazumdar M, Fleischut PM, Poultsides L, Memtsoudis SG ScientificWorldJournal. 2014 Jan 21;2014:623460. In this publication, the same database was used to look into the particular problem of perioperative transfusions and various outcomes, and encountered multiple problems/limitations. While insights were gained into practice patterns, the inability to address the issue of confounding had to be acknowledged and limited interpretations. Thus the authors should present a very convincing and strong argument regarding the novelty and additional importance of their study to the field.

We greatly appreciate Dr Memtsoudis' thoughtful comments. The difficulty of determining causality using retrospective cohort studies is a well-known limitation of our study design, and we readily accept that a study such as this cannot lend causal inference. Large datasets additionally trade granularity and clinical detail for size; Premier, as a particularly large dataset, unfortunately does not collect clinical data that would otherwise have been helpful for this study, as Dr Memtsoudis points out. We thank him for drawing our attention to his recent publication (using Premier data from 2005-2010) in Scientific World Journal. Compared with his work, we do have the additional advantages of knowing the number of units of blood transfused (through the use of billing records available through Premier), as well as robust present-on-admission comorbidity documentation, but his point is nonetheless well taken: additional clinical data would strengthen our conclusions. In the setting of rare complications like stroke and MI, however, we would argue that there are few, if any, datasets large enough to answer the research question we pose while providing significantly more clinical detail. Dr Kamel's 2012 study in Circulation, which inspired the present study, used NSQIP data (with somewhat greater clinical granularity) and while his focus was on transfusion of >4 units, his findings are consistent with ours. The novelty of our findings lies in the demonstration of an effect at small (e.g., 1 unit) volumes of transfusion. While this has been shown in the PCI population, it has not yet been demonstrated in a perioperative (surgical) cohort. However, given the prevalence of elective, scheduled surgeries, we as physicians have the opportunity to intervene upon a major modifiable risk factor for perioperative transfusion (i.e., correction of preexisting anemia). If intraoperative transfusions contribute to increased risk of perioperative complications, diligent preoperative treatment of anemia affords a low-risk opportunity to improve perioperative safety and reduce postoperative morbidity. We argue that this is an issue of great importance to any clinician who is positioned to provide preoperative risk stratification or optimization, which includes surgeons, anesthesiologists, primary care physicians, and hospitalists, among others.

Other comments:

1) Having mentioned the issue regarding the lack of the ability to establish causal relationships, did the authors consider that actual effect in their model would be the surgical intervention, stroke/MI the outcome and that the transfusion would function as a moderator? Would it be feasible to employ methodological methods to take this relationship into account?

Propensity modeling takes advantage of the “natural experiment” in which factors which may sometimes lead to transfusion – e.g., preexisting anemia, CAD, or particular surgical procedures – can be collectively estimated and patients matched on the basis of actually receiving (or not receiving) a transfusion, given equal probability of being
transfused. This approach does consider transfusion an intermediate step in the relationship between surgery and stroke/MI. We feel the propensity score analysis is very complementary to our hypothesis that transfusion is a potentially modifiable factor in the causal pathway between surgery and perioperative stroke/MI. However, a propensity score model has limitations as well, as we discuss further in response to Reviewer 3’s question below.

2) The authors present a possible pathophysiologic process by which transfusions may lead to adverse outcomes. This should be lauded and is important. In the case of MI/stroke and other cardiovascular problems however, other mechanisms are equally plausible. Anemia and resulting hypotension may be additional factors, which in this case may be the trigger for blood transfusions. This should be more thoroughly discussed.

There is no way to know what the specific indication was for any given transfusion described in our dataset. Dr Kamel, in his paper, advanced the hypothesis that transfusion of 4 or more units of blood represented is more likely to represent massive hemorrhage, a hypothesis with which we agree. At the other extreme, we performed subgroup analysis of patients receiving only one unit of blood, which we believe is less likely to represent transfusions given in the setting of hemodynamically unstable hemorrhage, and attempted to contextualize our findings using the recent study of stroke and MI following transfusion after PCI. We have added another sentence in the Limitations section which explicitly sets forth this caveat.

3) Did the authors consider evaluating the use of blood conservation methods, like cell saver, the use of colloids, use of anticoagulants/platelet agents and the utilization of resources such as ICUs? All these factors may affect the outcome or choice to transfuse PRBCs and are available in Premier.

