

RESEARCH ARTICLE

Adverse event detection by medical record review is reproducible, but the assessment of their preventability is not

Dorthe O. Klein¹*, Roger J. M. W. Rennenberg², Richard P. Koopmans², Martin H. Prins³

1 Department of Clinical Epidemiology and Medical Technology Assessment (KEMTA), Maastricht University Medical Centre+, Maastricht, the Netherlands, **2** Department of Internal Medicine, Maastricht University Medical Centre+, Maastricht, The Netherlands, **3** Department of Epidemiology, School for Public Health and Primary Care, Maastricht University, Maastricht, the Netherlands

* These authors contributed equally to this work.

* dorthe.klein@mumc.nl



Abstract

Objective

To assess the reproducibility of adverse event evaluation by a medical record review committee.

Design

Cross-sectional reanalysis of medical records.

Intervention

Reviewers re-examined fifty medical records of deceased patients regarding the presence of adverse events, their potential preventability and their possible contribution to death. Also we investigated the root causes of the preventable AEs. Differences between the first and second assessment were calculated.

Results

The Kappa on the presence of an adverse event was 0.64 and 0.32 for the potential preventability. The intrarater agreement showed a Kappa of 0.61 on the adverse event presence and 0.64 for the potential preventability. Interrater agreement showed a Kappa of 0.66 for the adverse event presence and 0.03 for the potential preventability.

Conclusion

We found a fair reproducibility for the detection of adverse events, but a poor reproducibility for the potential preventability. Possibly this was caused by lack of a definition for the preventability of adverse events. We think giving feedback to professionals using the results of medical record review remains valuable, but an improvement of its reproducibility is essential. To our opinion an international consensus on what exactly constitutes preventability of

OPEN ACCESS

Citation: Klein DO, Rennenberg RJMW, Koopmans RP, Prins MH (2018) Adverse event detection by medical record review is reproducible, but the assessment of their preventability is not. PLoS ONE 13(11): e0208087. <https://doi.org/10.1371/journal.pone.0208087>

Editor: Kristien Verdonck, Institute of Tropical Medicine Antwerp, BELGIUM

Received: February 28, 2018

Accepted: November 12, 2018

Published: November 29, 2018

Copyright: © 2018 Klein et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

adverse events and agreement on a definition is necessary. This would result in more comparable studies in this field and could then be more informative on the ideal procedure to avoid certain potentially preventable adverse events in the future.

Introduction

The quality and safety of patient care have gained attention in the past decades. Many methods are used with the aim to improve the quality of care. One of these methods is medical record review.

In several university hospitals in the Netherlands the focus of medical record review is on the detection of preventable adverse events (AEs) in patients who have died during their stay. The records of these group of patients are assumed to contain more (preventable) AEs in comparison with discharged (alive) patients.[1]

Often, trigger systems are used for medical record review because these lower the burden of review, by selecting records. Triggers are clues to alert the screeners for potential AEs so the medical record can be reviewed to determine if an actual AE has occurred.[2–5] Nonetheless, the conclusion that a potentially preventable AE has occurred is stressful for the involved professionals. A common reaction is therefore to dispute the judgment of the committee instead of evaluating the case itself and subsequently improving care.[6, 7] To prevent discussions regarding the accuracy of the judgment post hoc, this judgment should be reliable and reproducible.[8–13]

It is likely that if the committee judgment would prove to be reproducible, professionals are more inclined to accept the outcome. There are several studies which investigated the inter-rater reliability using the Harvard medical practice study (HMPS) trigger tool regarding the presence and the preventability of an AE in a mixed (alive and deceased) population. However, these studies showed a moderate reliability for the presence and preventability of an AE.[14–29] Moreover, improving quality and safety of care in an already reasonable safe environment leads to an exponential increase in costs. Therefore, we think health care providers should select the most optimal tests to measure and improve their performance. In order to determine the best methods, qualifying them, making them comparable, and optimising their use, further research is necessary. In our search for better, more reliable methods we hypothesized that our method would show better outcomes compared to previous studies because in our centre the entire review committee discusses the results of the medical record evaluation instead of in pairs or by oneself as is usually reported.[4, 30–32]

Therefore, our main question in this study was: “What is the reproducibility of the judgment of medical records by an internal committee reviewing medical records of deceased patients regarding AE presence. Moreover, we also evaluate the root cause of AEs and their reproducibility of these AEs concerning their preventability and their contribution to death?”

