PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Incidence of adverse events in Sweden during 2013–2016: a cohort study describing the implementation of a national trigger tool.
AUTHORS	Nilsson, Lena; Risberg, Madeleine; Soop, Michael; Nylen, Urban; Ålenius, Carina; Rutberg, Hans

VERSION 1 – REVIEW

REVIEWER	Fernando Vazquez Servicio de Microbiología, Hospital Universitario Central de Asturias and Departamento de Biología Funcional, Area de Microbiología, Facultad de Medicina Oviedo Spain
REVIEW RETURNED	15-Dec-2017

GENERAL COMMENTS	This study uses GTT, at national level, in Sweden and in different hospitals. There are several strengths: -the largest published trigger tool study -data in all the country -data about "off-site" care, there are no data in the literature and the effect on the AEs. This is an important aspect because to check them in a daily basis p.e. could be an strategy -analysis of the triggers: the reduction and the inclusion of new triggers Also there are several weakness: -some weakness are included an in the discussion of the study such as the inter-rater reliability scores -there are a gradual decrease in the rate of AEs but the authors do not wright the change of the hospitals in the period of study (two hospitals stopped reviewing) and the possibility of the changes in the results? Do you know the reason of this aspect? -authors don´t give data about the day when the AEs occur. There are two studies with this aspect Kennerly DA et al. J Patient Saf 2013; 9:1-9 Suárez C et al. JAGS 2014; 62:896-900 Minor aspects: -Abstract: To include the least severe AEs e.i. Categories A-E??? -Which were the criteria to use these minimum monthly number of records 40, 30 and 20? The IHI white paper for example ? -The paragraph "The number of reviewed records was reduced by 50%" is not clear. -What type of chart are using in the paper? Normally in GTT is used

	control chart. And which is the reason in the case the chart is another type? -I think the group of more than 65 years is an heterogeneous group and it is important to stratify in more than 65 and the frail elderly or in decades

REVIEWER	John T. James, PhD Patient Safety America Founder and CEO Houston, TX USA
REVIEW RETURNED	11-Jan-2018

GENERAL COMMENTS	The study provides interesting new insights into applications of the Global Trigger Tool (GTT) as it applies to the detection of adverse events in hospitalized adults. The study utilized review of medical records from almost 65,000 hospital admissions in Sweden from 2013 through 2016. Findings that were particularly interesting to me were the decrease in adverse events in the final 3 years when compared to the rate in 2013, but the decrease did not involve the most severe adverse events. Patients treated in a hospital service inappropriate to their condition (off site) were much more likely to experience an adverse event. The portion of adverse events that were preventable or probably preventable was typically above 60%, with a higher portion of preventable events in adults 65 and older. The occurrence of an adverse event prolonged the hospital stay, especially if the adverse event was deemed preventable. The figures were particularly clear in delivering insights. There are a few points of concern. There was no test of interrater reliability to assess the consistency of detection of adverse events with the GTT. This does not diminish the general quality of the study. On page 8, lines 4 to 8, I cannot follow the calculations. If 39.4% of the adverse events were probably preventable and 22% were certainly preventable, then these two together are 61.4%, not 66.6% as stated on line 8. “Ethics” at the bottom of page 6 seems misplaced. The “P” values on page 8 should be expressed to more significant figures to avoid P = 0.00. In table 2, the authors should extend the percentage numbers to two significant figures. For example, what does 0.1 (0.1-0.1) or 0.2 (0.2-0.3) mean? If 0.25 % of the 65,000 records studied contributed to death or 0.35 % did, then the difference would be the difference between 162 and 227 deaths, respectively. This potential difference is not captured in the tabulated 0.3 %. Page 10, line 1 mentions that “only documented AEs can be identified.” The discussion should build more on this point. As I described when I used 4 GTT studies to estimate harm initiated in U.S. hospitals (https://www.ncbi.nlm.nih.gov/pubmed/23860193), GTTs do miss AEs not documented in the medical record and, this may be a huge factor based on a study by Joel Weisman (https://www.ncbi.nlm.nih.gov/pubmed/18626049). He and his colleagues showed that the ability of trained nurses and physicians to identify known, seriously harmful events from medical records on cardiology patients was limited. In fact, only about 1/3 of the known events were detected in medical records. Part of this is because harmful events initiated during hospitalization may not be apparent until much later. An extreme example of that is offered by Fred Southwick, MD. He describes how the amputation of his leg in 2012 was necessitated because harm was initiated by prolonged use of a pressure cuff 17 years earlier, during surgery on his leg