We acknowledge that there are many variables in Premier that we could have investigated, or used in our models. However, we favored a more parsimonious approach, and tried to use only variables about which we had strong hypotheses and which would be relatively less confounded by clinical indications. Adjustment factors were based on a literature search for risk factors for stroke and MI. Accordingly, we did not obtain granular data on crystalloid, colloid, or CellSaver use, or ICU admission.

Premier has information only on in-hospital anticoagulants and antiplatelet agents, and we hypothesized that the decision to institute or restart any anticoagulant or antiplatelet would be highly confounded by the indication and patient’s perceived risk, the surgery, whether intra- or postoperative bleeding was greater than expected, and institutional practices. We had an extremely limited ability to adjust for those factors, and accordingly did not attempt to use anticoagulant data in our model. It would be very interesting to adjust for duration of anticoagulant hiatus prior to surgery, because practice patterns vary widely depending on indication and patient’s perceived risk, but we do not have those data.

4) The authors present in their methodology an attempt to control for hospital level variables, but it remains unclear if they just include hospital type information or actually performed a multilevel analysis accounting for each individual hospital, which may take into account differences in local practice patterns.
We apologize for the lack of clarity. All presented analyses are multilevel (hierarchical) models, adjusting for random effects by hospital. Instead of using the term “mixed effects” to describe the regression model used throughout, we now use “hierarchical logistic regression” in an attempt to improve clarity.

5) While the attempt to control for present on admission comorbidities is laudable, from my experience this variable has a high missingness in Premier. Can the authors comment on this and report missingness and how this issue was handled in general?

POA coding was, in fact, remarkably complete in the data from the years we used. Of 1,583,819 patients, 1,557,520 (98.3%) had at least one ICD-9 diagnosis explicitly flagged as “present-on-admission”. An additional 1,476 patients (0.09%) had at least one ICD-9 diagnosis explicitly flagged as “not-present-on-admission” but no “present-on-admission” flags, for a total of 98.4% of patients included in the study that had explicit documentation of at least one condition present or not present on admission. This leaves 26,299 patients, of whom 1,632 had codes for “exempt” POA reporting, and 24,667 without explicit characterization of any ICD-9 diagnosis as present, non-present, or exempt (1.56%). The difference between Dr Memtsoudis’ experience and our own likely reflects the evolution of Centers for Medicare and Medicaid Services (CMS) policy in the late 2000s regarding reimbursement for hospital-acquired conditions, which essentially incentivized hospitals to pursue thorough POA reporting.

6) What was the reason that the authors included stroke, MI and VF/VF in the composite variables. There is no right or wrong, but I think an explanation of this rational would help the reader understand. It is likely that by themselves these outcomes were to infrequent to allow for appropriate modeling, but one could ask why not other cardiovascular complications were also included.

(This was a question asked by Reviewer 2 as well, and we have reproduced elements of our response for both reviewers.)

VT/VF was included as a surrogate for coronary ischemia, which may be so rapidly lethal as to prevent diagnosis of myocardial ischemia or may be transient (and thus not result in a myocardial infarction identifiable on autopsy). This was nicely reviewed in Myerburg and Junttila, Circulation 2012, to which we now refer in the Methods section by way of justification, as the reviewer recommends. This justification is briefly restated in the Abstract and at the beginning of the Discussion as well.

Because of this link, we planned to include VT/VF in the composite outcome before any data exploration took place (i.e., prespecified). After the hierarchical logistic regression model had been fully developed, we did confirm that the VT/VF population “behaves” similarly to the stroke and MI population; not surprisingly, a history of CAD/MI also confers elevated risk for VT/VF, though not as strongly as it does for MI (reflecting the acknowledged heterogeneity of pathways which may lead to VT/VF). In response to the reviewer’s question, we ran the primary hierarchical logistic regression model on the population of patients who suffered either a stroke or MI, but not VT/VF (n = 3,755).

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBC transfusion</td>
<td></td>
</tr>
<tr>
<td>1 unit</td>
<td>1.83 (1.34-2.48)</td>
</tr>
<tr>
<td>2 units</td>
<td>1.98 (1.55-2.53)</td>
</tr>
<tr>
<td>3 units</td>
<td>2.99 (1.90-4.70)</td>
</tr>
</tbody>
</table>
These findings are consistent with the results of the primary model and Appendix 1 composite outcome subgroups, although the confidence intervals are somewhat wider.