Methods

Medical record review method

Since 2008, a stable team of experienced nurses (each with more than 10 years of clinical experience) with an affinity for record review and patient safety, screen the medical records of all deceased patients for the presence of triggers. Medical record review was introduced in our hospital in 2008 as a project called “preventing medical injury.” The aim of this project was to

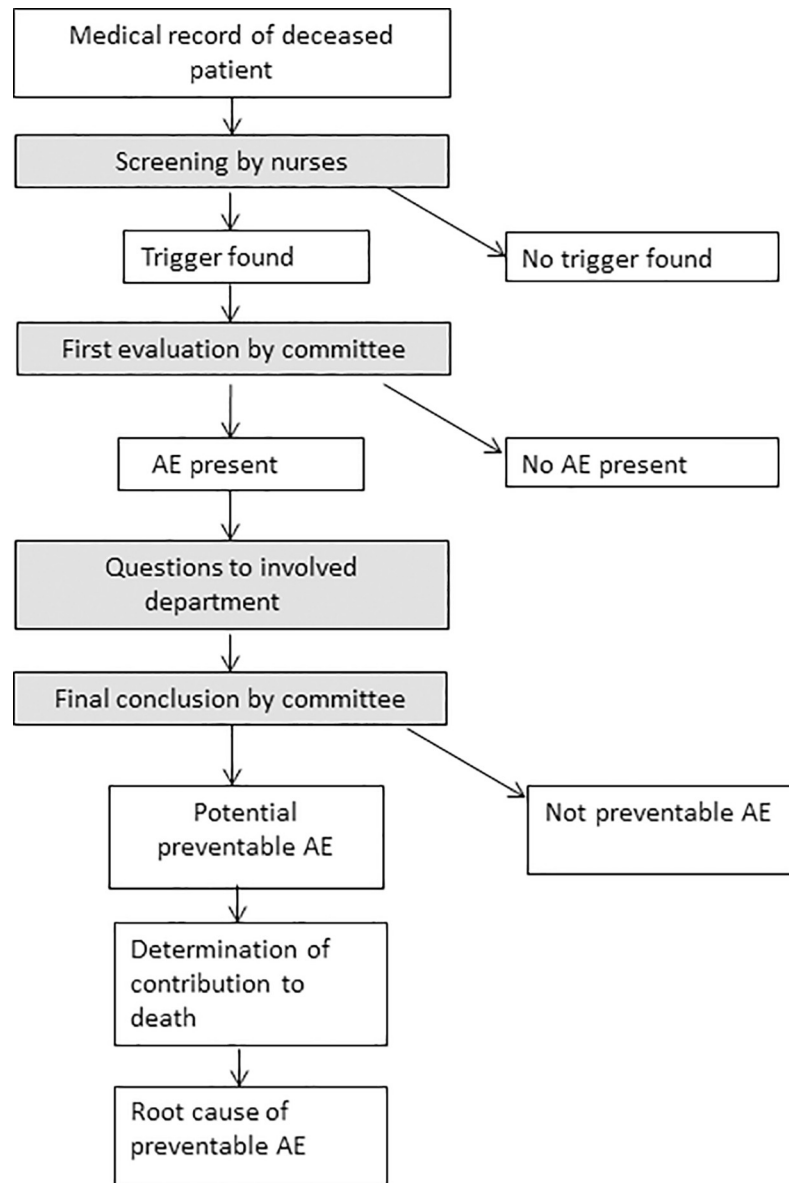


Fig 1. Flowchart of the medical record review procedure.

<https://doi.org/10.1371/journal.pone.0208087.g001>

improve the quality and safety of patient care by reducing and prevention of unintended medical injury happening to patients in our hospital.

We use a slightly adapted version of the HMPS trigger system[16] to make it suitable for the screening of deceased patients. Therefore, triggers regarding transfer to another acute care hospital and unplanned inappropriate discharge to home were omitted, as they have no relevance in deceased patients. In case a medical record has at least one trigger, it is forwarded to the review committee. Further explanation on the trigger system is published in larger detail elsewhere.[33] The procedure of the medical record review is also shown in Fig 1. The study protocol was approved by the medical ethics committee of our hospital. The committee explicitly gave their oral consent to participate in this study and publication of the results.

