PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Costs and quality of life associated with acute upper gastrointestinal bleeding in the UK: cohort analysis of patients in a cluster randomised trial.
AUTHORS	Campbell, Helen; Stokes, Elizabeth; Bargo, Danielle; Logan, Richard; Mora, Ana; Hodge, Renate; Gray, Alasdair; James, Martin; Stanley, Adrian; Everett, Simon; Bailey, Adam; Dallal, Helen; Greenaway, John; Dyer, Claire; Llewelyn, Charlotte; Walsh, Tim; Travis, Simon; Murphy, Mike; Jairath, Vipul

VERSION 1 - REVIEW

REVIEWER	Johannes W. Rey HSK Clinic Department of Internal Medicine II Germany
REVIEW RETURNED	27-Dec-2014

GENERAL COMMENTS	This is a well-designed and interesting study by Campbell and colleagues. The results are relevant for a variety of public health care systems, although data were evaluated for hospitals in the UK. However, from my point of view a few aspects should be clarified. Inpatient stay, endoscopy and red blood cell transfusions are the key cost drivers. Do the authors have recommendations for cost-reduction strategies? More timely endoscopy has been reported to reduce hospital inpatient stay. What was the time of endoscopy in your cohort? Which cost savings can be expected when endoscopy is performed earlier and the length of in hospital stay is reduced? Approximately 50% of upper gastrointestinal bleedings are peptic ulcer bleedings. Do you have any reasons for only 22,7% of this aetiology in your whole cohort?, However, the Blatchford Score allows an assessment of whether patients should require immediately treatment or not. Please discuss if more patients may also be treated as outpatients for reducing in hospital cost.

REVIEWER	Lu, Yidan and Barkun, Alan Division of Gastroenterology McGill University Health Center Montreal, Quebec Canada
REVIEW RETURNED	27-Jan-2015

GENERAL COMMENTS	This is a nested prospective evaluation of health and social care resource use and costs than includes unpaid and informal care utilization as well as time away from employment and health related quality of life in acute upper gastrointestinal bleeding in the UK. This was performed in the context of a multicenter pragmatic cluster randomized pilot trial enrolling 936 patients between 2012 to 2013, with a time horizon of 28 days. The overall manuscript is well written, with detailed description of cost methods included in the appendix and summarized in the methods section. The results address the study?s objectives and the discussion brings up the principal, important limitations (missing data, and discrepancies with 2007 UK UGIB registry). My major comment relates to the high proportion of missing information both for cost and HRQoL (40%). The author used a multiple imputation with chained regression (MICE) to impute missing value. I do not have the expertise to assess the adequacy of such methods. However from my reading (Azur 2011), MICE assumes that the missing data is Missing At Random. In fact, the authors note that the group who provided a 28-day follow-up seem different from those who did not (see Discussion). Also, they used a normal rather than a more traditional gamma distribution for all values, which may not be adequate, as brought up in the discussion. ?MICE operates under the assumption that given the variables used in the imputation procedure, the missing data are Missing At Random (MAR), which means that the probability that a value is missing depends only on observed values and not on unobserved values (Schafer & Graham, 2002). In other words, after controlling for all of the available data (i.e., the variables included in the imputation model) ?any remaining missingness is completely random? (Graham, 2009). Implementing MICE when data are not MAR could result in biased estimates.? Regardless, I would have liked to see a sensitivity analysis that included no imputation without the missing information. One last comment on the days of work lost, the results section states 125 hrs or 17days whereas the discussion mentions 11.6 days. I suspect that the latter is a combination of (no return to work + partial return to work). It is not clearly described. The concern of double counting cost is a real one with regards to transfusions and procedures. The explanation given to avoid this is not entirely clear. Can the authors comment on a rough magnitude of costs of patients staying beyond 28 days because of issues related to the episode of bleeding? Has the EQ-5D been validated for a 28-day recall? And are imputation methods validly applied to EQ-5D scores? This reviewer does not possess the statistical expertise to comment on the multivariable hierarchal analysis, but the rationale of adopted methodology appears sound. Similarly, I am not familiar with UK costing databases, but the tabular documentation seems quite appropriate and thorough.

	Minor comments: P3 Line 46: SE (abbreviation) please spell out Line 48: There seems to be missing a verb (e.g were key cost drivers). Line 50: The seems to be a missing verb in the sentence P4 Line 12: main cost drivers (missing a verb) P8 Line 30: I cannot access reference 8 P11 Line37: reference 20 (typo: ?Hernshaw?,?acure?)