Because our work was inspired and contextualized by that of Dr Kamel (Circulation 2012; ref 2 in our paper), which focused on stroke and MI as outcomes, we did not include any other cardiovascular complications (e.g., atrial fibrillation).

Appendix 1 presents the results of the primary hierarchical regression model looking at each component of the composite outcome separately; although our planned primary analysis was that of the composite outcome, the effect was sufficiently strong (and the dataset sufficiently large) that individual outcomes could have been examined with adequate power. However, as this was not our intended primary analysis, results are presented as an Appendix only.

7) The authors should be lauded for using sophisticated methodological, including sensitivity analyses and methods to reduce the chance of residual cofounding. 8) Looking at the tables, it becomes obvious that the presence of cardiovascular diseases are amongst the variables with the highest OR for adverse outcomes (higher than that for blood transfusions). Again this puts into question if it is not really the underlying disease that prompts clinicians to be more liberal in their transfusion practice to keep Hgb levels higher and ultimately drives the morbidity aspect. Would calculating attributable risk be of benefit? Of course this would also be limited but could shed some more light on this issue.

We calculated raw and adjusted population attributable fraction; the former using the standard equation and substituting the adjusted odds ratio from our hierarchical logistic regression, and the latter with the Stata command *punaf*, using a simplified (clustered) logit model because the hierarchical model and interaction terms cannot be used with *punaf*. Unadjusted population attributable fraction was 3.37%. Covariate-adjusted estimates (i.e. using *punaf*) are below and are now reported in the Results.

<table>
<thead>
<tr>
<th>Variable</th>
<th>PAF</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRBC transfusion</td>
<td>2.4%</td>
<td>1.8 - 3.0%</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
<td>9.7%</td>
<td>8.6 - 10.7%</td>
</tr>
<tr>
<td>CAD history</td>
<td>20.2%</td>
<td>18.5 - 22.0%</td>
</tr>
<tr>
<td>All comorbidity covariates (i.e., CVD, CAD,</td>
<td>43.3%</td>
<td>41.1 - 45.4%</td>
</tr>
<tr>
<td>“high-risk” composite, obesity, anemia, smoking)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We are not surprised that the attributable fraction for transfusion is quite small. It makes sense that, without a background of vulnerability (e.g. CAD or cerebrovascular disease), a single transfusion is unlikely to overcome compensatory anti-inflammatory or other mechanisms. Furthermore, the NNT calculated based on the propensity score model suggests that, assuming a causal relationship, over 250 transfusions must occur to cause one additional stroke or MI. In the Discussion, we put this finding in context by pointing out that, while the attributable risk is much smaller than that for specific medical comorbidities, pRBC transfusion is (unlike comorbidity) potentially modifiable.
We thank Dr Memtsoudis for suggesting this helpful analytic approach, which we think complements other findings discussed in the manuscript.

9) Although the authors attempted to control for procedure type in their analysis, it must be noted that within those categories major differences determining morbidity may exist. For example spine surgery can span the spectrum between laminectomies to extensive scoliosis correction surgery with massive blood loss.

Even subgroups can be quite heterogeneous, as Dr Memtsoudis points out. ICD-9 codes only provide limited granularity on which to stratify these groups. We could subdivide the “spine” group further, but would have to accept the loss of power (and the statistical implications) inherent in exploring that group further. If it is helpful for the reviewer in interpreting this specific group, of the “spine” patients, 153,738 underwent codes for spinal fusion and 36,184 underwent laminectomy without fusion. The relatively lower prevalence of laminectomy without fusion likely reflects the practice of performing these as outpatient procedures. We would expect this to bias our findings toward the null. But, fundamentally, our inability to adjust for highly granular clinical variables such as estimated blood loss remains a limitation of the work.

Additional Questions:
Please enter your name: Stavros Memtsoudis

Job Title: Clinical Professor of Anesthesiology and Public Health/ Senior Scientist

Institution: Cornell/ HSS

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF='http://www.bmj.com/sites/default/files/attachments/resources/2011/07/bmipolicyondeclarationofinterestsmarch2014.pdf' target='_new'>(please see BMJ policy)</a> please declare them here:
Reviewer: 2

Recommendation:

Comments:

General comments:

The authors report a retrospective cohort study on the association between perioperative transfusion of packed red blood cells (pRBC) and stroke or myocardial infarction (MI). Using data from a voluntary healthcare benchmarking database capturing approximately 20% of U.S. hospitalizations, the authors used time-stamped lab/procedure claims and ICD-9-CM diagnosis codes to ascertain pRBC transfusion and the outcomes of interest. They found a substantial association between pRBC transfusion and the risk of stroke and MI. This is a very well-written manuscript detailing a thoughtful analysis on an important clinical topic. Below are a few specific comments that may help to improve this paper.