Definitions

The following definitions were applied in the medical record review:

An AE was defined as an unintended outcome caused by the (non-)action of a caregiver and/ or the healthcare system resulting in temporary or permanent disability or death of the patient. An AE was considered as preventable, if in retrospect after a systematic analysis of the events, it seems that certain measures could have resulted in the prevention of the AE.[34]

Committee

The medical records with a trigger which are forwarded to the review committee, are redirected to the member with experience and expertise concerning the case. For example, if a patient died during a surgical procedure, a surgeon analyses the medical record for AEs. We chose this approach because we believe that the considerations on the presence of an AE are best made by a specialist experienced with the usual procedures, protocols and the possible treatments for a certain disease. The committee consists of 9 specialists with different specialties (internist, neurologist, pulmonologist, pediatrician/neonatologist, internist/oncologist, anesthetist, surgeon, cardiologist, cardiothoracic surgeon and internist/geriatrician) representing the departments with most of the in-hospital deaths. The members are both active (working in our hospital) and recently retired (who used to work in our hospital) experienced specialists.

After a thorough investigation by this committee member, the results are presented to the rest of the committee. A first conclusion on the potential presence of an AE is established in a weekly meeting with at least 4 committee members present. Subsequently, after consulting the involved physicians, the committee finally decides on the presence of an AE, the potential preventability, and the possible contribution of this AE to the death of the patient. If an AE was considered potentially preventable the committee identified the preventable cause according to a standard list of factors ([S1 Appendix](#)).

Data

The results of the triggering by the nurses and the review by the specialists are saved using software provided by Medirede, Clinical File Search version 3 (Mediround BV, 2015). This software was designed to store the data of the medical record review in a clear and easily accessible way.

Selection of records for this study

With regard to our main research question we aimed to get a point estimate with a 95% confidence interval of 6% to each side. Hence, we needed to include 50 cases for the evaluation of the reproducibility of the AE. We randomly selected these medical records out of the database which contained 1262 previously triggered and evaluated medical records of deceased patients between 2013 and 2015. This period was chosen to avoid the presence of recently debated cases in the study sample and therefore minimizing the risk of recognition of cases and remembering former decisions by the committee members. [Fig 2](#) shows the random selection procedure. We used the Excel random generator for the random selection of the study records. Of the 50 randomly selected cases, 25 had an AE after the first evaluation by the committee. We blinded the committee for the previous results of their medical record review. Therefore, the digital results of the first screening round were inaccessible and thus irretrievable for the reviewers.

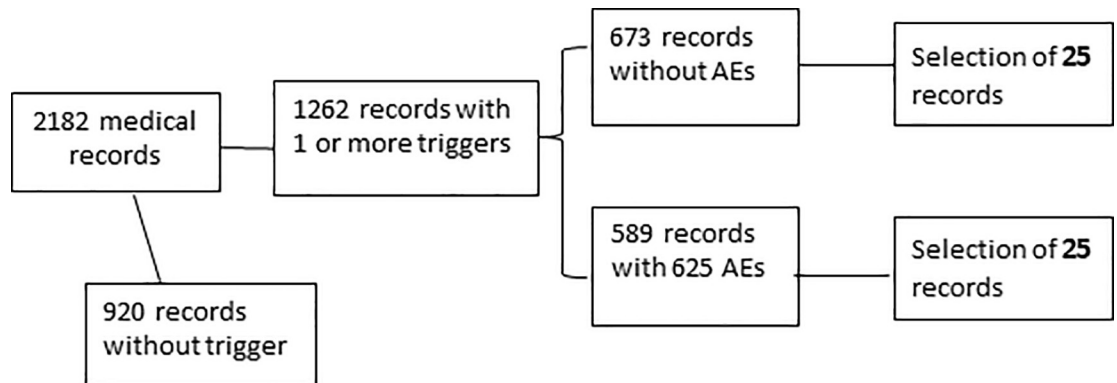


Fig 2. Flowchart of the selection of the medical records in this study.

<https://doi.org/10.1371/journal.pone.0208087.g002>

Analysis

After this re-evaluation of the medical records by the committee, we compared the outcomes of the two rounds regarding the presence of AEs, the preventability and their contribution to the death of the patient on a team level. The preventability of an AE was scored as being potentially preventable or not preventable. The contribution of an AE to the death of the patient was also scored dichotomous as potential contributing to death or no contribution.

Hence, we calculated the observed overall agreement (in %) to get an impression of the reproducibility of the judgment. By using cross tabulation, we calculated the observed overall agreement (accuracy) within the three groups (presence, preventability, and contribution) with the corresponding 95% confidence interval. Therefore, the number of both rounds negative (without AE) and both rounds positive (with AE) was summed and divided by the total number of medical records. Furthermore, we calculated the positive and negative agreement separately. To determine the reliability for AE presence, preventability, and contribution to death, we executed Cohen’s Kappa statistics and the McNemar’s test. A p-value of <0.05 was considered statistically significant. Because the judgment of the committee might be influenced by the presenter of a case, we also evaluated the outcomes of the medical record review after the review and presentation to the committee by different specialists (inter-rater reliability) and by the same specialist (intra-rater reliability). Thereafter, we checked if the AEs found in the second evaluation were the same as the original ones found. Finally, we compared the preventable causes to determine agreement between the first and second evaluation.

