PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Inversed relationship between completeness of follow-up and coverage of postoperative complications in gallstone surgery and ERCP. A potential source of bias in patient registers.
AUTHORS	Enochsson, Lars; Blohm, My; Sandblom, Gabriel; Jonas, Eduard; Hallerbäck, Bengt; Lundell, Lars; Österberg, Johanna

VERSION 1 – REVIEW

REVIEWER	Kishan Patel Centre for Public Health, Queen's University Belfast, Northern Ireland
REVIEW RETURNED	10-Oct-2017

GENERAL COMMENTS	Thank you for the opportunity to read this interesting paper. The study aims to analyse the factors that can affect the validity of national registers, specifically for follow-up data. I think the study is especially important in today's climate, when registry studies are increasing in popularity. Whilst the analysis of data is very clear, I do have some minor comments: 1) I do think there are points in the paper where it is difficult to understand the overall reasons for conducting the study. For example, in the first paragraph of the discussion, there is a discussion of the primary finding of the paper, but it quickly transitions to presenting results about specifically ERCP and cholecystectomy procedures, which in my mind, is not the focus of the study. 2) I would like to see an aims section in the paper. In the last line of the introduction, the aim is stated very briefly. I'd like to see some elaboration on this, just to ease the reader into the methods section. I feel that the transition between a discussion of the pros and cons of register data to the methods section is slightly confusing. Maybe some elaboration on the "factors" that were to be analysed would help bridge the gap? 3) I don't like the use of the term 'self-reported'. I understand why it has been used, but I'd be hesitant to use the term when, as alluded to in the methods section, the registers are filled in by medical professionals.

	4) I'm not sure why the unadjusted models are presented in the tables when it doesn't add anything to the analysis. I think the fully adjusted model presented alone would suffice. If the authors would like to use multiple models, perhaps the first model could be adjusted for only age and sex, the second for age, sex and ASA class, and a third as a fully adjusted model. I think the pattern between those models could be interesting. 5) There are a couple of points where websites are given in-text, when I think referencing them and adding the websites to the bibliography instead would help the readability of the text.

REVIEWER	Katarina Steen Carlsson Lund University Department of Clinical Sciences, Malmö Health Economic Unit Medicon Village SE-223 81 LUND Sweden
REVIEW RETURNED	11-Oct-2017

GENERAL COMMENTS	The paper presents the results of from analyses of data in one Swedish quality register, the Gallriks registry on gallstone surgery and endoscopic retrograde cholangiopancreatgraphy (ERCP). The objective is to analyze factors that may affect the validity of follow-up data in Gallriks. The authors raise this issue also as a more general issue for quality registers and other real-world data sources as opposed to clinical trials where strict protocols, close monitoring and small sample sizes make sure that data is complete. The authors conclude that there is a need for assessing degree of complete follow-up as an additional measure of validity for quality registers. The paper raises an important and relevant issue pertinent in an era of increasing interest for real-world data and especially registry-based research. They use a measure relevant to surgical interventions, completeness of 30-day follow-up, as outcome measure to explore whether hospitals with >=90% complete 30-day follow-up differ in terms of reported adverse events from those with lower level complete follow-up. The 10-year retrospect gives an impressing number of observations where less than 5% of observations failed to report a complete 30-day follow-up. However, the authors also find that 20% of hospitals performing cholecystectomy and 17% of those performing ERCPs had complete follow-up of less than 90%. The paper finds indications that reports of adverse events by ERCP with/without 90% complete follow-up differ significantly. In Method the authors may clarify the motive for outcome measure "complete 30 day follow-up" for people without detailed expert knowledge of gallstone surgery. Is the 30-day follow-up mandatory based on guidelines? Are there legitimate reasons for not having 30 day follow-up? Is it possible that some patients have been followed-up earlier and declared well enough not to need a follow-up at 30 days? Is it possible that some patients do not show up at 30-day follow-up? How was mortality accounted for? How many cases out of the 150 000+ cases were deceased within 30 days – or were they excluded?