Specific comments:

1. Page 7, line 9 states that a subgroup analysis was performed limited to patients who did not receive transfusion on postoperative days 1-7, but on page Page 5, line 15 it states that all patients undergoing transfusion on postoperative days 2-7 were excluded. Could the authors please clarify this? Do they simply mean that the primary analysis excluded those transfused from day 2-7, and the subgroup analysis additionally excluded those transfused on postoperative day 1? Also, what about patients with transfusion beyond postoperative day 7? I realize this would be rare, but it would be helpful to clarify.

We apologize for the confusion and have attempted to clarify the phrasing on page 7: the reviewer is correct that the subgroup analysis additionally excluded those transfused on postoperative day 1.

Although we did not extract data regarding transfusions beyond postoperative day 7, we here present a histogram of day-of-first-transfusion (within 7 days) in the study population before exclusion of patients transfused on days 2 or beyond. In this graph, transfuseday is postoperative day – as in, transfuseday 0 is day-of-surgery, transfuseday 1 is postoperative day 1, etc.
Very few patients were first transfused on postoperative days 5-7. We suspect even fewer would have been first transfused on day 8 or beyond. We also speculate that transfusions so temporally distant from surgery are less likely due to surgical blood loss (particularly the modifiable contributor of preoperative anemia) than other hospital-related factors like incident complications, recurrent phlebotomy, etc.

2. The statistical approach is nicely done and very clearly explained. I just have a few questions:

a. Why not model the number of pRBCs as a continuous variable? In our paper (reference 2 in your paper), besides the primary finding of increased stroke/MI risk with hemorrhage requiring >4 units of pRBCs, in a secondary analysis we found a nicely linear relationship between the overall number of pRBCs transfused and the risk of stroke/MI.

We wish to take this opportunity to thank Dr Kamel for his excellent paper, which inspired this work. We did not model pRBCs as a continuous variable because the distribution’s extremely long tail (i.e., the very small number of patients transfused more than, e.g., 10 units) suggested the variable would be better treated as an ordinal. Our data also suggest that, in this population, the relationship between stroke/MI risk and pRBCs transfused is not linear; the OR for 1 unit is statistically indistinguishable from that for 2 units. We certainly do agree with Dr Kamel that there is a dose-response relationship demonstrated by these data, which he showed in his work with the linear relationship to which he refers.

b. Why meqrlogit instead of xtmelogit?

Stata 13 renamed xtmelogit (as it was in Stata 12) to meqrlogit.
3. Page 10, line 49: Just to be completely correct, would add something like “Although not perfectly population-representative, these findings …”

We have qualified our statement: “Furthermore, although it is not perfectly representative of the United States population, the dataset we used…”. We do feel that the results are very likely to be generalizable, despite this limitation.

4. The inclusion of VT/VF in the composite outcome distracts from the overall message of the paper. At the very least, there should be an explanation of the rationale for including it, and it should be more consistently referred to (e.g., the the Results of the Abstract refer only to stroke/MI, as does Table 4). I would favor removing VT/VF altogether, or including it as a secondary outcome.

(This was a question asked by Reviewer 1 as well, and we have reproduced elements of our response for both reviewers.)

VT/VF was included as a surrogate for coronary ischemia, which may be so rapidly lethal as to prevent diagnosis of myocardial ischemia or may be transient (and thus not result in a myocardial infarction identifiable on autopsy). This was nicely reviewed in Myerburg and Junttila, Circulation 2012, to which we now refer in the Methods section by way of justification, as the reviewer recommends. This justification is briefly restated in the Abstract and at the beginning of the Discussion as well.

Because of this link, we planned to include VT/VF in the composite outcome before any data exploration took place (i.e., prespecified). After the hierarchical logistic regression model had been fully developed, we did confirm that the VT/VF population “behaves” similarly to the stroke and MI population; not surprisingly, a history of CAD/MI also confers elevated risk for VT/VF, though not as strongly as it does for MI (reflecting the acknowledged heterogeneity of pathways which may lead to VT/VF). In response to the reviewer’s question, we ran the primary hierarchical logistic regression model to estimate OR for the outcome of either a stroke or MI, but not VT/VF (n = 3,755).

