OPEN

A Systematic Review of Methods for Medical Record Analysis to Detect Adverse Events in Hospitalized Patients

Dorthe O. Klein, MSc, PhD,* Roger J.M.W. Rennenberg, MD, PhD,†
Richard P. Koopmans, MD, PhD,† and Martin H. Prins, MD, PhD‡

Objective: In this systematic review, we evaluate 2 of the most used trigger tools according to the criteria of the World Health Organization for evaluating methods.

Methods: We searched Embase, PubMed, and Cochrane databases for studies (2000–2017). Studies were included if medical record review (MRR) was performed with either the Global Trigger Tool or the Harvard Medical Practice Study in a hospital population. Quality assessment was performed in duplicate. Fifty studies were included, and results were reported for every criterion separately.

Results: Medical record review reveals more adverse events (AEs) than any other method. However, at the same time, it detects different AEs. The costs of an AE were on average €4296. Considerable efforts have been made worldwide in health care to improve safety and to reduce errors. These have resulted in some positive effects. The literature showed that MRR is focused on several domains of quality of care and seems suitable for both small and large cohorts. Furthermore, we found a moderate to substantial agreement for the presence of a trigger and a moderate to good agreement for the presence of an AE.

Conclusions: Medical record review with a trigger tool is a reasonably well-researched method for the evaluation of the medical records for AEs. However, looking at the World Health Organization criteria, much research is still lacking or of moderate quality. Especially for the cost of detecting AEs, valuable information is missing. Moreover, knowledge of how MRR changes quality and safety of care should be evaluated.

Key Words: trigger tool, adverse event, patient safety, Global Trigger Tool, hospital

(*J Patient Saf* 2021;17: e1234–e1240)

Several studies have shown significant rates of adverse events (AEs) that cause harm to patients during their stay in the hospital.[1–3] Therefore, interest in the implementation of quality and safety programs that prevent these harmful events has grown. According to the report *To Err Is Human* (1992) of the Institute of Medicine, at least 44,000 people (and possible even 98,000) died of AEs each year in U.S. hospitals.[3] An update, 15 years after this first report, showed little improvement and stresses the need for further improving patient safety.[4] In the Netherlands, a report by the Dutch Institute for Research in Healthcare evaluated care-related harm in Dutch hospitals. It was estimated that yearly approximately 1700 patients (4.1% of the total number of inpatient deaths) die of unintentional preventable harm. Follow-up after 6 years showed slight improvement, but still 2.6% of the inpatient deaths seemed preventable. In other countries, an incidence between 2.5% and 13.5% has been found.[5–10]

To diminish care-related harm, there are several instruments to detect AEs: direct observation, incident reporting systems, autopsy reports, and mortality and morbidity conferences are some examples.[11–14] For medical record review (MRR), trigger tools are often used to prevent time- and cost-consuming investigation of all records. The most well-known trigger systems are the Harvard Medical Practice Study (HMPS) trigger system[15] with 18 triggers[15,16] and the Global Trigger Tool (GTT), developed by the Institute for Healthcare Improvement (IHI) with 54 triggers.[17]

Information on sensitivity, specificity, and positive and negative predictive values of these screening tools for detecting AEs is important in striving for effective and affordable methods to detect AEs. Also, MRR, independently of the method used, is a costly process because it takes time of both experienced physicians and nurses; hence, it is of importance that the process is indeed improving quality and safety.

However, adequate detection of AEs is only a small part of the review process. The total procedure also involves feedback to the medical departments, adjustments in the delivery of care, and hence improved outcome for patients resulting in less AEs. Other recent reviews highlighted MRR on the reliability perspective[18] or describe the use and potential value of the GTT.[19,20] Our aim was to describe the 2 most used trigger tools (GTT and HMPS) with the whole process in mind, starting with the detection of the triggers to the feedback to the departments and the effect on the patient safety. We were interested in all stages of this screening method. The World Health Organization (WHO) criteria for evaluating methods offer, in our opinion, a suitable framework to analyze all stages of this screening process. Therefore, we investigated the total screening process based on MRR and if this is evidence based according to the WHO criteria for evaluating methods.[21] With these criteria in mind, we searched the literature for evidence about the use of this specific method to improve patient safety.

METHODS

Search Strategy and Information Sources

We identified potentially eligible studies by searching PubMed, Embase, and the Cochrane Library for every criterion (and the explanation) described in Box 1. Our search was restricted to studies in English or Dutch published between 2000 and 2017 because we assumed that older results might not be applicable because of rapidly changing health care. The search strategy and corresponding search terms are shown in Table S1 (Supplemental Digital Content 1, http://links.lww.com/JPS/A263). The flowchart for every criterion separately is shown in Figure S2 (Supplemental Digital Content 2, http://links.lww.com/JPS/A264).