	(http://www.nytimes.com/2013/02/20/opinion/losing-my-leg-to-a-medical-error.html). It is also well known that directly observed medication errors are often undetectable in medical records (https://www.ncbi.nlm.nih.gov/pubmed/11887410) . Furthermore, the GTT is not efficient at detecting errors of omission (https://www.ncbi.nlm.nih.gov/pubmed/22618920), such as failure to follow evidence-based guidelines. The authors do not want to convey the impression that the vast majority of adverse events initiated during hospitalization have been detected by their GTT. Their study is an important step on the pathway to full understanding of preventable adverse events initiated during hospitalization, but it is far from the last step. The study is worthy of publication with consideration of the points above.

VERSION 1 – AUTHOR RESPONSE

Comments to the author	Author response
Editor Comments to Author:	
- Please rewrite your title. It should state the research question, study design and setting/country. This is the preferred format of the journal.	We have changed accordingly.
Reviewer: 1	
There are several strengths: -the largest published trigger tool study -data in all the country -data about "off-site" care, there are no data in the literature and the effect on the AEs. This is an important aspect because to check them in a daily basis p.e. could be an strategy -analysis of the triggers: the reduction and the inclusion of new triggers	
Also there are several weakness: -some weakness are included an in the discussion of the study such as the inter-rater reliability scores	As we write under methods, there was no assessment of interrater-reliability. We agree that this is a weakness and have pointed at this under limitations in the revised discussion.
-there are a gradual decrease in the rate of AEs but the authors do not wright the change of the hospitals in the period of study (two hospitals stopped reviewing) and the possibility of the changes in the results? Do you know the reason of this aspect?	Both hospitals that stopped reviewing were minor and we are confident that the impact on the results is negligible. We have added information regarding the size of the hospitals.
-authors don´t give data about the day when the AEs occur. There are two studies with this aspect Kennerly DA et al. J Patient Saf 2013; 9:1-9 Suárez C et al. JAGS 2014; 62:896-900	We agree that this is an interesting aspect, especially in connection with length of hospital stay. We did not collect data on day of AE and have added this as a remark in the discussion regarding length of hospital stay.
Minor aspects: -Abstract: To include the least severe AEs e.i. Categories A-E???	The reviewers did not include events in categories A-D. Category E is included in the figures.

-Which were the criteria to use these minimum monthly number of records 40, 30 and 20? The IHI white paper for example ?	The IHI white paper recommends 20-40 records per month, depending on the available resources. We followed this recommendation, and have added a reference to the IHI white paper.
-The paragraph "The number of reviewed records was reduced by 50%" is not clear.	We have rewritten in order to clarify the sentence.
-What type of chart are using in the paper? Normally in GTT is used control chart. And which is the reason in the case the chart is another type?	We have used the types of charts presented in the paper, for example in Fig 1. Statistical process control would have been an alternative, but when referring the results to the review teams and hospital staff we preferred a simple presentation that was easy to understand.
-I think the group of more than 65 years is an heterogeneous group and it is important to stratify in more than 65 and the frail elderly or in decades	We agree that this is a heterogeneous age group. We have added information about rate of AEs in age groups 65-74, 75-84 and ≥ 85 years, and also information about length of stay in connection with AEs.
Reviewer: 2	
The study provides interesting new insights into applications of the Global Trigger Tool (GTT) as it applies to the detection of adverse events in hospitalized adults. The study utilized review of medical records from almost 65,000 hospital admissions in Sweden from 2013 through 2016. Findings that were particularly interesting to me were the decrease in adverse events in the final 3 years when compared to the rate in 2013, but the decrease did not involve the most severe adverse events. Patients treated in a hospital service inappropriate to their condition (off site) were much more likely to experience an adverse event. The portion of adverse events that were preventable or probably preventable was typically above 60%, with a higher portion of preventable events in adults 65 and older. The occurrence of an adverse event prolonged the hospital stay, especially if the adverse event was deemed preventable. The figures were particularly clear in delivering insights.	
There are a few points of concern. There was no test of interrater reliability to assess the consistency of detection of adverse events with the GTT. This does not diminish the general quality of the study.	We agree that this is a weakness and have pointed at this under limitations in the revised discussion.
On page 8, lines 4 to 8, I cannot follow the calculations. If 39.4% of the adverse events were probably preventable and 22% were certainly preventable, then these two together are 61.4%, not 66.6% as stated on line 8.	Thank you for noticing our mistake. This has been corrected.
“Ethics” at the bottom of page 6 seems misplaced.	We agree and have moved the section to a placement after “statistics”