REVIEWER	Andrew Meltzer George Washington University, USA
REVIEW RETURNED	28-Jan-2015

GENERAL COMMENTS	Comprehensive approach to cost for upper GI hemorrhage. This is a secondary analysis of TRIGGER study. Good study. I was surprised that only 30% of patients required a therapeutic intervention. Despite that, mean length of stay is greater than 5 days. The 70% who did not require intervention seem like a target for cost reduction by reducing hospitalizations. Also, while I see total number of units transfused, I do not see how many patients were transfused. I suspect the mean per patient transfusion is misleading because some patients received no blood and some received a lot of blood. In our experience, there are very few "mean" patients since patients were either not very sick or extremely sick. I was surprised that 160 patients returned to the ED within 28 days. Were these episodes of recurrent bleeding? in which case a repeat therapeutic procedure or transfusion may have been required -- I don't see this cost considered. In figure 1, we see that most people have no problems and few people have extreme problems. This skewed data makes it difficult to consider the average GI bleed patient.

VERSION 1 – AUTHOR RESPONSE

Reviewer Name Johannes W. Rey
Institution and Country HSK Clinic
Department of Internal Medicine II
Germany

?Inpatient stay, endoscopy and red blood cell transfusions are the key cost drivers. Do the authors have recommendations for cost-reduction strategies? ?
Response: We appreciate the reviewer?s comments, however the purpose of this study was not to make direct recommendations about cost-reduction strategies and we believe that is beyond the

scope of the conclusions that can be drawn from this study.

?More timely endoscopy has been reported to reduce hospital inpatient stay. What was the time of endoscopy in your cohort? Which cost savings can be expected when endoscopy is performed earlier and the length of in hospital stay is reduced??
Response: We refrain from reporting about time to endoscopy, since this is detailed in the publication of the main clinical paper arising from the trial which will shortly be published. The issue regarding time to endoscopy has been previously addressed in a very large UK study which showed that delayed endoscopy (>24 hours) was associated with a significant increase in the risk adjusted length of stay (1.7 days longer) compared to endoscopy within 12 hours (Jairath et al, Endoscopy 2012; 44(8):723-30.) Of course health care providers would need to consider any cost reductions from shorter hospital stay against potential increases in expenditure required to facilitate faster access to endoscopy in their own hospital settings.

?Approximately 50% of upper gastrointestinal bleedings are peptic ulcer bleedings. Do you have any reasons for only 22,7% of this aetiology in your whole cohort??
Response: In westernised populations, the prevalence of peptic ulcer disease has been steadily declining and is now reported in the order of 30-40%, the discrepancies likely reflective of the differing methodologies of case ascertainment amongst studies. In the large UK audit of acute upper gastrointestinal bleeding in 2007, the prevalence of peptic ulcer bleeding (PUB) was 36%. We can only speculate on this issue, but the lower prevalence in this study may in part be reflective of the centres who took part where there appeared to be a higher burden of liver disease and variceal bleeding.

?However, the Blatchford Score allows an assessment of whether patients should require immediately treatment or not. Please discuss if more patients may also be treated as outpatients for reducing in hospital cost.?
Response: The utility of the Blatchford score lies in its ability to identify low risk patients who may be suitable for out-patient management. Existing evidence suggests that patients with a score of 0-1 may be safely managed out-patients, with early access outpatient endoscopy, provided there are no other medical reasons to be admitted to hospital. Thus in theory this group could be treated as outpatients and reduce the cost associated with hospitalisation.

Reviewer Name Lu, Yidan and Barkun, Alan
Institution and Country Division of Gastroenterology
McGill University Health Center
Montreal, Quebec
Canada

Please state any competing interests or state ?None declared?: None declared

Please leave your comments for the authors below
Title: Cost and quality of life associated with acute upper gastrointestinal bleeding in the UK: cohort analysis of patients in a cluster randomized trial.

Comments:

My major comment relates to the high proportion of missing information both for cost and HRQoL (40%). The author used a multiple imputation with chained regression (MICE) to impute missing value. I do not have the expertise to assess the adequacy of such methods. However from my reading (Azur 2011), MICE assumes that the missing data is Missing At Random. In fact, the authors note that the group who provided a 28-day follow-up seem different from those who did not (see Discussion). Also, they used a normal rather than a more traditional gamma distribution for all values, which may not be adequate, as brought up in the discussion.