	In Discussion, the authors may elaborate to what extent other outcome measures, if any, were considered for this assessment. In addition, the chosen outcome measure seems very relevant in the assessment of quality of data after a surgical intervention where guidelines stipulate such a standard follow-up. In a broader context, for what other conditions and interventions may this particular set up be relevant. Does it apply to all surgical interventions? Why/why not? What kind of measures would be relevant for chronic disease? The results presented indicate that best reporting frequency and completeness of data is achieved by hospitals with greater patient volumes. This is the only explanation consistent with the fact that less than 5% of observations in Gallriks have incomplete follow-up, but as much as 20% and 17 % of hospitals have less than 90% complete 30-day follow-up. However, this conclusion is not made by the authors. For a broader understanding of the context, it would be helpful if the authors also addressed the organization issues. That is, do hospitals need to have a critical mass in order to contribute to group-level follow-up and is there a case for identification of required patient volumes and structures for effective management of data. Figures 3a, 3b do not seem to add much information apart from what is communicated verbally in the text. May be transferred to a Table and offered as supplementary material.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:
1) I do think there are points in the paper where it is difficult to understand the overall reasons for conducting the study. For example, in the first paragraph of the discussion, there is a discussion of the primary finding of the paper, but it quickly transitions to presenting results about specifically ERCP and cholecystectomy procedures, which in my mind, is not the focus of the study.

Response: We agree with the reviewer that the focus of this manuscript is primarily to state the importance of a thorough follow-up and not to compare the outcome of cholecystectomies with that of ERCP. However, our analysis of the register data found that the importance of a thorough 30-day follow-up seems to have a higher impact on the outcome in the ERCP group, which we found interesting. About these differences, one can only speculate, but we think that it could be due to the fact that, in general, ERCP procedures are technically more demanding and have a higher risk of complication. Furthermore, the complications during cholecystectomies are usually detected intraoperatively whereas complications in conjunction with ERCP, like pancreatitis and cholangitis, are usually detected postoperatively. Therefore it is even more important to have coordinators that can do an accurate 30-day follow-up in the ERCP group. We agree with the reviewer that the aim of the study is to focus on the importance of follow-up and have therefore rewritten the first paragraph in the discussion.

2) I would like to see an aims section in the paper. In the last line of the introduction, the aim is stated very briefly. I'd like to see some elaboration on this, just to ease the reader into the methods section. I feel that the transition between a discussion of the pros and cons of register data to the methods section is slightly confusing. Maybe some elaboration on the "factors" that were to be analysed would help bridge the gap?

Response:We agree with the reviewer number one, that the aim of the study could be further specified. We therefore, have included an aims section at the end of the introduction where we have tried to, in a better way, explain the purpose of the manuscript. We have excluded the word “factors” and replaced it with a more specific explanation in the aims section.

3) I don't like the use of the term 'self-reported'. I understand why it has been used, but I'd be hesitant to use the term when, as alluded to in the methods section, the registers are filled in by medical professionals.

Response:We agree with the reviewer and have excluded the word “self-reported”.

4) I'm not sure why the unadjusted models are presented in the tables when it doesn't add anything to the analysis. I think the fully adjusted model presented alone would suffice. If the authors would like to use multiple models, perhaps the first model could be adjusted for only age and sex, the second for age, sex and ASA class, and a third as a fully adjusted model. I think the pattern between those models could be interesting.

Response:We agree with the reviewer that it’s not necessary to include the unadjusted values and have therefore omitted them in tables 2 and 3. Regarding the adjusted Odds Ratios they have been adjusted for sex, age, ASA-Class, indications and whether the procedures were scheduled or acute. In the final adjustments we have only included, except for sex and age which always were included, the cofounders that were statistically significant.