Results

Analyses comparing the outcomes of the first, and the second round on a team level

Presence of an AE (Table 1). During the first screening of the medical records (previous to the study), the committee found 25 AEs, the second time (during the study) 28 AEs were

Table 1. AEs found in first and second round.

		Second round		Total
		Absent	Present	
First round	Absent	19	6	25
	Present	3	22	25
Total		22	28	50

<https://doi.org/10.1371/journal.pone.0208087.t001>

Table 2. Potential preventability of the AEs found.

		Second round		Total
		Not preventable	Potentially preventable	
First round	Not preventable	5	1	6
	Potentially preventable	7	10	17
Total		12	11	23

<https://doi.org/10.1371/journal.pone.0208087.t002>

found. There was an overlap of 22 AEs which were found both times. The observed agreement for the presence or absence of an AE was 82.0%. The corresponding Cohen’s Kappa was 0.64 (95%CI 0.48–0.80). McNemar’s test showed a value of 0.51, this means there was no significant difference in the proportion of AEs found in both rounds.

Presence of potential preventability. In the first round, 17 of the 25 AEs had the indication to be potentially preventable, in the second round 11 of the 28 AEs were found to be potentially preventable. The corresponding observed overall agreement of the preventable AEs was 65% (Table 2), and the Cohen’s Kappa was 0.32 (95%CI 0–0.65). McNemar’s test showed a value of 0.070, which indicates there was no significant difference in the proportion of preventability found in both rounds.

Possible contribution to death of the patient. For the calculation of the agreement regarding the possible contribution of the AE to the death of the patient, we compared the cases (22) with the same AEs in both rounds. 21 of these AEs were considered to possibly have contributed to the death of the patient, in the first round. (McNemar = 0.125) During the second screening, the committee concluded that 22 of the 22 AEs possibly contributed to the death of the patient. The accuracy was, therefore, 95% (Table 3).

Intrarater reliability–repeatability (individual level)

31 of the 50 cases were reviewed by the same committee member during the evaluation in both rounds. The observed overall agreement for the presence of an AE in this subgroup was 81%, and the corresponding Kappa was 0.61 (95%CI 0.33–0.89). The agreement on the presence of an AE was 79% and the agreement on the absence of an AE was 82%. McNemar’s test showed a value of 1.000, thus there was no significant difference found in the proportion of AEs in both rounds. For the preventability of the AEs the observed overall agreement was 57%, and the corresponding Kappa 0.64 (95%CI 0.18–1). The agreement on the presence of an AE was 100% and the agreement on the absence of an AE was 60%. An exact McNemar’s test determined that there was no significant difference in terms of the preventability of the AEs found the first and second time ($p = 0.500$). 11 cases in which a possible contribution to death was concluded in the first screening, the second screening found the same result, this means a 100% agreement.

Interrater reliability–reproducibility (individual level)

19 of the 50 medical records were scrutinized by different doctors during their evaluation in both rounds. The observed overall agreement for the presence of an AE was 84%, and the

Table 3. Potential contribution to death in both rounds.

		Second round		Total
		No contribution	Potential contribution	
First round	No contribution	0	1	1
	Potential contribution	0	21	21
Total		0	22	22

<https://doi.org/10.1371/journal.pone.0208087.t003>

Cohen's Kappa was 0.66 (95%CI 0.30–1) in this subgroup. The agreement on the presence of an AE was 79% and the agreement on the absence of an AE was 100%. McNemar's test showed a p-value of 0.250, no significant difference in the proportion of AEs found in both rounds. For the preventability, the observed overall agreement was 45%, and the Cohen's Kappa was 0.03 (95%CI 0–0.55). The agreement on the presence of an AE was 75% and the agreement on the absence of an AE was 29%. An exact McNemar's test determined that there was no significant difference in the proportion of preventability of the AEs found the first and second time ($p = 0.219$). In 10 out of the 11 cases in which a possible contribution to the death of the patient was determined in the first screening, the second screening concluded the same, resulting in 91% agreement.

Root cause analysis

The total number of cases with a potentially preventable adverse event in both sessions, hence labeled with a suspected cause, was nine. The overall agreement on this cause was 78%, with a Kappa of 0.5 (95% CI 0–1). McNemar's test was 1.00 indicating no significant difference.

