BMJ open

Strengths and weaknesses of working with the Global Trigger Tool method for retrospective record review: focus group interviews with team members

Kristina Schildmeijer,[1] Lena Nilsson,[2] Joep Perk,[1] Kristofer Årestedt,[1,3,4] Gunilla Nilsson[1]

To cite: Schildmeijer K, Nilsson L, Perk J, *et al*. Strengths and weaknesses of working with the Global Trigger Tool method for retrospective record review: focus group interviews with team members. *BMJ Open* 2013;**3**:e003131. doi:10.1136/bmjopen-2013-003131

▸ Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2013-003131).

Received 26 April 2013
Revised 13 August 2013
Accepted 14 August 2013

[1]Faculty of Health and Life Sciences, School of Health and Caring Sciences, Linnaeus University, Kalmar, Sweden
[2]Department of Medicine and Health Sciences, Department of Anesthesia and Intensive Care, Linköping University, County Council of Östergötland, Linköping, Sweden
[3]Division of Nursing Science, Department of Medical and Health Sciences, Linköping University, Linköping, Sweden
[4]Palliative Research Centre, Ersta Sköndal University College and Ersta hospital, Stockholm, Sweden

Correspondence to
Dr Kristina Schildmeijer; kristina.schildmeijer@lnu.se

ABSTRACT

Objectives: The aim was to describe the strengths and weaknesses, from team member perspectives, of working with the Global Trigger Tool (GTT) method of retrospective record review to identify adverse events causing patient harm.

Design: A qualitative, descriptive approach with focus group interviews using content analysis.

Setting: 5 Swedish hospitals in 2011.

Participants: 5 GTT teams, with 5 physicians and 11 registered nurses.

Intervention: 5 focus group interviews were carried out with the five teams. Interviews were taped and transcribed verbatim.

Results: 8 categories emerged relating to the strengths and weaknesses of the GTT method. The categories found were: *Usefulness of the GTT, Application of the GTT, Triggers, Preventability of harm, Team composition, Team tasks, Team members' knowledge development and Documentation*. Gradually, changes in the methodology were made by the teams, for example, the teams reported how the registered nurses divided up the charts into two sets, each being read respectively. The teams described the method as important and well functioning. Not only the most important, but also the most difficult, was the task of bringing the results back to the clinic. The teams found it easier to discuss findings at their own clinics.

Conclusions: The GTT method functions well for identifying adverse events and is strengthened by its adaptability to different specialties. However, small, gradual methodological changes together with continuingly developed expertise and adaption to looking at harm from a patient's perspective may contribute to large differences in assessment over time.

ARTICLE SUMMARY

Strengths and limitations of this study

- Experienced reviewers from different sized hospitals were interviewed. All team members took the opportunity to speak as a team.
- The analysis was based on propositions from the five focus group teams. However, it was not clear whether their experiences of working with the GTT method reflected the opinions of teams from other hospitals or under other circumstances. This needs to be verified through further studies.
- The study was carried out in Sweden where preventability has been added to the GTT method. Judgment about this was considered as subjective by the teams. However, apart from the addition of preventability, the Swedish GTT method does not differ from the original GTT method.

BACKGROUND

Adverse events (AEs) are common in healthcare. European studies report a prevalence of patient harm of 9–12%.[1–4] AEs causing patient harm can be identified through retrospective patient record reviews.[3 5] The Global Trigger Tool (GTT), developed by the Institute for Healthcare Improvement (IHI),[6] is one such method, primarily intended for following the level of AEs on a hospital level, that has been increasingly used.[7 8] It identifies AEs causing patient harm caused by medical treatment, not the underlying medical condition of the patient. In GTT 20 medical records from randomly chosen hospital admissions are retrospectively reviewed every month by experienced teams often consisting of two registered nurses (RNs) and one physician. An advantage of the GTT is the measurement of the rate of AEs over time within an organisation, and the use of GTT for evaluating and measuring patient safety has been promoted in several countries.[9–13] GTT has been shown superior to other incident-reporting systems in identifying AEs.[7 14] On the other hand, the methodology of record reviews, including GTT, has been criticised for not being

sufficiently robust, whereby judgements have been found to differ between reviewing teams.[15–18] Moderate inter-rater agreement has been shown,[17–19] and in the study according to oncological care a Bland-Altman analysis of the GTT method showed large random errors when comparing review teams.[17] This reduces its ability to track a true change in the level of patient harm.

In order to gain a nuanced view of record review as a tool for evaluating patient safety work, we find it important to gain a deeper understanding of the experiences from reviewing teams working with the GTT method. This knowledge could provide us with a better understanding of the strengths and weaknesses of the GTT method, which might help us to better understand causes of disagreement between teams, and provide valuable knowledge for a broader view of the GTT method and its role in patient safety work.

Objective

The aim was to describe strengths and weaknesses, from the perspective of team members working with the GTT method.