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBC transfusion</td>
<td></td>
</tr>
<tr>
<td>1 unit</td>
<td>1.83 (1.34-2.48)</td>
</tr>
<tr>
<td>2 units</td>
<td>1.98 (1.55-2.53)</td>
</tr>
<tr>
<td>3 units</td>
<td>2.99 (1.90-4.70)</td>
</tr>
<tr>
<td>4+ units</td>
<td>3.98 (2.78-5.71)</td>
</tr>
<tr>
<td>CVD</td>
<td>17.4 (16.0-19.1)</td>
</tr>
<tr>
<td>CAD/MI hx</td>
<td>1.74 (1.61-1.88)</td>
</tr>
</tbody>
</table>

These findings are in line with the results of the primary model and Appendix 1 composite outcome subgroups.

5. I agree with the authors that a substantial portion of the association between transfusion and stroke/MI is likely to be causal. But it should be acknowledged that it is not possible to completely tease apart the deleterious effect of transfusion from the deleterious effect of bleeding. Even a single-unit transfusion may reflect surgical oozing that may activate innate hemostatic mechanisms. The data cited on transfusion after PCI is compelling, but it is not in a directly comparable population since one would expect
generally more blood loss with surgery than with PCI. I think it would make the paper stronger to mention this in the limitations section.

This is a nuanced point. Furthermore, the inflammatory response resulting from tissue damage may also be expected to activate hemostatic mechanisms. We have briefly addressed this in our Limitations section as the reviewer recommends; the pathophysiology is extremely complex.

6. Conclusion: I think it is fine to call for more randomized trials, but given existing randomized trial results (refs 18 and 19), I think a somewhat stronger claim for transfusion-sparing strategies can be made. Something to the effect of, “Our findings of rare but serious adverse outcomes, combined with several randomized trials showing the safety of transfusion-sparing approaches in critical care and postoperative medicine, argue for …”

Your phrasing is very thoughtful; we have adapted it and additionally cited a recent meta-analysis published in BMJ which supports the safety of a restrictive approach.

Additional Questions:
Please enter your name: Hooman Kamel

Job Title: Assistant Professor

Institution: Weill Cornell Medical College

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF="http://www.bmj.com/sites/default/files/attachments/resources/2011/07/bmjpolicydeclarationofinterestsmarch2014.pdf"target='_new'>(please see BMJ policy)</a> please declare them here:

Reviewer: 3
Recommendation:

Comments:
The authors have performed a thoughtful and detailed analysis of an important question -- does discretionary day-of-surgery transfusion have an impact. By excluding patients that receive 5+ units over the hospital admission, or receive PRBCs on later postoperative days, the impact of procedure-related anemia transfusion is more clearly identified.

Methods
- there is a multitude of testing performed. most of it appears to have been predefined. was a protocol submitted or registered that documents the various and many subgroup and sensitivity analyses

Our a priori analytic approach was to develop our primary regression model, then test the model’s robustness in key clinical subsets where bleeding or transfusion risks were likely to differ because of surgery type as well as in subsets that tested the impact of our primary outcome construction (based on the number, type, or timing of transfusion given). We also decided a priori to use propensity score analysis to further limit the risk of allocation biases; these analyses proceeded along standard methodologic lines and included adjusting and matching steps. However, no formal protocol was registered; we have thus revised our phrasing in the manuscript.

- it is unclear to me whether all documented discharge diagnoses were available for the elixhauser comorbidity scoring -- how many diagnoses were the min/max/median?. the methods suggest primary and secondary were used. unclear if secondary includes comorbidity codes as opposed to pure "secondary" codes

POA coding was based on both primary and secondary diagnoses. Elixhauser comorbidity codes were derived from ICD-9 secondary diagnosis codes. Our dataset included up to 25 diagnosis codes for each patient. The median number of POA diagnoses was 5 [interquartile range 3-8]. 80% of patients had between 2 and 11 diagnoses. A histogram of number of POA diagnoses per patient (i.e, flagged with “present on admission” flag – this does not include diagnoses with the “not-present-on-admission” flag) is below. We now briefly describe the distribution of POA codes in the Methods section of the manuscript.
- why is the non-mixed model logistic regression presented or used at all? we know that there is clustering of patients at the hospital level and that fixed hospital covariates do not incorporate all of the unmeasured hospital-level effects. The authors are to be commended for using a mixed effects model, but I would recommend for simplicity and readability, that the mixed effects model, with hospital as a random effect, is the primary analyses presented. although it is described as the "primary logistic regression model", the methods section seems to describe a very iterative process where different pieces are added/removed. I presume the authors did not retain variables that were predictors in the standard model but eliminated in the mixed effects. However, the "manual iterative forward technique" leaves it unclear if the same thresholds were applied through all phases (fixed and mixed).