Explanation of the WHO criteria (Box 1)

1. Effectiveness in capturing the extent of harm (in different environments)
   Comparative studies, presenting the numbers of AEs found with MRR and another method
2. Availability of reliable data
   Interrater reliability of MRR
3. Suitability for large-scale or small, repeated studies
   Acceptability and feasibility of MRR by institutions and professionals
4. Costs (financial, human resources, time, and burden on system)
5. Effectiveness in influencing policy
   Effect of (the results of) MRR on national, regional, or local policy
6. Effectiveness in influencing hospital and local safety procedures and outcomes
   Intervention studies that evaluate the effect of MRR on daily practice and the AEs
7. Synergy with other domains of quality of care
   The IHI defined 6 domains of quality of care (safe, effective, patient-centered, efficient, timely, and equitable).[22] In this criterion was described whether MRR can be of influence on several of these domains.

Selection Criteria and Process

Studies were included if MRR was performed with either the GTT or the HMPS in a hospital population with a wide variety of patient groups. Suitable study designs were observational studies. Reviews, posters, comments, studies solely focusing on adverse drug events, and patient younger than 18 years were also excluded. Furthermore, we excluded studies that used computer detection for finding triggers and/or AEs. Duplicate references were removed using the software program EndNote (EndNote X8; Thomson Reuters, New York, New York). Afterward, all retrieved citations based on the titles and abstracts were screened. Study selection was performed in duplicate by 2 independent reviewers (D.O.K. and R.J.M.W.R.) according to the aforementioned inclusion and exclusion criteria. Study eligibility was thereafter assessed by reading the full text. Disagreements were resolved by discussion and consensus in which both reviewers had an equal vote.

Quality Assessment

The quality of the included studies was evaluated using the Gilbert criteria complemented with criteria composed by Worster and Badcock. These criteria have been developed to assess the quality of MRR studies and include, for example, the following: assessment on training of abstractors, presence of abstraction forms, and evaluation of the interrater agreement.[23–25] Quality assessment was performed in duplicate by D.O.K. and R.J.M.W.R. Rating categories used in the assessment were "present" or "missing." These were transformed into 1 and 0 and, after scoring the numbers, were added together, resulting in a score between 1 and 15. Scores between 0 and 5 were seen as weak; between 6 and 10, reasonable; and between 11 and 15, good. The results are shown in Table S3 (Supplemental Digital Content 3, http://links.lww.com/JPS/A266).

Data Extraction and Analysis

Data extraction was performed by D.O.K. General data extracted from full-text included the following: author, year of publication, country, study aim, study design, instrument(s), sample characteristics, number of screening criteria, number of reviewers, and number of records analyzed for the measurement of the reliability. Extracted outcome data included the following: prevalence of AEs, κ agreement for AEs, percentage agreement, summary of main results, and costs of AEs (Table S4, Supplemental Digital Content 4, http://links.lww.com/JPS/A265) Using basic descriptive statistics (mean with confidence intervals [CIs] if available), we summarized these variables. Costs were not corrected for inflation. Different currencies were transformed to euros to make comparison easier. We used the exchange of August 2018 (1€ = 1.13 U.S.$). The results of this systematic review are presented according to the *Preferred Reporting Items for Systematic Reviews and Meta-analyses* guidelines.[26] (Supplemental Digital Content 5, http://links.lww.com/JPS/A267).

RESULTS

Our search resulted in a total of 832 citations (steps 1–4), and after title and abstract screening, 752 were discarded, leaving 80 full-text articles to assess their eligibility. Another 63 were excluded after full-text evaluation, leaving 17 for inclusion. After reading the references in these studies, we found 76 additional studies. The included studies originated from 24 countries, and sample size varied from 96[27] to 210 million[28] patients. Some studies were relevant for more than 1 criterion. Hereinafter, we report the results for every criterion separately and according to the trigger tool (GTT or HMPS). The average quality of the included studies was reasonable (score 8.3).

Criterion 1. Effectiveness in Capturing the Extent of Harm (in Different Environments)

Sixteen studies of reasonable quality were identified, which compare the use of the MRR with the use of another method to detect AEs in a hospital setting. A wide diversity of methods was compared with the GTT, such as reporting notification systems,[29,30] a hospital survey on patient safety,[31] patient safety indicators (PSIs),[29,32] and complaints and claims by patients and their relative and incident reports.[33,34] For the HMPS, the outcomes were primarily compared with incident reports, as well as patient-reported AEs[35] and patient complaints.[36]

GTT

Kurutkan et al[30] found that the GTT was 19 times more sensitive compared with internal voluntary reporting for the detection of AEs. Another study by Farup[31] detected an inverse association between the patient safety culture survey and AEs. Kennerly et al[29] found that voluntary reports and PSIs captured less than 5% of the total AEs, and this was also found by other studies.[32,37] Rutberg et al[34] found that 6.3% of the AEs detected by the GTT were reported with voluntary reporting. Mull et al[33] compared 3 methods (quality improvement programs, PSIs, and voluntary incident reporting) with GTT and concluded that 12% of the AEs were also detected by one of the other methods.