?MICE operates under the assumption that given the variables used in the imputation procedure, the missing data are Missing At Random (MAR), which means that the probability that a value is missing depends only on observed values and not on unobserved values (Schafer & Graham, 2002). In other words, after controlling for all of the available data (i.e., the variables included in the imputation model) ?any remaining missingness is completely random? (Graham, 2009). Implementing MICE when data are not MAR could result in biased estimates.?

Regardless, I would have liked to see a sensitivity analysis that included no imputation without the missing information.

Response: The vast majority of costs (on average 86%) were incurred during the initial hospital episode and there was very little missing data for resource use in hospital, just 1.4%. For the vast majority of costs therefore, only a small amount of imputation was required.

As outlined in the quoted paragraph above, the term Missing at Random suggests that the values of missing data points can be determined by or are conditional on the values of other variables observed in the study, and beyond this the missingness is random. In this case we know by looking at observed baseline and demographic variables in our study, that patients with missing follow-up resource use and quality of life data were younger, with lower Blatchford Scores, and fewer co-morbidities. Several of these variables were used within the imputation regression equations so as to retain any correlation structures within the data, for example if younger patients with lower Blatchford scores have low numbers of GP visits post discharge, patients with missing data who share these characteristics will have low numbers of visits imputed.

We would have liked to fit a gamma model, but until statistical methods and software have further developed this is not straightforward. We have used the most appropriate methods available to us at this time.

In response to the reviewer?s suggestion about a sensitivity analysis comparing results with and without imputation, we point to the final paragraph of page 21 of the manuscript in the results section where we report mean levels of health related quality of life before and after imputation. The former value is 0.683, and the latter is 0.701, suggesting that the characteristics of non-responding patients were associated with higher levels of quality of life.

In terms of total health care costs to 28 days, an analysis comprising of only those patients with complete data on all resource use items includes only 361/936 (38.5%) patients and results in a total cost estimate of ?3185. With almost two thirds of patients missing this estimate is inefficient (many patients with just one or two missing variables are excluded) and is more likely to be subject to bias than the imputed total cost estimate, which, as stated above is mainly comprised of initial in-hospital costs where only 1.4% of data were missing.

We chose not to impute informal care costs and productivity losses given the magnitude of the missing data, and also the lack of additional information on which to base those imputations for

example existing living arrangements.

One last comment on the days of work lost, the results section states 125 hrs or 17days whereas the discussion mentions 11.6 days. I suspect that the latter is a combination of (no return to work + partial return to work). It is not clearly described.
Response: The reviewer is correct. Whilst the first figure stated above, 125 hrs or 17 days relates to lost days of work for patients in paid employment who did not return to work at all during the study period, the second figure of 11.6 is an average of work days lost amongst patients in paid employment who did not return to work and who did return to work (either at their pre-bleed hours or reduced hours). We have added the following sentence to the final paragraph on page 20 to make this clearer.

?Average working time lost across all patients who had been in paid employment was 11.6 days.?

The concern of double counting cost is a real one with regards to transfusions and procedures. The explanation given to avoid this is not entirely clear.
Response: When costing procedures (e.g. radiological interventions and interventions for serious adverse events such as stroke and myocardial infarction), we ensured against double counting as far as we were able, by removing average bed day costs already included in the published unit costs that we used to value or cost observed procedures. This was because inpatient bed days were costed separately for each patient in our study. This process is described in the detailed costing section of the online appendix.

?Can the authors comment on a rough magnitude of costs of patients staying beyond 28 days because of issues related to the episode of bleeding? ?
Response: This would not be possible as patients were only followed up to 28 days as per the trial protocol and not beyond.

Has the EQ-5D been validated for a 28-day recall? And are imputation methods validly applied to EQ-5D scores?
Response: The EuroQol EQ-5D asks respondents to report on their health state at the time at which they are completing the questionnaire (in this study 28 days), and therefore is not concerned with recall. It is widely used across health economic evaluations and studies in a range of disease areas and specialties and at a range of time points. EQ-5D is recommended in England and Wales as the instrument for measuring single index health utility scores in economic evaluations by the National Institute for Health and Care Excellence (NICE). Imputation methods which are tailored to suit the type of variable to be imputed are as valid for EQ-5D scores as for resource use.

This reviewer does not possess the statistical expertise to comment on the multivariable hierarchal analysis, but the rationale of adopted methodology appears sound. Similarly, I am not familiar with UK costing databases, but the tabular documentation seems quite appropriate and thorough.