Because of the computational intensity of the hierarchical model, initially the model was developed as a fixed logistic regression until covariates were finalized according to the ROC criteria described. Interaction terms were then explored and added to the fixed logistic regression if criteria were met. After we were happy with the performance of that model, a hierarchical mixed-effects component, addressing random effects by hospital, was finally added and retained according to the criteria described in the methods section.

All results presented are from said hierarchical logistic regression, which we have attempted to clarify in response to Reviewer 1’s comment. The fixed logistic regression model was only an intermediate step in the development of our primary model.

As Dr Kheterpal points out, development of the models was according to an iterative process, whereby covariates were finalized, then interaction terms, then random effects
by hospital. Covariates were retained according to a ROC criterion, interaction terms according to statistical significance (p<0.05, used throughout the manuscript except where noted) in the model, and the hierarchical component according to a change in odds ratios of >10%. The model was developed in conjunction with a biostatistician who is acknowledged; model development and refinement according to these methods was under her guidance. No “backward” iteration occurred (i.e., once it was “in the model” according to criteria above, no predictor nor interaction term was removed). The derivation of the model was complex and we hope that detailed description in the methods gives an accurate idea of what was done. We are very happy to revise our description if any of the reviewers feel there are specific areas that are confusing or misleading.

- correct "compliment" to "complement"

Thank you; we have done so.

- the propensity score analysis is concerning given how poor the ROC for that regression is.

It is certainly true that a propensity score model is only as good as its predictors. We “forced” all covariates into the model but still found relatively limited predictive power. We have updated the propensity score analysis with a composite maximum surgical blood ordering schedule (MSBOS), the details of which are now described in the Methods section. The MSBOS is an institution-specific index of surgical procedures and the maximum expected number of units of pRBCs that would typically be ordered. It is developed by clinical consensus within the institution to assist with management of blood bank resources and does not consider the patient’s underlying comorbidities, surgeon characteristics or any other variable besides the description of the surgical procedure itself.

During the process of designing this study, we felt a MSBOS was a logical choice to adjust for estimated blood loss by surgical procedure since patient comorbidities are already used in the propensity model. However, it proved to be extremely difficult to find a complete MSBOS indexed by ICD-9 code. Research groups in the United States with whom we spoke all responded that their MSBOS is indexed by proprietary codes used by individual EMR systems and could not be linked to ICD-9 codes. This is also the case at UCSF, and because the EMR codes are indexed many-to-many with CCS and/or ICD-9 codes, we were unable to perform any automated linkage between our institutional MSBOS and ICD-9 codes. We abandoned the attempt, and used subgroup analysis (i.e., 1 unit vs 0 units only, day-of-surgery transfusion only) to provide an additional analytic angle on the issue.

As a result of Dr Kheterpal’s comment, we redoubled our efforts to obtain a MSBOS, expanding our efforts to researchers outside the US. Drs Seungok Lee and Yonggoo Kim partially indexed a MSBOS by ICD-9 code in 2008 and published it in the Korean Journal of Blood Transfusion; this is the only published instance of a MSBOS indexed by ICD-9 code that we were able to locate. 798,436 patients (50.4%) undergoing 59 procedures were categorized according to the MSBOS published by Drs Lee and Kim. We then manually matched UCSF MSBOS values to ICD-9 codes for procedures with a frequency of 1000 or greater in the dataset (627,791 patients; 150 procedures). This left
157,592 patients (10.0%) undergoing 1512 distinct procedures unmatched to a MSBOS value.

The ROC AUC of our propensity model improved from 0.688 to 0.715, though our results are largely unchanged. A followup propensity score sensitivity analysis eliminating patients without a matched MSBOS value (i.e., including only patients with a known MSBOS value) achieved a ROC AUC of 0.722, again with largely unchanged results, which are reported in Appendix 2. We now use this MSBOS in the reported propensity score model despite its motley pedigree since it is an intuitively important component and we do not believe it introduces inappropriate bias, based on the consistency of our results in the sensitivity analyses shown in Appendix 2.