Harvard Medical Practice Study

Christiaans-Dingelhoff et al[38] linked 4 reporting systems with MRR: informal and formal complaints by patients or relatives, medicolegal claims by patients or relatives, and incident reports by health care professionals. Less than 4% of the AEs identified by record review were found in at least 1 of the 4 reporting systems.

Michel et al[14] compared 3 methods: cross-sectional, prospective, and retrospective review of records. The prospective and retrospective methods identified similar numbers of AEs (70% versus 66%

of the total), but the prospective method detected more preventable AEs (64% versus 40%). The cross-sectional method showed a large number of false positives and identified none of the most serious AEs.

Several studies compared the number of AEs found when using incident reports compared with the use of HMPS. Blais et al[39] showed that in 15.5% of the cases with an AE, an incident report was present. Also, Sari et al[40] detected underreporting by voluntary incident reporting; only 24% of all patient safety incidents and only 5% of those resulting in patient harm were detected with the HMPS method. de Feijter et al[36] compared both patient complaints, MRR and incident reporting systems; they found that the type of AE found was dependent on the method used. Therefore, they recommend using a combination of methods when assessing patient safety in a hospital. The findings by Macharia et al[41] were in line with the other studies.

Weissman et al[35] found that 23% of the patients had at least 1 patient-reported AE and 11% according to MRR. The agreement between these 2 were poor for occurrence of any type of AE and slightly better for life-threatening or serious AEs. Bjertnaes et al[42] showed a significant correlation between the 2 measurement methods.

Criterion 2. Availability of Reliable Data

Twenty-seven suitable studies of reasonable quality regarding the availability of reliable data in studies that used a trigger tool method for MRR were found.

Trigger Tool

The IHI method showed a positive predictive value of 30.4% (95% CI, 13.3%–47.6%).[30,32,43–48] There was only one study that calculated the negative predictive value, which was 99%.[32] The agreement on the presence of a trigger had on average a moderate agreement (κ = 0.48; 95% CI, 0.18–0.77).[2,44–46,48,49]

In studies using the HMPS, the positive predictive value was on average 33.4% (95% CI, 21.1%–45.8%).[5–7,9,40,50–56] The agreement between the nurses on the presence of a trigger showed on average a substantial agreement (κ = 0.63; 95% CI, 0.55–0.71).[5,40,51,52,54,55]

The AE Assessment Strategy

Within the studies using the IHI trigger tool,[17] the κ on the presence of an AE was on average 0.67 (good agreement; 95% CI, 0.56–0.82).[1,2,30,37,43–46,49,57–60] In 3 studies, the agreement on the severity of the AE was investigated, which showed an average κ of 0.40 (fair agreement; 95% CI, 0.06–0.73).[1,43,58,59]

In HMPS studies, there was a moderate agreement between medical doctors on the presence of an AE in the medical record, with a κ value on average of 0.58 (95% CI, 0.33–0.82).[5,9,40,48,52–54,61]

Criterion 3. Suitability for Large-Scale or Small, Repeated Studies

Forty-three suitable studies of reasonable quality regarding the suitability of MRR for large-scale or small, repeated studies were found.

In the last decades, several large-scale studies have been executed to assess the prevalence of AEs in hospitals on a national level or the financial impact of AEs. The smaller studies were used for the training of the reviewers,[49] for comparison with the detection rate of other methods,[29,32,37,38,62] or for assessing the interrater reliability[45] of the review.

Size

The size of the studies varied from 15[49] (training) records to 40,851,[63] and the number of hospitals investigated also varied from 1 to 25 hospitals.

Cross-Country Comparison

Deilkas et al[64] compared the AE rate between Norway and Sweden. Norway had significantly higher AE rates of surgical complications. Swedish hospitals had significantly higher rates of pressure ulcers, falls, and "other" AEs. No significant difference between overall AE rates was found between the 2 countries.[64]

Criterion 4. Costs (Financial, Human Resources, Time, and Burden on System)

Of the 12 studies found concerning the costs of AEs, 6 reported these results based on MRR; the quality of these 6 studies was reasonable. Others reported the costs based on a hospital claim database,[28] a national medical and drugs claim database, and a hospital cost accounting system.[65] Furthermore, some studies used consensus-based methodology, diagnostic coding error,[66] and estimation of social cost.[67] The costs of an AE were on average €4296 (range, €2600–€6436).[6,55,60,68–70]

Next to the cost of an AE itself (the clinical consequences and the patient harm), the cost of finding an AE using MRR is also important. With this we mean the cost of nurses for trigger analysis, and physicians for the AE detection in records and their administration. Keep in mind that for detecting one AE, many records without AEs have to be analyzed. No studies on the cost of detecting an AE were available. Based on our own data (unpublished), the cost of detecting an AE with HMPS was €150.000 on a yearly basis, and that for a single potentially preventable AE was approximately €1800.