Discussion

This is the first study assessing the inter- and intrarater reliability for the presence of AEs, their preventability and potential contribution to death among deceased hospital patients. We showed that independent of the committee member involved in the second evaluation, there is a good reliability for the presence of an AE, but only a fair reliability regarding the preventability of an AE. Potential contribution to death was highly reproducible. However, it should be realized that this finding could be distorted by the small number of cases (11) with the same AE in both rounds which were considered to contribute to the death of the patient. Moreover, there was almost no AE that did not contribute to the death of a patient. At the same time, it was difficult to exclude that an AE did not contribute to the death of the patient. It seems, therefore, that the trigger system especially selects cases with AEs that contribute to or cause death. In addition, we couldn't calculate the Kappa for this section due to the low numbers in the cells of the two by two tables.

Our committee found more AE's when scrutinizing the same records for the second time. Therefore, the question remains if the number of AEs found eventually is the true number of AEs. Even if the inter-rater reliability would be acceptable, there is no evidence that this kind of record review really detects all AEs.[14]

A strong point of our study is, in contrast to most other studies, the fact that reviewers discussed the outcomes of the medical record analyses in a group session. In these other studies, usually pairs of reviewers perform the medical record review. A study by Hofer et al (2000) found that discussion between reviewers didn't improve the overall reliability, with a Kappa of 0.36 before and a Kappa of 0.40 after discussion.[35] Interestingly, in our study where cases were discussed twice (during the first and second round), we found a substantially higher level of agreement with a Kappa of 0.64. However, the discussion in our study was between the same committee members both times, whereas in the study of Hofer et al, it was between different pairs of reviewers.

The results from our subgroup analyses suggest a rather low inter-rater reproducibility concerning the preventability of the AEs, in contrast to a higher intra-rater reproducibility. The same phenomenon is also seen in other studies in which both hospital survivors and deceased patients were included. [23–30] Although our system is different in some crucial parts, the judgement concerning the preventability of AEs is comparable with the results in these studies.

Several studies use different classifications for preventability. 3-level classification systems and also a 5- and 6- level classification systems are used in studies.[23, 36, 37] This makes comparisons between studies difficult. There is no gold standard concerning potential preventability which leaves us with the consensus of the medical record research committee as second best. Differences in opinion will therefore always exist, this might give rise to an ongoing discussion. Nevertheless, the effort to detect AEs and their preventability is in our opinion useful because any preventable cause that can be abolished is important, therefore adequate feedback, resulting in consecutive change of procedures if necessary, is needed.

Our study had some limitations. A few methodological aspects of our study require attention. We blinded our reviewers and used records from several years ago, which we consider as a strong point, although we cannot exclude that they may have been able to recognize cases from their memory. Furthermore, there is the lack of a clear definition (which is also not available in international literature) to determine preventability. We think this is the reason why the reliability of the judgment on preventability is disappointing. Although the committee judgment on preventability became based on more experience over the past years, there are no strict guidelines or rules on how to judge. Previous studies sometimes use Likert scales or percentages which in our opinion creates a false sense of precision, with inconsistent reproducibility.[13, 20, 38–42] Secondly, the sample size of the re-examined cases was rather small, especially in the subgroup analysis. This was reflected in the root cause analysis in which only nine cases remained to be analysed. Therefore, we were unable to answer all of our sub questions, especially the one focusing on the reproducibility of the contribution of the AE to the death of the patient. The small number of cases was the result of powering the study to our main question. Moreover, increasing the number of cases to answer our sub-questions with sufficient power would have been very time and hence cost consuming.

In conclusion, we showed that the judgment on the presence of an AE by a committee investigating medical records of deceased patients was reproducible. However, their judgment on preventability is less so, possibly because there is a lack of clear definitions on this subject. It is difficult to comment on the reliability of contribution to death because all AEs seem somehow to contribute to the death of the patient in both rounds. Despite this, we think giving feedback to professionals based on the review of preventable AEs is mandatory. Yet, we should focus on finding more reliable methods to identify preventable AEs. The next step should be, a clear international accepted definition on preventability which could aid to increase the reproducibility of these methods and thus make them more valuable. Thereafter, respectful communication with the involved professionals in order to improve the quality and safety of care is of utmost importance to increase the chance that care will really become better.