However, while the risk of an unmeasured (and therefore unadjusted) confounder cannot be understated, the improved though imperfect propensity model may also reflect idiosyncratic differences in transfusion practice between patients – i.e., the “natural experiment” upon which propensity scores rely.

Results
- the term "packed red blood cells" should not be introduced in the results section. either use pRBC throughout, or stay with RBC. it makes the reader wonder

We are sorry for the confusion, and have tried to standardize our terms throughout the manuscript: pRBC is now defined twice (once in the Abstract, once in the Introduction) and used throughout. We have eliminated the term RBC.

- the obesity % in table 1 seem very underreported. is this purely ICD9 data, or are actual BMIs used?

Unfortunately, Premier does not contain BMI data. Obesity coding therefore relies on ICD-9 codes, and is underreported (as you note). However, we do not have an alternative metric for obesity; although the Elixhauser method is validated and widely used, we have attempted to address the limitations of ICD9 coding in the Discussion.

- is table 2 the mixed effects model results, or standard fixed effects logistic regression only?

The hierarchical (mixed effects) model is the only model reported. Apologies for the confusion; we hope to have clarified by standardizing the way we refer to the model throughout the manuscript.

- I think univariate odds ratios are not helpful in table 2 and make it harder to read. thank you for the transparency, but it comes at the expense of readability

We have removed the column of univariate odds ratios from the table.

- table 2: by orthopedic, you mean joint replacement, correct?

Table 2 reflects all surgeries in the dataset, based on ICD-9 code with selected exclusions for minor procedures (e.g., closed reduction of fracture) – for orthopedic surgery, this includes joint replacement as well as any other operation on bones or joint structures, ORIF, etc. The later subgroup analysis of joint replacement procedures
further restricts this group to primary or revision hip or knee replacement, again by ICD-9 code.

- the title of the paper and the data presented are inconsistent. the title is regarding "single unit" while patients up to 4 units are included. maybe discretionary transfusion is a better title?

We fully appreciate your point; however, we have struggled to come up with a title that highlights the most novel finding (i.e., the risk associated with single-unit transfusion), and yet does not inadvertently provide commentary on the rationale behind administering the transfusion (which we are unable to address with our data). We had considered "small-volume" and "low-volume", but as there is no accepted definition for a "small-volume" transfusion, this term was thought to be unclear. We are happy to consider title suggestions from the reviewers or the editorial team which may better reflect our findings!

Discussion
- I think the discussion is divergent from the actual data presented and their unique value. instead of summarizing the known literature, I think the authors should be more focused on their data: that even 1 unit of PRBC confers risk. the dose dependent relationship they have demonstrated is also impressive. The authors should focus on the "discretionary" transfusions aspect of their data.

Although we agree with Dr Kheterpal’s hypothesis that many of the single-unit transfusions might have been avoidable (either through tolerance of a lower hematocrit or preoperative optimization), we do not have the data to label them “discretionary”. Furthermore, we remain intrigued by the prevalence of 2-unit transfusions which, at about 50%, was greater than the prevalence of 1-unit transfusions (30%). The odds ratio for a 2-unit transfusion was, however, statistically indistinguishable from the OR for one unit, so we cannot approach the issue of the “discretionary” unit from that angle. In the context of the dogma that “if you’re going to transfuse, give 2 units” which has now been widely discredited, we have a strong suspicion that at least the second unit of the many 2-unit transfusions was in fact not clinically indicated – but, in the absence of clinical data, this unfortunately remains speculative.

- other limitations include the absence of hemodynamic data. it remains unclear whether the alternatives to transfusion (crystalloid?) are preferable. the absence of any clinical data prevents us from knowing that the patients had similar hemodynamic profiles prior to transfusion

We have now explicitly referred to this limitation – absence of hemodynamic data – in the Discussion. Unfortunately, our work cannot shed light on the question of whether crystalloid resuscitation in combination with restrictive use of pRBC transfusions is preferable in the setting of anemia with hemodynamic instability.

Additional Questions:
Please enter your name: Sachin Kheterpal

Job Title: Associate Professor of Anesthesiology

Institution: University of Michigan
Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF="http://www.bmj.com/sites/default/files/attachments/resources/2011/07/bmjpolicyondeclarationofinterestsmarch2014.pdf" target='_new'>(please see BMJ policy)</a> please declare them here: None. Sponsored research funds paid to our department are in an unrelated clinical subject.