Criterion 5. Effectiveness in Influencing Policy

We found no trials or studies but only reports of projects concerning this issue.

Short Overview of Studies Mentioned in the Context of the European Union Network for Patient Safety and Quality of Care

In the Netherlands, the reports by the Dutch Institute for Research in Healthcare have been published in the context of a research program named patient safety in the Netherlands. The first study was first performed to gain insight into AEs in Dutch hospitals.[54] The 2 follow-up studies were executed to evaluate whether the safety programs had a positive influence on these AEs.[71,72]

Based on experience in Norway and Sweden, MRR by the GTT method gives a valuable overview of the kind and incidence of AEs affecting patients and a good starting point for intensified patient safety improvement work.[63,73]

Criterion 6. Effectiveness in Influencing Hospital and Local Safety Procedures and Outcomes

Our search revealed 5 studies investigating changes in AE rates during the study period.

Kennerly et al[46] showed a 7% reduction in AEs in 2 years (on average, 3.5% per year). Suarez et al[74] found during a 6-year study period a decrease of 2.5% (on average, 0.4% per year). Deilkas et al[63] showed that AE rates decreased from 16.1% to 13.0% in 2 years (1.55% per year).

However, Rutberg et al[34] found no improvement during the 4-year study period in which the GTT was used, despite several initiatives for improving the quality in the hospital, similar to Landrigan et al[1] and Mortaro et al.[75] Three national studies in the Netherlands (2004–2012) showed no changes in overall AE but did show a decrease of 45% regarding preventable AEs.[71] However, the latest update in 2017 did not show a further decrease of the preventable AEs and preventable deaths but did show a 2%

decrease in overall AE.[76] Landrigan et al[1] also found slight improvements in a 5-year period in the United States.

Criterion 7. Synergy With Other Domains of Quality of Care

We found no trials or studies but only reports of projects concerning this issue. The IHI has defined 6 domains of quality of care.[22] Medical record review has common ground with a few of these domains; safe, effective, timely, and equitable. During MRR, a committee assesses whether AEs have occurred. The goal is to improve care, making it safer. Furthermore, it is evaluated whether the specific treatment for a particular patient was correct, right on time (effective and timely), and independent of personal characteristics (equitable). The project "Deepening our Understanding of Quality Improvement in Europe" investigated the relation between quality systems and patient-related outcomes. Almost 200 hospitals in 8 European countries participated in the Deepening our Understanding of Quality Improvement in Europe project. Beside questionnaires, MRR and data registries were analyzed. One of the conclusions of this project was that the presence of quality systems has a positive effect on the safety culture in a hospital.[77]

DISCUSSION

Our study clearly shows that there is abundant literature concerning MRR in hospitals. However, almost 75% of these articles were of rather moderate quality (steps 1–4). The first 4 WHO criteria relate to the characteristics of MRR concerning validity, reliability, and costs. They could effortlessly be extracted from the existing literature. The last 3 criteria relate to the ability of MRR to generate improvements in safety procedures and the quality of safety programs. Data concerning criteria 5 and 7 were indirectly described or concealed in reports, and therefore, we were unable to evaluate the quality according to the quality checklist.

The literature we found in relation to the first criterion showed that MRR reveals more AEs than any other method. However, at the same time, MRR detects different AEs compared with other methods. Furthermore, we found a moderate agreement for the presence of a trigger and a good agreement for the presence of an AE for the GTT. For the HMPS, we found a substantial agreement for the presence of a trigger and a moderate agreement for the presence of an AE. Also, MRR seems suitable for both small and large cohorts as shown in several studies with different sample sizes.

The costs concerning AEs can be the cost of the event itself (which is usually the topic of the literature we found), but also the cost of MRR. It is striking that most studies investigating costs of AEs only evaluated costs related to the event. The only study we found evaluating also the costs of the detection method was published by Bates et al,[78] but was not included in the current study. The costs for the detection of a single AE in 1995 was 103€, and that for a preventable AE was 241€ (€11.10 for every admission). Translating this to the current situation means a considerable amount of money for the detection instrument, let alone the costs of the AEs themselves. Because there is no agreement on which costs exactly should be taken into account concerning AEs, comparison between studies is difficult. For future research, it is important to have a complete overview of all costs involved. It should contain not only the cost of the detection instrument and the direct cause of the AE itself, but also loss of working days, up to the implementation of other protocols and their costs to prevent AEs. Only with these total costs we will be able to estimate the costs per quality-adjusted life-year to see if this is acceptable. The aforementioned is also underlined in a report by Øvretveit[79]

and a more recent report by the Organization for Economic Co-operation and Development.