Supporting information

S1 Appendix.
(DOCX)

S1 Data.
(XLSX)

Acknowledgments

This study couldn't have been executed without the help of the medical specialists in our team. Therefore, I would like to thank professor Dr. H. Hillen, Dr. P. Bener, Dr. M. van der Hoeven, Dr. P. Hupperets, prof. Dr. P. Kitslaar, Prof. Dr. P. Roekaerts, Prof. Dr. P. Soeters, Dr. C. de Zwaan and for the scrutinization of the medical records.

Author Contributions

Conceptualization: Dorthe O. Klein, Roger J. M. W. Rennenberg, Richard P. Koopmans.

Formal analysis: Dorthe O. Klein.

Investigation: Dorthe O. Klein.

Methodology: Dorthe O. Klein.

Project administration: Dorthe O. Klein.

Supervision: Roger J. M. W. Rennenberg, Richard P. Koopmans.

Writing – original draft: Dorthe O. Klein, Roger J. M. W. Rennenberg.

Writing – review & editing: Dorthe O. Klein, Roger J. M. W. Rennenberg, Richard P. Koopmans, Martin H. Prins.

References

1. Baines RJ, Langelaan M, de Bruijne MC, Wagner C. Is researching adverse events in hospital deaths a good way to describe patient safety in hospitals: a retrospective patient record review study. *Bmj Open*. 2015; 5(7). ARTN e007380 <https://doi.org/10.1136/bmjopen-2014-007380> WOS:000363482000021. PMID: 26159451
2. Brennan TA, Leape LL. Adverse events, negligence in hospitalized patients: results from the Harvard Medical Practice Study. *Perspect Healthc Risk Manage*. 1991; 11(2):2–8. PMID: 10109934.
3. Classen DC LR, Provost L, Griffin FA, Resar R. Development and evaluation of the Institute for Healthcare Improvement Global Trigger Tool. *Journal of Patient Safety* 2008; 4(3):169–77.
4. de Wet C, Bowie P. The preliminary development and testing of a global trigger tool to detect error and patient harm in primary-care records. *Postgrad Med J*. 2009; 85(1002):176–80. <https://doi.org/10.1136/pgmj.2008.075788> PMID: 19417164.
5. Resar RK, Rozich JD, Classen D. Methodology and rationale for the measurement of harm with trigger tools. *Qual Saf Health Care*. 2003; 12 Suppl 2:ii39–45. https://doi.org/10.1136/qhc.12.suppl_2.ii39 PMID: 14645894; PubMed Central PMCID: PMCPMC1765771.
6. Millwood S. Developing a Platform for Learning from Mistakes: changing the culture of patient safety amongst junior doctors. *BMJ Qual Improv Rep*. 2014; 3(1). <https://doi.org/10.1136/bmjquality.u203658.w2114> PMID: 27493733; PubMed Central PMCID: PMCPMC4949613.
7. Wolf ZR, Hughes RG. Error Reporting and Disclosure. In: Hughes RG, editor. *Patient Safety and Quality: An Evidence-Based Handbook for Nurses*. Advances in Patient Safety. Rockville (MD)2008.
8. Contreary K, Collins A, Rich EC. Barriers to evidence-based physician decision-making at the point of care: a narrative literature review. *J Comp Eff Res*. 2016. <https://doi.org/10.2217/cer-2016-0043> PMID: 27935741.
9. He L, Gannon S, Shannon CN, Rocque BG, Riva-Cambrin J, Naftel RP. Surgeon interrater reliability in the endoscopic assessment of cistern scarring and aqueduct patency. *J Neurosurg Pediatr*. 2016; 18(3):320–4. <https://doi.org/10.3171/2016.3.PEDS15648> PMID: 27231825.
10. Mooney MA, Hardesty DA, Sheehy JP, Bird R, Chapple K, White WL, et al. Interrater and intrarater reliability of the Knosp scale for pituitary adenoma grading. *J Neurosurg*. 2016:1–6. <https://doi.org/10.3171/2015.9.JNS152183.test> PubMed PMID: 28306407.
11. Ramnarayan P, Kapoor RR, Coren M, Nanduri V, Tomlinson AL, Taylor PM, et al. Measuring the impact of diagnostic decision support on the quality of clinical decision making: development of a reliable and valid composite score. *J Am Med Inform Assoc*. 2003; 10(6):563–72. <https://doi.org/10.1197/jamia.M1338> PMID: 12925549; PubMed Central PMCID: PMCPMC264434.
12. Coutts SB, Simon JE, Tomanek AI, Barber PA, Chan J, Hudon ME, et al. Reliability of assessing percentage of diffusion-perfusion mismatch. *Stroke*. 2003; 34(7):1681–3. <https://doi.org/10.1161/01.STR.0000078840.96473.20> PMID: 12805485.
13. Sharek PJ, Parry G, Goldmann D, Bones K, Hackbarth A, Resar R, et al. Performance characteristics of a methodology to quantify adverse events over time in hospitalized patients. *Health Serv Res*. 2011; 46(2):654–78. <https://doi.org/10.1111/j.1475-6773.2010.01156.x> PMID: 20722749; PubMed Central PMCID: PMCPMC3064924.

