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)	POSTOPERATIVE WOUND DEHISCENCE AFTER LAPAROTOMY: A USEFUL HEALTH CARE QUALITY INDICATOR? A COHORT STUDY BASED ON NORWEGIAN HOSPITAL ADMINISTRATIVE DATA
AUTHORS	Helgeland , Jon; Tomic, Oliver; Hansen, Tonya; Kristoffersen, Doris; Hassani, Sahar; Lindahl, Anne

VERSION 1 - REVIEW

REVIEWER	Nadia A. Henriksen Zealand University Hospital, Koege
REVIEW RETURNED	22-Oct-2018

GENERAL COMMENTS	It is very well-written manuscript. Postoperative wound dehiscence is an important complication to laparotomy and this large database study examines wound dehiscence as a health care indicator. The study is strengthened by the large data material, but this is also the limitation. Has the data validity of the Norwegian Patient Registry ever been studied? Is smoking status and obesity available? Please comment on this. Another important aspect is the definition of wound dehiscence. Are superficial dehiscences with intact fascia included or not? Or is it only fascial dehiscences/deep wound dehiscence/burst abdomen that are included? Please comment on that and strictly specify, what you consider as postoperative wound dehiscence. I think it is important to stress that wound dehiscence possibly is a factor that is modifiable and therefore is an important health indicator. It seems that a standardized surgical method may decrease the number wound dehiscences. I think the statistical method is difficult to fully understand, especially the tables. I would suggest statistical review. Further comments: P4 L74-77: what about emergency surgery? P4 L78: the objective of the study is to investigate the incidence of PWD. This should also be included in the Objectives section of the abstract. The Objectives section of the abstract is not entirely clear. Please clarify that. P5 L100-105: Definition of wound dehiscence. Are superficial wound dehiscences included? Is the code for 'reclosure' specific for fascial closure or are superficial dehiscences requiring 'skin reclosure' also included? Table 1 is not easy to read. Absolute numbers with percentages in parentheses would be useful. What is the selection of variables based on? For instance, I don't understand why 'hemiplegia/paraplegia' is included. Is that relevant? Table 2 is also difficult to follow. Is it Crude OR or adjusted OR? A column with absolute numbers could be inserted. 'Spline function

	1-4' should be explained below the table. I guess, it refers to hospital volume, but this should be explained and also it would be useful to have the volume intervals specified. Considering 'operation type', what is the reference operation, OR =1 ? Please specify what 'several organs' includes? Please specify if 'hernia' is 'groin hernia', and is 'abdominal wall' ventral hernias? Or what does that refer to?

REVIEWER	Ben Byrne University of Bristol
REVIEW RETURNED	08-Nov-2018

GENERAL COMMENTS	Thank you for the opportunity to review this interesting manuscript. I think it makes an important contribution, but have some concerns about the methods which I would like to raise. Most importantly, how have the authors considered / allowed for minimally invasive surgery within the study population? I could not find any reference to laparoscopic surgery. The risk of wound dehiscence is likely to be significantly associated with minimally invasive techniques. In HES in the UK, this is indicated by associated OPCS codes indicating a laparoscopic approach. Differences in use of laparoscopy / robotics across units may account for variation in dehiscence rates. Please can the authors provide more justification for excluding immunocompromised patients? I would probably rather these were not excluded, as it is not possible to determine which patients were also on steroids or other immunomodulatory medications that could have influenced dehiscence rates. Therefore, there may be bias in excluding some immunocompromised patients, but not others, due to the administrative nature of the data. I am not familiar with all data available in this dataset. Are there any socio-economic variables that could be entered into the analysis? These are likely related to other unmeasured variables, such as smoking and obesity, and may therefore be important in risk-adjustment. The dataset is naturally heterogeneous. However, I am unsure if the current grouping of operations is optimal. Clean and contaminated operations are grouped together, and the level of bacterial contamination is likely to be an important factor associated with dehiscence. For example, GI procedures include cardiomyotomy (a clean operation in which the GI tract should not be breached) with colonic resection (which may be clean-contaminated, contaminated, or dirty). I consider the label 'hernia repair' a little misleading - the codes listed seem to represent diaphragmatic or hiatal hernia and this should be more explicit. Otherwise, the reader may consider it to include other types of hernia, such as incisional hernia. While hospital volume is appropriately covered in the discussion, I would be interested to see a re-run of the regression after excluding the 4 very low volume hospitals. I suspect that volume would then not be significantly associated with outcome.