The results of the fifth criterion show that MRR can have an effect on health care policy. In Europe, the Network for Patient Safety and Quality of Care has been active for 4 years. This network was cofunded and supported by the European Commission within the Public Health Program. Its focus was "to improve **Patient Safety and Quality of Care in Europe** by supporting the implementation of good organizational practices and safe clinical practices in health care organizations and through sharing of information and experiences." The Network for Patient Safety and Quality of Care builds on the experience of the European Union Network for Patient Safety project (2008–2010), which established patient safety platforms in several European member states. The main outcome will be the consolidation of the *permanent network for patient safety.*

Considerable efforts have been made worldwide in the health care systems to improve safety and to reduce errors in the treatment for patients. As is shown in criterion 6, these efforts have translated into only slight improvements in the overall safety of patients or in better quality of care. Finally, the last criterion shows that MRR has focused on several domains of quality of care.

Michel[80] has composed an overview of strength and weaknesses of available methods for assessing AEs. Medical record review is one of these methods. Since then (2003), no update has been performed to investigate how well the trigger tools comply according to the WHO criteria. Instead of giving an overview of the available methods, we decided to focus on the 2 most used trigger tools. Also, in this overview, Michel did not give an insight into the quality of the included studies, rather compared the methods with each other. Furthermore, he evaluated the evidence-based rating of the methods for estimating AEs, for each criterion.

Our systematic review had several strengths and some limitations. We used extensive search criteria (Table S1, Supplemental Digital Content 1, http://links.lww.com/JPS/A263) in several databases and also screened the references of the finally selected articles, which, to our opinion, minimizes the risk of missing important literature. Moreover, we used an accepted WHO strategy to evaluate this screening tool. Furthermore, we combined several quality checklists for the assessment of the quality of the included studies.

An important limitation for drawing aggregate conclusions was the different methods studies deployed. Although the same trigger tool method was used, almost every study adapted the triggers or the review process slightly. Besides that, different definitions and scales were used for both AEs and their preventability. Moreover, some studies used external reviewers, some internal reviewers, or experienced reviewers. Although the intention was to evaluate the quality of all included studies, this was only possible for 5 of 7 WHO criteria. The reason was that the others were not suitable for evaluation according to our 15 quality requirements. These were actually more reports than studies. For other studies, we found that important detailed information was not reported (e.g., information on positive and negative agreement, and reproducibility regarding the individual triggers). We only searched for studies from the year 2000 and onward. There is a chance that important older studies are therefore missed. However, we doubt that these findings would still be generalizable to today's care because of rapidly developing health care and quality and safety improvements. Also, we were forced to write a systematic narrative review; because of the number of different studies, we were not able to combine the numbers and create a meta-analysis.

Between 2000 and 2017, many hospitals have switched from paper medical records to electronic medical records. In this period, many hospitals switched from paper health records to electronic health records. Electronic health records have several advantages

for MRR because the records are accessible, readable, and easier to process.[81–83] A next step will be automatic detection of triggers and AEs, but at the moment, this is not optimal.[84–88]

In conclusion, MRR with one of the trigger tools is a well-researched method for the evaluation of medical records for AEs. However, looking at the WHO criteria for the evaluation of methods, much research is still lacking or of moderate quality. More information concerning costs of the detection method and the improvements in care and patient outcome is needed. Only with this information that MRR could be evaluated on its cost-effectiveness. Moreover, more insight into how MRR changes quality and safety of care is needed. We found no studies analyzing the whole string that starts with the application of the triggers and ends with quality and safety improvement for individual patients at acceptable costs.

REFERENCES

1. Landrigan CP, Parry GJ, Bones CB, et al. Temporal trends in rates of patient harm resulting from medical care. *N Engl J Med*. 2010;363: 2124–2134.

2. Najjar S, Hamdan M, Euwema MC, 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;25:640–647.

3. Kohn LT, Corrigan JM, Donaldson M. To Err Is Human: Building a Safer Health System. Washington, DC: Institute of Medicine; 1999.

4. National Patient Safety Foundation. *Free From Harm: Accelerating Patient Safety Improvement Fifteen Years After To Err Is Human*. Boston, MA: National Patient Safety Foundation; 2015.

5. Baker GR, Norton PG, Flintoft V, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ*. 2004;170:1678–1686.

6. Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ*. 2001;322:517–519.

7. Davis P, Lay-Yee R, Briant R, et al. Adverse events in New Zealand public hospitals I: occurrence and impact. *N Z Med J*. 2002;115:U271.

8. Wilson RM, Runciman WB, Gibberd RW, et al. The quality in Australian health care study. *Med J Aust*. 1995;163:458–471.

9. Thomas EJ, Studdert DM, Burstin HR, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care*. 2000;38:261–271.

10. Levinson DR. Adverse events in hospitals: national incidence among Medicare benificiaries. Department of Health and Human Services–Office of Inspector General. 2010. Available at: https://psnet.ahrq.gov/issue/adverse-events-hospitals-national-incidence-among-medicare-beneficiaries. Accessed August 1, 2019.

11. Forster AJ, Andrade J, Van Walraven C. Validation of a discharge summary term search method to detect adverse events. *J Am Med Inform Assoc*. 2005;12:200–206.