14. Hanskamp-Sebregts M, Zegers M, Vincent C, van Gorp PJ, de Vet HC, Wollersheim H. Measurement of patient safety: a systematic review of the reliability and validity of adverse event detection with record review. *Bmj Open*. 2016; 6(8):e011078. <https://doi.org/10.1136/bmjopen-2016-011078> PMID: 27550650; PubMed Central PMCID: PMC5013509.
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33(1):159–74. PMID: 843571.
16. Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *N Engl J Med*. 1991; 324(6):370–6. <https://doi.org/10.1056/NEJM199102073240604> PMID: 1987460.
17. Sari AB, Sheldon TA, Cracknell A, Turnbull A. Sensitivity of routine system for reporting patient safety incidents in an NHS hospital: retrospective patient case note review. *BMJ (Clinical research ed)*. 2007; 334(7584):79. <https://doi.org/10.1136/bmj.39031.507153.AE> PMID: 17175566; PubMed Central PMCID: PMC1767248.
18. Brennan TA, Localio RJ, Laird NL. Reliability and validity of judgments concerning adverse events suffered by hospitalized patients. *Med Care*. 1989; 27(12):1148–58. PMID: 2593729.
19. Baines RJ, Langelaan M, de Bruijne MC, Asscheman H, Spreeuwenberg P, van de Steeg L, et al. Changes in adverse event rates in hospitals over time: a longitudinal retrospective patient record review study. *BMJ Qual Saf*. 2013; 22(4):290–8. <https://doi.org/10.1136/bmjqs-2012-001126> PMID: 23293136.
20. Zegers M, de Bruijne MC, Wagner C, Groenewegen PP, van der Wal G, de Vet HC. The inter-rater agreement of retrospective assessments of adverse events does not improve with two reviewers per patient record. *J Clin Epidemiol*. 2010; 63(1):94–102. <https://doi.org/10.1016/j.jclinepi.2009.03.004> PMID: 19473812.
21. Thomas EJ, Studdert DM, Burstin HR, Orav EJ, Zeena T, Williams EJ, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. *Medical care*. 2000; 38(3):261–71. PMID: 10718351.
22. Localio AR, Weaver SL, Landis JR, Lawthers AG, Brenhan TA, Hebert L, et al. Identifying adverse events caused by medical care: degree of physician agreement in a retrospective chart review. *Annals of internal medicine*. 1996; 125(6):457–64. Epub 1996/09/15. PMID: 8779457.
23. Hogan H, Healey F, Neale G, Thomson R, Vincent C, Black N. Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study. *BMJ Qual Saf*. 2012; 21(9):737–45. <https://doi.org/10.1136/bmjqs-2011-001159> PMID: 22927487; PubMed Central PMCID: PMC3436096.
24. Soop M, Fryksmark U, Koster M, Haglund B. The incidence of adverse events in Swedish hospitals: A retrospective medical record review study. *International Journal for Quality in Health Care*. 2009; 21(4):285–91. <http://dx.doi.org/10.1093/intqhc/mzp025>. PMID: 19556405.
25. Forster AJ, Asmis TR, Clark HD, Al Saied G, Code CC, Caughey SC, et al. Ottawa Hospital Patient Safety Study: incidence and timing of adverse events in patients admitted to a Canadian teaching hospital. *CMAJ: Canadian Medical Association journal = journal de l'Association medicale canadienne*. 2004; 170(8):1235–40. Epub 2004/04/14. <https://doi.org/10.1503/cmaj.1030683> PMID: 15078845; PubMed Central PMCID: PMC385353.
26. Baker GR, Norton PG, Flintoft V, Blais R, Brown A, Cox J, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ*. 2004; 170(11):1678–86. <https://doi.org/10.1503/cmaj.1040498> PMID: 15159366; PubMed Central PMCID: PMC408508.
27. Davis P. LYR, Briant R. Adverse events in New Zealand public hospitals: principal findings from a national survey. New Zealand: Ministry of Health 2001.
28. Thomas EJ, Lipsitz SR, Studdert DM, Brennan TA. The reliability of medical record review for estimating adverse event rates. *Ann Intern Med*. 2002; 136(11):812–6. PMID: 12044129.
29. Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L, Hamilton JD. The Quality in Australian Health Care Study. *Med J Aust*. 1995; 163(9):458–71. PMID: 7476634.
30. Mortaro A, Moretti F, Pasco D, Tessari L, Tardivo S, Pancheri S, et al. Adverse Events Detection Through Global Trigger Tool Methodology: Results From a 5-Year Study in an Italian Hospital and Opportunities to Improve Interrater Reliability. *J Patient Saf*. 2017. <https://doi.org/10.1097/PTS.0000000000000381> PMID: 28598897.
31. Rutberg H, Borgstedt Risberg M, Sjodahl R, Nordqvist P, Valter L, Nilsson L. Characterisations of adverse events detected in a university hospital: a 4-year study using the Global Trigger Tool method. *Bmj Open*. 2014; 4(5):e004879. <https://doi.org/10.1136/bmjopen-2014-004879> PMID: 24871538; PubMed Central PMCID: PMC4039822.
32. Najjar S, Hamdan M, Euwema MC, Vleugels A, Sermeus W, Massoud R, et al. The Global Trigger Tool shows that one out of seven patients suffers harm in Palestinian hospitals: challenges for launching a