	The limitations provided are well considered, but there may be other limitations not discussed such as non-operative management of dehiscence that is not captured using the study methods, or patients who die after dehiscing their wound without reoperation. Regarding the style and writing, I had two minor points: - The sentence spanning lines 147-8 seems to be a fragment. - I do not think that sentences should end with 'etc'. There are better ways to itemise a partial list.

REVIEWER	Professor David A Watters Deakin University and Barwon Health Australia
REVIEW RETURNED	15-Nov-2018

GENERAL COMMENTS	My difficulties with your paper is whether or not you differentiate emergency and elective laparotomy, which has a great impact on morbidity and mortality. I am not convinced that a second procedure for wound dehiscence truly reflects quality unless the case-mix and risk is standardised Your statistical methods seem very reasonabe, however I would like to comment on Fig 1 - there is one significant hospital in the top box "small hospitals with an acute function" - I understand the method the authors propose to identify outliers is based on limiting the false discovery rate (which I assume this hospital must have passed), but possibly using an orthogonal terchnique to identify outliers might identify this as an outlier (3g 1.5 x interquartile range above the third quartile) - is this hospital with a PWD rate of almost 3% really not an outlier? I think your variation in your primary outcome measures may simply reflect variation in case-mix, and perhaps urgency. I am not sure that you achieve your objective of determining a health care quality indicator. For me, you have not defined a" laparotomy" and how this might be influenced by laparoscopic or robotic approaches, even where conversion or a laparotomy approach is used in conjunction with a minimally invasive approach, and this makes repeating such a study elsewhere difficult to standardise. I think this could be achieved if you looked at, or reported by and compared hospitals according to, for example, emergency laparotomy, or elective open colorectal surgery.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1

Reviewer Name: Nadia A. Henriksen

Institution and Country: Zealand University Hospital, Koege

Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below It is very well-written manuscript. Postoperative wound dehiscence is an important complication to laparotomy and this large database study examines wound dehiscence as a health care indicator.

1. The study is strengthened by the large data material, but this is also the limitation. Has the data validity of the Norwegian Patient Registry ever been studied?

Case identification based on diagnoses and/or procedures has been extensively validated. We have included a better explanation in Discussion (l.233-237)

2. Is smoking status and obesity available? Please comment on this.

We have analysed data from the patient administrative database, and thus we do not have information on smoking, obesity or other potentially relevant clinical factors, other than previous diagnoses and admissions. (noted in Material and Methods l.169-171, Discussion l.238-253)

3. Another important aspect is the definition of wound dehiscence. Are superficial dehiscences with intact fascia included or not? Or is it only fascial dehiscences/deep wound dehiscence/burst abdomen that are included? Please comment on that and strictly specify, what you consider as postoperative wound dehiscence.

Deep dehiscence includes only wounds with dehiscence of the facia and communication with adominal cavity. (as explained in Material and Methods l.104-110). The code texts are in the Online supplement, Table 6

4. I think it is important to stress that wound dehiscence possibly is a factor that is modifiable and therefore is an important health indicator. It seems that a standardized surgical method may decrease the number wound dehiscences.

This has been included in Introduction (l.75-77) and Conclusions (l. 265-268)

I think the statistical method is difficult to fully understand, especially the tables. I would suggest statistical review.

Further comments:

5. P4 L74-77: what about emergency surgery?

We have included laparotomies from all episodes, both acute and elective, in the data as well in the risk adjustment model. This has been elaborated in Material and Methods (l.95, 163, 170-171) We have also mentioned that the model has been tested for interactions (Materials and methods,l.138). If acute and elective operations were to follow different models, this would have shown up as interaction between the variable “acute episode” and the other risk adjustment variables. This was not the case (Results, l.190).