12. Forster AJ, Worthington JR, Hawken S, et al. Using prospective clinical surveillance to identify adverse events in hospital. *BMJ Qual Saf*. 2011;20: 756–763.

13. Murff HJ, Patel VL, Hripcsak G, et al. Detecting adverse events for patient safety research: a review of current methodologies. *J Biomed Inform*. 2003; 36:131–143.

14. Michel P, Quenon JL, de Sarasqueta AM, et al. Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. *BMJ*. 2004;328:199–190.

15. Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *N Engl J Med*. 1991;324:370–376.

16. Brennan TA, Leape LL. Adverse events, negligence in hospitalized patients: results from the Harvard Medical Practice Study. *Perspect Healthc Risk Manage*. 1991;11:2–8.

17. Griffin FA. *IHI Global Trigger Tool for Measuring Adverse Events (Second Edition). IHI Innovation Series White Paper*. Cambridge, MA: Institute for Healthcare Improvement; 2009.

18. Hanskamp-Sebregts M, Zegers M, Vincent C, et al. Measurement of patient safety: a systematic review of the reliability and validity of adverse event detection with record review. *BMJ Open*. 2016;6:e011078.

19. Hibbert PD, Molloy CJ, Hooper TD, et al. The application of the Global Trigger Tool: a systematic review. *Int J Qual Health Care*. 2016;28: 640–649.

20. Madden C, Lydon S, Curran C, et al. Potential value of patient record review to assess and improve patient safety in general practice: a systematic review. *Eur J Gen Pract*. 2018;24:192–201.

21. *Patient Safety: Rapid Assessment Methods for Estimating Hazards*. Geneva: World Health Organization, Department of Human Resources for Health; 2003.

22. Institue for Healthcare Improvement. Across the chasm: six aims for changing the health care system. Available at: http://www.ihi.org/resources/Pages/ImprovementStories/AcrosstheChasmSixAimsforChangingtheHealthCareSystem.aspx. Accessed December 5, 2019.

23. Gilbert EH, Lowenstein SR, Koziol-McLain J, et al. Chart reviews in emergency medicine research: where are the methods? *Ann Emerg Med*. 1996;27:305–308.

24. Badcock D, Kelly AM, Kerr D, et al. The quality of medical record review studies in the international emergency medicine literature. *Ann Emerg Med*. 2005;45:444–447.

25. Worster A, Bledsoe RD, Cleve P, et al. Reassessing the methods of medical record review studies in emergency medicine research. *Ann Emerg Med*. 2005;45:448–451.

26. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol*. 2009;62:e1–e34.

27. Ock M, Lee SI, Jo MW, et al. Assessing reliability of medical record reviews for the detection of hospital adverse events. *J Prev Med Public Health*. 2015;48:239–248.

28. David G, Gunnarsson CL, Waters HC, et al. Economic measurement of medical errors using a hospital claims database. *Value Health*. 2013;16: 305–310.

29. Kennerly DA, Kudyakov R, da Graca B, et al. Characterization of adverse events detected in a large health care delivery system using an enhanced global trigger tool over a five-year interval. *Health Serv Res*. 2014;49: 1407–1425.

30. Kurutkan MN, Usta E, Orhan F, et al. Application of the IHI Global Trigger Tool in measuring the adverse event rate in a Turkish healthcare setting. *Int J Risk Saf Med*. 2015;27:11–21.

31. Farup PG. Are measurements of patient safety culture and adverse events valid and reliable? Results from a cross sectional study. *BMC Health Serv Res*. 2015;15:186.

32. Naessens JM, Campbell CR, Huddleston JM, et al. A comparison of hospital adverse events identified by three widely used detection methods. *Int J Qual Health Care*. 2009;21:301–307.

33. Mull HJ, Brennan CW, Folkes T, et al. Identifying previously undetected harm: piloting the Institute for Healthcare Improvement's Global Trigger Tool in the Veterans Health Administration. *Qual Manag Health Care*. 2015;24:140–146.

34. Rutberg H, Borgstedt Risberg M, Sjodahl R, et al. Characterisations of adverse events detected in a university hospital: a 4-year study using the Global Trigger Tool method. *BMJ open*. 2014;4:e004879.

35. Weissman JS, Schneider EC, Weingart SN, et al. Comparing patient-reported hospital adverse events with medical record review: do

patients know something that hospitals do not? *Ann Intern Med*. 2008;149: 100–108.

36. de Feijter JM, de Grave WS, Muijtjens AM, et al. A comprehensive overview of medical error in hospitals using incident-reporting systems, patient complaints and chart review of inpatient deaths. *PLoS One*. 2012; 7:e31125.

37. Classen DC, Resar R, Griffin F, et al. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff (Millwood)*. 2011;30:581–589.

38. Christiaans-Dingelhoff I, Smits M, Zwaan L, et al. To what extent are adverse events found in patient records reported by patients and healthcare professionals via complaints, claims and incident reports? *BMC Health Serv Res*. 2011;11:49.