The "P" values on page 8 should be expressed to more significant figures to avoid P = 0.00.	This has been corrected.
In table 2, the authors should extend the percentage numbers to two significant figures. For example, what does 0.1 (0.1-0.1) or 0.2 (0.2-0.3) mean? If 0.25 % of the 65,000 records studied contributed to death or 0.35 % did, then the difference would be the difference between 162 and 227 deaths, respectively. This potential difference is not captured in the tabulated 0.3 %.	This has been corrected.
Page 10, line 1 mentions that "only documented AEs can be identified." The discussion should build more on this point. As I described when I used 4 GTT studies to estimate harm initiated in U.S. hospitals (https://www.ncbi.nlm.nih.gov/pubmed/23860193), GTTs do miss AEs not documented in the medical record and, this may be a huge factor based on a study by Joel Weisman (https://www.ncbi.nlm.nih.gov/pubmed/18626049). He and his colleagues showed that the ability of trained nurses and physicians to identify known, seriously harmful events from medical records on cardiology patients was limited. In fact, only about 1/3 of the known events were detected in medical records. Part of this is because harmful events initiated during hospitalization may not be apparent until much later. An extreme example of that is offered by Fred Southwick, MD. He describes how the amputation of his leg in 2012 was necessitated because harm was initiated by prolonged use of a pressure cuff 17 years earlier, during surgery on his leg (http://www.nytimes.com/2013/02/20/opinion/losing-my-leg-to-a-medical-error.html). It is also well known that directly observed medication errors are often undetectable in medical records (https://www.ncbi.nlm.nih.gov/pubmed/11887410). Furthermore, the GTT is not efficient at detecting errors of omission (https://www.ncbi.nlm.nih.gov/pubmed/22618920), such as failure to follow evidence-based guidelines. The authors do not want to convey the impression that the vast majority of adverse events initiated during hospitalization have been detected by their GTT. Their study is an important step on the pathway to full understanding of preventable adverse events initiated during hospitalization, but it is far from the last step.	Thank you for these important aspects. We agree and have expanded our discussion on the limitations of the retrospective record review method and added some of the suggested references.
The study is worthy of publication with consideration of the points above.	

VERSION 2 – REVIEW

REVIEWER	John T. James

	Patient Safety America USA
REVIEW RETURNED	09-Feb-2018

GENERAL COMMENTS	Thanks for adding the understanding that your approach overlooks adverse events that happen after discharge. Page 12, line 51: has should be have.

REVIEWER	Fernando Vazquez Servicio de Microbiología, Hospital Universitario Central de Asturias and Area de Microbiología, Departamento de Biología Funcional, Facultad de Medicina Oviedo, Spain
REVIEW RETURNED	10-Feb-2018

GENERAL COMMENTS	This is a revised paper and now is OK.

VERSION 2 – AUTHOR RESPONSE

Editor Comments to Author:
Your title could be edited further to improve readability. We suggest something like: 'Incidence of adverse events in Sweden during 2013–2016: a cohort study describing the implementation of a national trigger tool'.
Authors' response: We have changed according to your suggestion.

Reviewer: 1
This is a revised paper and now is OK.

Reviewer: 2
Thanks for adding the understanding that your approach overlooks adverse events that happen after discharge. Page 12, line 51: has should be have.
Authors' response: We have changed accordingly.