6. P4 L78: the objective of the study is to investigate the incidence of PWD. This should also be included in the Objectives section of the abstract. The Objectives section of the abstract is not entirely clear. Please clarify that.

The main objective has been to study variation. Accurate estimation of incidence would require a study of the sensitivity of the indicator, which is a daunting task outside the scope of this study. We see that our formulation was easy to misunderstand, this has now been amended (Introduction, l.86)

7. P5 L100-105: Definition of wound dehiscence. Are superficial wound dehiscences included? Is the code for 'reclosure' specific for fascial closure or are superficial dehiscences requiring 'skin reclosure' also included?

Deep dehiscence includes only wounds with dehiscence of the facia.and communication with adominal cavity. (as explained in Material and Methods l.104-110, supplement Table 6)

8. Table 1 is not easy to read. Absolute numbers with percentages in parentheses would be useful.

Agreed. Tables have been thoroughly revised

9. What is the selection of variables based on? For instance, I don't understand why 'hemiplegia/paraplegia' is included. Is that relevant?

The literature has identified several comorbidities as risk factors, these have now been listed in the Introduction (l.77-78). We agree that hemiplegia/paraplegia is not very relevant, and have dropped it from the table

10. Table 2 is also difficult to follow. Is it Crude OR or adjusted OR?

The table shows adjusted ORs, this is now included in the table heading

11. A column with absolute numbers could be inserted.

A table showing frequencies of the various operation types has been included in the online supplement (Table 2).

12. 'Spline function 1-4' should be explained below the table. I guess, it refers to hospital volume, but this should be explained and also it would be useful to have the volume intervals specified.

The splines are used for modelling age dependency. This is now noted in the table. For natural splines, support intervals are not very intuitive. The knots are age quartiles, this is now noted in Material and Methods (l.164-165). Also, we now show the values of the estimated spline function at several ages, with age 40 as the reference.

13. Considering 'operation type', what is the reference operation, OR =1 ?

The variable operation type has been standardised by mean zero (on the log-odds scale), now noted in the table and in Material and Methods (l.166-167). It was felt that no type was a natural reference. The ORs are thus standardised to geometric mean 1 which means "average type"

14. Please specify what 'several organs' includes?

It means "more than one of the types, excluding exploratory laparotomy". The tables has been revised.

15. Please specify if 'hernia' is 'groin hernia', and

It is diaphragmal hernia. The tables has been revised accordingly, Also, the codes can be found in the Online supplement, Table 8.

16. Is 'abdominal wall' ventral hernias? Or what does that refer to?

The codes for this category has been revised (see Online supplement, Table 7), and the category relabeled “Exploratory laparotomy”, which probably is the most common use

Reviewer: 2

Reviewer Name: Ben Byrne

Institution and Country: University of Bristol

Please state any competing interests or state ‘None declared’: None

Please leave your comments for the authors below Thank you for the opportunity to review this interesting manuscript. I think it makes an important contribution, but have some concerns about the methods which I would like to raise.

17. Most importantly, how have the authors considered / allowed for minimally invasive surgery within the study population? I could not find any reference to laparoscopic surgery. The risk of wound dehiscence is likely to be significantly associated with minimally invasive techniques. In HES in the UK, this is indicated by associated OPCS codes indicating a laparoscopic approach. Differences in use of laparoscopy / robotics across units may account for variation in dehiscence rates.

All laparosopic and endoscopic surgery has been excluded. The PWD rates only pertain to laparotomies. Conceivably, open surgery is used for the more complicated or more urgent cases, leading to laparotomies having higher risk, and hospitals with high laparoscopy rates being unfavourably compared. We cannot control for this other than by risk adjusting for operation type. The increased use of laparotomy is why we have included calendar year in the model. It turns out, however, that PWD rate is decreasing along with the laparotomy rate (see Table 2). This theme, although interesting, was omitted as it was felt that the paper would be too long and less clear

18. Please can the authors provide more justification for excluding immunocompromised patients? I would probably rather these were not excluded, as it is not possible to determine which patients were also on steroids or other immunomodulatory medications that could have influenced dehiscence rates. Therefore, there may be bias in excluding some immunocompromised patients, but not others, due to the administrative nature of the data.