39. Blais R, Bruno D, Bartlett G, et al. Can we use incident reports to detect hospital adverse events? *J Patient Saf*. 2008;4:9–12.

40. Sari AB, Sheldon TA, Cracknell A, et al. Sensitivity of routine system for reporting patient safety incidents in an NHS hospital: retrospective patient case note review. *BMJ*. 2007;334:79.

41. Macharia WM, Muteshi CM, Wanyonyi SZ, et al. Comparison of the prevalence and characteristics of inpatient adverse events using medical records review and incident reporting. *S Afr Med J*. 2016;106:1021–1036.

42. Bjertnaes O, Deilkas ET, Skudal KE, et al. The association between patient-reported incidents in hospitals and estimated rates of patient harm. *Int J Qual Health Care*. 2015;27:26–30.

43. Sharek PJ, Parry G, Goldmann D, et al. Performance characteristics of a methodology to quantify adverse events over time in hospitalized patients. *Health Serv Res*. 2011;46:654–678.

44. Hwang JI, Chin HJ, Chang YS. Characteristics associated with the occurrence of adverse events: a retrospective medical record review using the Global Trigger Tool in a fully digitalized tertiary teaching hospital in Korea. *J Eval Clin Pract*. 2014;20:27–35.

45. Schildmeijer K, Nilsson L, Arestedt K, et al. Assessment of adverse events in medical care: lack of consistency between experienced teams using the global trigger tool. *BMJ Qual Saf*. 2012;21:307–314.

46. Kennerly DA, Saldana M, Kudyakov R, et al. Description and evaluation of adaptations to the global trigger tool to enhance value to adverse event reduction efforts. *J Patient Saf*. 2013;9:87–95.

47. Unbeck M, Schildmeijer K, Henriksson P, et al. Is detection of adverse events affected by record review methodology? An evaluation of the " Harvard Medical Practice Study" method and the "Global Trigger Tool". *Patient Saf Surg*. 2013;7:10.

48. Naessens JM, O'Byrne TJ, Johnson MG, et al. Measuring hospital adverse events: assessing inter-rater reliability and trigger performance of the Global Trigger Tool. *Int J Qual Health Care*. 2010;22: 266–274.

49. Classen. Development and evaluation of the Institute for Healthcare Improvement Global Trigger Tool. *J Patient Saf*. 2008;4:169–177.

50. Akbari Sari A, Doshmangir L, Torabi F, et al. The incidence, nature and consequences of adverse events in Iranian hospitals. *Arch Iran Med*. 2015; 18:811–815.

51. Thomas EJ, Lipsitz SR, Studdert DM, et al. The reliability of medical record review for estimating adverse event rates. *Ann Intern Med*. 2002; 136:812–816.

52. Soop M, Fryksmark U, Koster M, et al. The incidence of adverse events in Swedish hospitals: a retrospective medical record review study. *Int J Qual Health Care*. 2009;21:285–291.

53. Zegers M, de Bruijne MC, Wagner C, et al. The inter-rater agreement of retrospective assessments of adverse events does not improve with two reviewers per patient record. *J Clin Epidemiol*. 2010;63:94–102.

54. Zegers M, de Bruijne MC, Wagner C, et al. Adverse events and potentially preventable deaths in Dutch hospitals: results of a retrospective patient record review study. *Qual Saf Health Care*. 2009;18:297–302.

55. Sousa P, Uva AS, Serranheira F, et al. Estimating the incidence of adverse events in Portuguese hospitals: a contribution to improving quality and patient safety. *BMC Health Serv Res*. 2014;14:311.

56. Klein DO, Rennenberg R, Koopmans RP, et al. The ability of triggers to retrospectively predict potentially preventable adverse events in a sample of deceased patients. *Prev Med Rep*. 2017;8:250–255.

57. Asavaroengchai S, Sriratanaban J, Hiransuthikul N, et al. Identifying adverse events in hospitalized patients using global trigger tool in Thailand. *Asian Biomed*. 2009;3:545–550.

58. O'Leary KJ, Devisetty VK, Patel AR, et al. Comparison of traditional trigger tool to data warehouse based screening for identifying hospital adverse events. *BMJ Qual Saf*. 2013;22:130–138.

59. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.

60. Adler L, Yi D, Li M, et al. Impact of inpatient harms on hospital finances and patient clinical outcomes. *J Patient Saf*. 2018;14:67–73.

61. Zwaan L, De Bruijne M, Wagner C, et al. Patient record review of the incidence, consequences, and causes of diagnostic adverse events. *Arch Intern Med*. 2010;170:1015–1021.

62. Kobayashi M, Ikeda S, Kitazawa N, et al. Validity of retrospective review of medical records as a means of identifying adverse events: comparison between medical records and accident reports. *J Eval Clin Pract*. 2008;14:126–130.

63. Deilkas ET, Bukholm G, Lindstrom JC, et al. Monitoring adverse events in Norwegian hospitals from 2010 to 2013. *BMJ Open*. 2015;5:e008576.