- strategic safety plan. *Int J Qual Health Care*. 2013; 25(6):640–7. <https://doi.org/10.1093/intqhc/mzt066> PMID: 24141012.
33. Klein DO, Rennenberg R, Koopmans RP, Prins MH. The ability of triggers to retrospectively predict potentially preventable adverse events in a sample of deceased patients. *Prev Med Rep*. 2017; 8:250–5. <https://doi.org/10.1016/j.pmedr.2017.10.016> PMID: 29181297; PubMed Central PMCID: PMC5700821.
 34. vdWG Wagner C. Voor een goed begrip. Bevordering patiëntveiligheid vraagt om heldere definities [For a good understanding. Improving patient safety requires clear definitions] *Med contact*. 2005; 60:1888–91.
 35. Hofer TP, Bernstein SJ, DeMonner S, Hayward RA. Discussion between reviewers does not improve reliability of peer review of hospital quality. *Med Care*. 2000; 38(2):152–61. PMID: 10659689.
 36. Manaseki-Holland S, Lilford RJ, Bishop JRB, Girling AJ, Chen YF, Chilton PJ, et al. Reviewing deaths in British and US hospitals: a study of two scales for assessing preventability. *BMJ Qual Saf*. 2017; 26(5):408–16. <https://doi.org/10.1136/bmjqs-2015-004849> PMID: 27334868.
 37. Teixeira PG, Inaba K, Hadjizacharia P, Brown C, Salim A, Rhee P, et al. Preventable or potentially preventable mortality at a mature trauma center. *J Trauma*. 2007; 63(6):1338–46; discussion 46–7. <https://doi.org/10.1097/TA.0b013e31815078ae> PMID: 18212658.
 38. Landrigan CP, Parry GJ, Bones CB, Hackbarth AD, Goldmann DA, Sharek PJ. Temporal trends in rates of patient harm resulting from medical care. *N Engl J Med*. 2010; 363(22):2124–34. <https://doi.org/10.1056/NEJMsa1004404> PMID: 21105794.
 39. O'Leary KJ, Devisetty VK, Patel AR, Malkenson D, Sama P, Thompson WK, et al. Comparison of traditional trigger tool to data warehouse based screening for identifying hospital adverse events. *BMJ Qual Saf*. 2013; 22(2):130–8. <https://doi.org/10.1136/bmjqs-2012-001102> PMID: 23038408.
 40. Zwaan L, De Bruijne M, Wagner C, Thijs A, Smits M, Van Der Wal G, et al. Patient record review of the incidence, consequences, and causes of diagnostic adverse events. *Archives of internal medicine*. 2010; 170(12):1015–21. <https://doi.org/10.1001/archinternmed.2010.146> PMID: 20585065.
 41. Mirza SK, Deyo RA, Heagerty PJ, Turner JA, Lee LA, Goodkin R. Towards standardized measurement of adverse events in spine surgery: conceptual model and pilot evaluation. *BMC Musculoskelet Disord*. 2006; 7:53. <https://doi.org/10.1186/1471-2474-7-53> PMID: 16787537; PubMed Central PMCID: PMC5700821.
 42. Kessomboon P, Panarunothai S, Wongkanaratanakul P. Detecting adverse events in Thai hospitals using medical record reviews: agreement among reviewers. *J Med Assoc Thai*. 2005; 88(10):1412–8. PMID: 16519388.