This is primarily because this is part of the OECD and AHRQ specifications. Immunodeficiency and use of immunosuppression as risk factors, is now noted in Introduction (l. 81). Unfortunately, we have no data for immunosuppressive medication, unless this is captured by one of the diagnoses. We believe the OECD approach is the best compromise

19. I am not familiar with all data available in this dataset. Are there any socio-economic variables that could be entered into the analysis? These are likely related to other unmeasured variables, such as smoking and obesity, and may therefore be important in risk-adjustment.

Unfortunately not, due to present data restrictions in Norway

20. The dataset is naturally heterogeneous. However, I am unsure if the current grouping of operations is optimal. Clean and contaminated operations are grouped together, and the level of bacterial contamination is likely to be an important factor associated with dehiscence. For example, GI procedures include cardiomyotomy (a clean operation in which the GI tract should not be breached) with colonic resection (which may be clean-contaminated, contaminated, or dirty).

There is no clinical information regarding clean vs contaminated operations available in our administrative dataset. We think that the procedure codes may not be sufficient to distinguish clean from contaminated operations and have thus not tried to differentiate.

21. I consider the label 'hernia repair' a little misleading - the codes listed seem to represent diaphragmatic or hiatal hernia and this should be more explicit. Otherwise, the reader may consider it to include other types of hernia, such as incisional hernia.

It is diaphragmal hernia. The tables has been revised accordingly, Also, the codes can be found in the Online supplement, Table 8.

22. While hospital volume is appropriately covered in the discussion, I would be interested to see a re-run of the regression after excluding the 4 very low volume hospitals. I suspect that volume would then not be significantly associated with outcome.

This has been done and noted in Results(l.204-205). Volume is still significant, this may be due to a small and non-systematic variation among the large and regional hospitals

23. The limitations provided are well considered, but there may be other limitations not discussed such as non-operative management of dehiscence that is not captured using the study methods, or patients who die after dehiscing their wound without reoperation.

This has now been included in Discussion (l. 250-251). Patients who die without reclosure are missed in this analysis. We think the number of patients this applies to is very small, and we do not think this will influence our results. However, the total care of the patient should be reflected in the dehiscense rate, including the perioperative management, thus reflecting the quality of the care including the surgery. Very few patients with deep dehiscenses would be cared for with only non-operative management.

24. Regarding the style and writing, I had two minor points:

- The sentence spanning lines 147-8 seems to be a fragment.

- I do not think that sentences should end with 'etc'. There are better ways to itemise a partial list.

This has been amended

Reviewer: 3

Reviewer Name: Professor David A Watters

Institution and Country: Deakin University and Barwon Health, Australia

Please state any competing interests or state ‘None declared’: None

Please leave your comments for the authors below

25. My difficulties with your paper is whether or not you differentiate emergency and elective laparotomy, which has a great impact on morbidity and mortality.

We have included laparotomies from all episodes, both acute and elective, in the data as well in the risk adjustment model. This has been elaborated in Material and Methods (l.95, 163, 170-171) We have also mentioned that the model has been tested for interactions (Materials and methods,l.138). If acute and elective operations were to follow different models, this would have shown up as interaction between the variable “acute episode” and the other risk adjustment variables. This was not the case (Results, l.190).