64. Deilkas ET, Risberg MB, Haugen M, et al. Exploring similarities and differences in hospital adverse event rates between Norway and Sweden using Global Trigger Tool. *BMJ Open*. 2017;7:e012492.

65. Pappas SH. The cost of nurse-sensitive adverse events. *J Nurs Adm*. 2008; 38:230–236.

66. Wardle G, Wodchis WP, Laporte A, et al. The sensitivity of adverse event cost estimates to diagnostic coding error. *Health Serv Res*. 2012;47(3 Pt 1): 984–1007.

67. Goodman JC, Villarreal P, Jones B. The social cost of adverse medical events, and what we can do about it. *Health Aff (Millwood)*. 2011;30:590–595.

68. Brown P, McArthur C, Newby L, et al. Cost of medical injury in New Zealand: a retrospective cohort study. *J Health Serv Res Policy*. 2002; 7(Suppl 1):S29–S34.

69. Hoogervorst-Schilp J, Langelaan M, Spreeuwenberg P, et al. Excess length of stay and economic consequences of adverse events in Dutch hospital patients. *BMC Health Serv Res*. 2015;15:531.

70. Rafter N, Hickey A, Conroy RM, et al. The Irish National Adverse Events Study (INAES): the frequency and nature of adverse events in Irish hospitals-a retrospective record review study. *BMJ Qual Saf*. 2017;26:111–119.

71. Baines R, Langelaan M, de Bruijne M, et al. How effective are patient safety initiatives? A retrospective patient record review study of changes to patient safety over time. *BMJ Qual Saf*. 2015;24:561–571.

72. Baines RJ, Langelaan M, de Bruijne MC, et al. Changes in adverse event rates in hospitals over time: a longitudinal retrospective patient record review study. *BMJ Qual Saf*. 2013;22:290–298.

73. Federal Ministry of Health, Germany. Best practices in patient safety. 2017. Available at: https://www.bundesgesundheitsministerium.de/fileadmin/ Dateien/3_Downloads/P/Patientensicherheit/Patient_Safety_Summit_ 2017_Final_Report.pdf. Accessed July 1, 2019.

74. Suarez C, Menendez MD, Alonso J, et al. Detection of adverse events in an acute geriatric hospital over a 6-year period using the global trigger tool. *J Am Geriatr Soc*. 2014;62:896–900.

75. Mortaro A, Moretti F, Pascu D, 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 [published online ahead of print June 9, 2017]. *J Patient Saf*.

76. Langelaan M, Broekens MA, Bruijne MC de, et al. Monitor zorggerelateerde schade 2015/2016: dossieronderzoek bij overleden patiënten in Nederlandse ziekenhuizen. www.nivel.nl: NIVEL; 2017. Available at: https://www.nivel.nl/nl/publicatie/monitor-zorggerelateerde-schade-20152016-dossieronderzoek-bij-overleden-patienten. Accessed January 4, 2019.

77. http://www.duque.eu/. Accessed October 8, 2018.

78. Bates DW, O'Neil AC, Petersen LA, et al. Evaluation of Screening Criteria for Adverse Events in Medical Patients. *Med Care*. 1995;33: 452–462.

79. Øvretveit J. *Does Improving Quality Save Money? A Review of Evidence of Which Improbement to Quality Reduce Costs to Health Service Providers.* London: The Health Foundation; 2009.

80. Michel P. *Strengths and Weaknesses of Available Methods for Assessing the Nature and Scale of Harm Caused by the Health System: Literature Review.* Geneva, Switzerland: World Health Organization; 2003.

81. Zahabi M, Kaber DB, Swangnetr M. Usability and safety in electronic medical records Interface design: a review of recent literature and guideline formulation. *Hum Factors*. 2015;57:805–834.

82. Tanner C, Gans D, White J, et al. Electronic health records and patient safety: co-occurrence of early EHR implementation with patient safety practices in primary care settings. *Appl Clin Inform*. 2015;6:136–147.

83. Sittig DF, Singh H. Electronic health records and national patient-safety goals. *N Engl J Med*. 2012;367:1854–1860.

84. Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. *J Am Med Inform Assoc*. 2005;12:448–457.

85. Penz JF, Wilcox AB, Hurdle JF. Automated identification of adverse events related to central venous catheters. *J Biomed Inform*. 2007;40: 174–182.

86. Rochefort CM, Buckeridge DL, Tanguay A, et al. Accuracy and generalizability of using automated methods for identifying adverse events from electronic health record data: a validation study protocol. *BMC Health Serv Res*. 2017;17:147.

87. Bates DW, Evans RS, Murff H, et al. Detecting adverse events using information technology. *J Am Med Inform Assoc*. 2003;10:115–128.

88. Gerdes LU, Hardahl C. Text mining electronic health records to identify hospital adverse events. *Stud Health Technol Inform*. 2013;192:1145.