5) There are a couple of points where websites are given in-text, when I think referencing them and adding the websites to the bibliography instead would help the readability of the text.

Response: We have now added the websites to the bibliography.

Reviewer 2:
1) In Method the authors may clarify the motive for outcome measure "complete 30 day follow-up" for people without detailed expert knowledge of gallstone surgery.

Response: We agree with the reviewer that the meaning of what is a complete 30-day follow-up could be better explained and we have now tried to do so in the first paragraph in the methods section.

2) Is the 30-day follow-up mandatory based on guidelines?

Response: Yes, the 30 day follow-up is a well-established standard that is common in many Surgical quality registers as well as in other studies. [1]

3) Are there legitimate reasons for not having 30 day follow-up?

Response: No, we do not think so.

4) Is it possible that some patients have been followed-up earlier and declared well enough not to need a follow-up at 30 days?

Response: No, the 30-day follow-up should be performed on all patients by the local coordinator at each participating unit. This is a mandatory follow-up where the local coordinator looks through the medical records to see if there have been any adverse events. Often, but not always, the coordinator also contacts the patient by phone. Thus, the patient does not have to attend to a regular physical exam at the hospital.

5) Is it possible that some patients do not show up at 30-day follow-up?

Response: As stated above none of the patients show up at the 30-day follow-up since the 30-day follow-up consists of screening through the medical records and contacting the patient by phone.

6) How was mortality accounted for?

Response: The mortality is automatically transferred to the register from the Swedish Central Death Register.

7) How many cases out of the 150 000+ cases were deceased within 30 days - or were they excluded?

Response: The overall mortality of cholecystectomies and ERCP is 2.3% and in total 3683 patients died. However, the mortality is significantly higher in patients where an ERCP was performed both due to the fact the there are more patients with malignancy in this group and that ERCP have significantly more postoperative complications. No patients are supposed to be excluded since the 30-day follow-up is performed by looking through the medical records and obviously, death within 30 days is considered as the “ultimate complication” and thus registered as such within the register.

8) In Discussion, the authors may elaborate to what extent other outcome measures, if any, were considered for this assessment. In addition, the chosen outcome measure seems very relevant in the assessment of quality of data after a surgical intervention where guidelines stipulate such a standard follow-up. In a broader context, for what other conditions and interventions may this particular set up be relevant. Does it apply to all surgical interventions? Why/why not? What kind of measures would be relevant for chronic disease?

Response: The purpose of the register for cholecystectomy and ERCP is to analyze the intra- and postoperative outcome of these procedures. In this respect, it mimics many other surgical and orthopedic registers like the hernia register, the fracture register and the hip replacement register to mention a few. Some of the complications presented in our manuscript is specifically related to the interventions registered in GallRiks, like postoperative pancreatitis, whereas others are more of a general nature. GallRiks is not a disease register like for example the Diabetes register etc. However, we think that the principle of completeness of follow-up and coverage of complications apply to all registers.

9) The results presented indicate that best reporting frequency and completeness of data is achieved by hospitals with greater patient volumes. This is the only explanation consistent with the fact that less than 5% of observations in Gallriks have incomplete follow-up, but as much as 20% and 17 % of hospitals have less than 90% complete 30-day follow-up. However, this conclusion is not made by the authors.

Response: We understand the comments made by the reviewer but we do not fully agree. In hospitals with small procedure volumes, the absolute numbers of missing follow-up registrations can be rather low but since the procedure volumes often are small it will have a higher impact on the percentages given. On the contrary, we sometimes have experienced that the hospitals with greater volumes sometimes can have logistic difficulties in completing the 30-day follow-up.

10) For a broader understanding of the context, it would be helpful if the authors also addressed the organization issues. That is, do hospitals need to have a critical mass in order to contribute to group-level follow-up and is there a case for identification of required patient volumes and structures for effective management of data.