26. I am not convinced that a second procedure for wound dehiscence truly reflects quality unless the case-mix and risk is standardised

We have adjusted for case mix as much as the available administrative data permit. Residual case mix is discussed in Discussion (l.238-253). The indicator has a long history and is used by AHRQ and OECD, without case mix adjustment. In the original paper by Hannan, the indicator was justified by association with measures of process quality. It is also associated with surgical technique, surgeon experience and perioperative care

27. Your statistical methods seem very reasonabe, however I would like to comment on Fig 1 - there is one significant hospital in the top box "small hospitals with an acute function" - I understand the method the authors propose to identify outliers is based on limiting the false discovery rate (which I assume this hospital must have passed), but possibly using an orthogonal terchnique to identify outliers might identify this as an outlier (3g 1.5 x interquartile range above the third quartile) - is this hospital with a PWD rate of almost 3% really not an outlier?

It is important to account for uncertainty in the PWD rates when identifying outliers. For the mentioned hospital, the standard deviation is almost equal to the estimated rate.

28. I think your variation in your primary outcome measures may simply reflect variation in case-mix, and perhaps urgency.

We think that all quality indicators most of all are warning flags, to be followed up by investigating resources and procedures in institutions that are flagged as outliers (see Discussion l.274-277). We have been able to adjust for many of the risk factors mentioned in the literature, including whether the hospital episode is acute or elective. The literature also reports effects of factors pertaining to the hospitals' care, such as surgeon experience and surgical technique, and events that only appear after the primary operation, such as deep wound infection and bleeding. None of these factors can be used to control for case mix in a comparison of hospitals. The strongest reported effect is deep wound infection, with an odds ratio of 6.43 (van Ramshorst et al). For an informal sensitivity estimate, we can imagine that some unmeasured factor is distributed among hospitals with the same relative frequencies as chronic pulmonary disease (overall rate of around 7%). This factor would have to have an odds ratio of 10 to explain the observed variation in adjusted PWD rates among hospitals.

29. I am not sure that you achieve your objective of determining a health care quality indicator. For me, you have not defined a" laparotomy" and how this might be influenced by laparoscopic or robotic approaches, even where conversion or a laparotomy approach is used in conjunction with a minimally invasive approach, and this makes repeating such a study elsewhere difficult to standardise.

All laparosopic and endoscopic surgery has been excluded. We have now included definitions of laparotomy and wound dehiscence in Material and Methods (l. 104-110), and tabulated the frequencies of robotic and converted operations, both very low (see Table 1)

30. I think this could be achieved if you looked at, or reported by and compared hospitals according to, for example, emergency laparotomy, or elective open colorectal surgery.

We have included laparotomies from all episodes, both acute and elective, in the data as well in the risk adjustment model. This has been elaborated in Material and Methods (l.95, 163, 170-171) We have also mentioned that the model has been tested for interactions (Materials and methods,l.138). If acute and elective operations were to follow different models, this would have shown up as interaction between the variable "acute episode" and the other risk adjustment variables. This was not the case (Results, l.190).

VERSION 2 – REVIEW

REVIEWER	Nadia A. Henriksen Dept. of Surgery, Zealand University Hospital
REVIEW RETURNED	10-Jan-2019

GENERAL COMMENTS	Congratulations with your paper. I have no further comment.

REVIEWER	Ben Byrne University of Bristol
REVIEW RETURNED	19-Dec-2018

GENERAL COMMENTS	Thank you for the revised manuscript which has sufficiently addressed the issues previously raised. I think this paper is now appropriate for publication.

REVIEWER	Prof David A Watters Deakin University and Barwon Health Australia
REVIEW RETURNED	14-Jan-2019

GENERAL COMMENTS	I remain uncertain as to the reason for the outliers for postoperative wound dehiscence you have identified. It would be interesting to know when the hospitals review their data in the light of being identified as an outlier whether they remain real or apparent outliers. However, you have addressed the issues I raised in my original review, given that you need to work with a large population based health database.

VERSION 2 – AUTHOR RESPONSE

Response to reviewer 3:

We agree that studying the causes of the observed variation would be very worthwhile. However, we have not decided yet on how to follow up on our study.

We have revised the manuscript to make it clear that no clinical data were available, and that causes of PWD could therefore not be investigated (Materials and Methods, l.103-104; Discussion, l.254-256).