Minor comments:

P3
Line 46: SE (abbreviation) please spell out - Response: ?standard error? has been added
Line 48: There seems to be missing a verb (e.g. were key cost drivers). ? Response: ?were? has

been added
Line 50: The seems to be a missing verb in the sentence ? Response: Amended to read ?Post-discharge healthcare costs were?..?
P4
Line 12: main cost drivers (missing a verb) ? Response: ?the? has been added
P8
Line 30: I cannot access reference 8 ? Response: the link for reference 8 has been updated to https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit
P11
Line37: reference 20 (typo: ?Hernshaw?,?acure?) ? Response: We have corrected Hernshaw to Hearnshaw and acure to acute

Reviewer Name Andrew Meltzer
Institution and Country George Washington University, USA
Please state any competing interests or state ?None declared?: None declared

Please leave your comments for the authors below Comprehensive approach to cost for upper GI hemorrhage. This is a secondary analysis of TRIGGER study. Good study.

?I was surprised that only 30% of patients required a therapeutic intervention. Despite that, mean length of stay is greater than 5 days. The 70% who did not require intervention seem like a target for cost reduction by reducing hospitalizations.?
Response: This figure is very similar to the data observed in the large UK national audit of acute upper gastrointestinal bleeding conducted in 2007, where the overall use of therapeutic endoscopy was approximately 25%. We appreciate the reviewer?s comments but it is possible that in the 70% who did not specifically require therapeutic intervention at endoscopy (as the lesion may have downgraded itself through natural course and as a result of acid suppression), they may still have required other interventions related to the GI bleed (e.g. acid suppression, transfusions), or indeed require admission to hospital due to concurrent comorbidity which may have decompensated as a result of the bleed.

Also, while I see total number of units transfused, I do not see how many patients were transfused. I suspect the mean per patient transfusion is misleading because some patients received no blood and some received a lot of blood. In our experience, there are very few "mean" patients since patients were either not very sick or extremely sick.
Response: In the total cohort 380 patients received at least one red cell transfusion (380/936= 40.5%) The mean number of RBC units transfused in these 380 patients was 3.89. The mean number of units transfused in Table 2 refers to the overall cohort, regardless of whether they were transfused or not.

I was surprised that 160 patients returned to the ED within 28 days. Were these episodes of recurrent bleeding? in which case a repeat therapeutic procedure or transfusion may have been required -- I don't see this cost considered.
Response: This figure refers to the total number of visits to the ED in the cohort up to day 28, regardless of the reason. We have added a footnote to Table 2 to make it clearer that we are reporting total number of visits rather than patients, and that some patients had more than one visit. In

terms of further bleeding there were 33 cases in the total cohort up to hospital discharge and then a further 11 cases up to day 28 (i.e. 44 cases in total). As this is a cohort with extensive medical comorbidity, this number of visits to the ED is likely a reflection of concurrent comorbidity.

Of the 11 cases of further bleeding occurring after hospital discharge, those requiring hospital treatment were costed only by multiplying reported hospital inpatient length of stay by an average inpatient bed day cost. The possibility exists therefore that additional costs for a very small number of therapeutic procedures and transfusions may not have been included, but given the small number of cases involved, the impact of such an under-estimation on the overall cost results from the study is likely to be negligible.

Footnote to Table 2

*** Data reported are total number of bed days and visits with some patients having numerous contacts

?In figure 1, we see that most people have no problems and few people have extreme problems. This skewed data makes it difficult to consider the average GI bleed patient.?
Response: In this cluster randomised study 59% of all consecutive patients presenting with AUGIB were enrolled into the trial. A major strength of this study design was that a high proportion of eligible patients were enrolled into the study, indicating the results are broadly generalizable to the spectrum of patients admitted with GI bleeding in the UK.

Figure 1 illustrates that many patients do have some problems, and allows affected domains to be identified (for example around a third of patients reported some problems with usual activities, pain and discomfort, and anxiety and depression). These data are converted into a single index score which allows the average quality of life of these patients to be considered against that of an age matched population (see sentence 1 on page 22).

VERSION 2 – REVIEW

REVIEWER	Alan Barkun and Yidan Lu McGill University Health Center Montreal, Qc Canada
REVIEW RETURNED	26-Feb-2015

GENERAL COMMENTS	The authors have addressed with great clarity all our previous concerns and have made corrections/modification when necessary. We therefore recommend the acceptance of this paper.