Response: In GallRiks all cholecystectomies and ERCPs done in Sweden are expected to be registered, no matter how few each participating unit performs. Of course in reality this goal is difficult to achieve but the coverage of GallRiks is around 90% (registrations in GallRiks compared to the National Patient Registry). The advantage of this policy is that when summing up the total numbers of registrations you get a good assessment of complications with good statistical power. However, we agree with the reviewer that the organizational concerns could be more emphasized since we know that there can be big differences in the way the logistics of the registrations are carried out and we think that this has a big impact on the quality of the register. We, therefore, have revised the manuscript somewhat to emphasize the importance of the organizational issues in the Discussion.

11) Figures 3a, 3b do not seem to add much information apart from what is communicated verbally in the text. May be transferred to a Table and offered as supplementary material.

Response: We respect the reviewer´s opinion but would prefer to keep these figures, if possible.

1. Ohlsson H, Winso O: Assessment of the Surgical Apgar Score in a Swedish setting. Acta Anaesthesiol Scand 2011, 55(5):524-529.

VERSION 2 – REVIEW

REVIEWER	Katarina Steen Carlsson Lund University, Sweden
REVIEW RETURNED	26-Nov-2017

GENERAL COMMENTS	This version is revised with adjustments throughout the different sections of the paper. The authors have also changed the title and the present version may better represent the results. However, since the analysis of the paper focus on one register of two surgical procedures relevant for one condition, the title would read better if reference was made to "…the example of gallstone surgery…" or "…in gallstone surgery…" I noted that the objective of the abstract and the wording of Aims does not match fully in this version. The authors may consider a complete correspondence. Two of the questions I had for the original version have been addressed in the response to reviewers but not in the main text: Mortality and its reporting. In the response to reviewers, the authors have clarified how 30-day mortality is reported; it is included in

	Adverse events. This information should be included in the paper as BMJ Open is a general journal and not all readers can be expected to be familiar with the traditions of gallstone surgery research. Following the authors explanations on data collection procedures, mortality should be complete in registration for all patients as it is based on population registry data. Maybe it is the standard of gallstone surgery research to mix death with other adverse events. To me it seems odd not to report 30-day mortality separately from other adverse events. From a patient, citizen and health policy perspective 30-day mortality should be reported separately. For the objective of the paper, it is not a problem that ERCP and cholecystectomies have different expected mortality. Moreover, if I am correct in my assumption that mortality is always reported, the paper may achieve an interesting and clear focus on lack of reporting on the other adverse events if death is reported separately. Reporting by small and large hospitals. In the response to reviewers, the authors do not fully agree with the analysis that the reported data show that larger hospitals have more complete data. However, it is mathematically impossible that 20% and 17 % of hospitals have less than 90% complete 30-day follow-up on cholecystectomies and ERCP at the same time as less than 5% of patients in Gallriks overall have incomplete follow-up, if it were not for a relatively greater number of smaller hospitals having trouble meeting the 90% complete follow-up goal. Of course, not all big hospitals need to be meeting the goal but the numbers would not match if smaller hospitals are not overrepresented in the group with incomplete follow-up. If the authors do not want to discuss this, I suggest the authors omit the reporting of the data on percentage of hospitals who do not meet the goal of complete follow-up. It would be a pity though since it is an interesting, and from a policy perspective highly relevant, finding.

VERSION 2 – AUTHOR RESPONSE

1. Regarding the title

Response: We agree with the reviewer and have changed the title.

2. The wording of Aims.

Response: We have changed the aims section somewhat so now that they correspond both in the abstract as well as in the text.

3. Mortality and its reporting.....

Response: We agree and have added a paragraph about reported mortality at the end of the results section on page 9 in the manuscript.

4. Reporting by small and large hospitals.

Response: We agree that differences in completeness between high- and low-volume units may account for the high completeness at the national level despite some units being very far from this level of completeness. We have commented on this in the discussion. We have also added a sentence in the legend to figure 2, clarifying how the hospitals are ordered on the x-axis.