METHOD

The study was conducted using a qualitative, descriptive approach with focus group interviews as described and analysed by Krueger and Casey.[20]

The Global Trigger Tool method

The GTT method is based on reviewing medical records from randomly selected admissions in a structured way in an effort to identify AEs by searching for 'triggers'. It is designed for utilising small samples over time from a hospital or a healthcare organisation.[21 22] A 'trigger' may indicate that an AE has occurred, for example, the trigger 'patient fall', indicating for instance a side effect of drugs that can cause patient harm from a fall. Teams consisting of one physician and two RNs review a patient's medical records. In the first stage, the RNs review independently noting 'triggers' and associated AEs on a chart. The time limit is set at 20 min. The nurses then discuss their findings, and after having reached a consensus, complete a new chart together. Charts with a potential AE are forwarded to the physician. In stage two, the physician determines if an AE has occurred and, if so, the level of harm. The Swedish GTT version is modified for Swedish conditions, listing 53 triggers rather than 54 as in the original GTT method.[6] It also contains an additional item referred to as 'preventability'. Preventability is graded on a scale from 1 to 6, with 1 being no real evidence for preventability and 6 being completely secure evidence for preventability. Preventability for each AE is judged by the physicians[3 9] (box 1).

Box 1 Description of preventability scale

Preventability scale

1. No real evidence for preventability
2. Weak to small evidence for preventability
3. Preventability less likely than 50%, but close
4. Preventability more probable than 50%, but close
5. Strong evidence for preventability
6. Completely secure evidence for preventability

Classes 1–3 are considered non-preventable harm.
Classes 4–6 are considered preventable harm.[3]

Participants

Five focus groups participated in this study. Each included one GTT review team from five different sized hospitals in south-eastern Sweden. The teams had participated in an earlier study comparing reviews of the same set of records.[18] Team members were contacted by email requesting participation in focus group interviews concerning their experiences of using the GTT. All team members agreed to participate in the interviews. The 16 participants consisted of 5 physicians and 11 RNs. All had long experience of working in healthcare, and 3–5 years' experience of reviewing medical record (table 1). Participants from the teams were part of their hospital's patient safety teams, and worked extensively with patient safety issues including the GTT on a hospital-wide level. Eleven of the 16 team members were educated in GTT reviewing by attending a 1-day programme and partaking yearly in a 2-day regional meeting devoted to training and collaboration. The other five team members had been trained by their respective colleagues. Of the teams, four (I, II, III and V) had participated in a regional GTT network for 4 years where issues including harm and preventability were discussed. Four of the hospitals are middle-sized with about 200 beds, and one is a university hospital with approximately 600 beds. All hospitals used electronic medical records.

Data collection

The interviews took place between January and March 2011 at the five hospitals where the participants worked. Five 80–95 min focus group interviews were conducted. One of the authors (KS) moderated all interviews, with an assistant moderator (GN or KÅ). The moderator was responsible for facilitating the discussion and prompting team members to speak, while the assistant moderator recorded sessions and took notes. KS had worked with the GTT method for 4 years on a hospital level, while GN and KÅ had experience of working with qualitative studies and focus groups. Besides taking notes they were to also ensure that the moderator's preunderstanding did not affect team members' responses. After an opening question, aimed to make team members feel comfortable, an introductory question was asked: What are your experiences of the strengths and weaknesses of the GTT method? The introductory question was followed by transition questions: "Tell us how it was when you started reviewing by using the GTT" and "How are you presently reviewing?" The purpose of transition

Table 1 Team member characteristics

Team	Profession	Age/ gender	Experiences of specialty (years)	Profession	Working with GTT (years)	Formal education in GTT
I	Physician	64M	30	Psychiatry	3	No
	Nurse	60F	39	Midwife	4	Yes
	Nurse	59F	37	Internal medicine care	3	No
	Nurse	63F	30	Psychiatric care	4	Yes
II	Physician	54F	25	Anaesthesiology	5	Yes
	Nurse	39F	19	Intensive care	3	No
	Nurse	38F	18	Intensive care	3	No
III	Physician	64M	37	ENT	4	Yes
	Nurse	44F	18	Emergency care	4	Yes
	Nurse	63F	39	Orthopaedic care	4	Yes
IV	Physician	66M	40	Internal medicine care	4	No
	Nurse	53F	27	Internal medicine care	4	Yes
	Nurse	58F	36	Intensive care	4	Yes
V	Physician	63M	35	Surgeon	4	Yes
	Nurse	37F	13	Paediatric care	4	Yes
	Nurse	58F	34	Midwife	4	Yes

ENT, ear, nose and throat; F, female; GTT, Global Trigger Tool; M, male.

questions was to form a link between the introductory and key questions. The key questions captured the major areas of concern. Examples of key questions were: "From your experiences, which strengths/weaknesses do you find with the method in its entirety?" "What are your experiences of the different triggers?" and "What is your opinion on the judgement of preventability?" During the interviews, the moderator asked probing questions, for example, "What do you mean?" or "Can you explain a little further?" Towards the ends of the interviews the moderator asked questions, for example, if there was more to discuss, or if something needed further clarification. The moderator also summarised the key points. All interviews were recorded and transcribed verbatim. Immediately after the focus group interviews, the moderator and the assistant moderator gave a debriefing of their first impressions and compared these interpretations from those found in earlier focus groups. The transcribed interviews were returned to the team members with the question of whether they felt that the text reflected their interviews. All accepted the text.

Analysis of the interviews

Interviews were analysed according to Krueger and Casey to identify patterns and discover relationships between ideas. Data analysis proceeded simultaneously with data collection until no new information emerged. The text was coded and opinions with similar meanings were grouped together until eight categories emerged. Comparisons were made throughout the analysis between categories and the text as a whole.[20]

Trustworthiness is needed in all research strategy.[23] To validate the findings investigator triangulation was used. Transcripts were read and reread by all researchers to gain a sense of content, sometimes returning to interview recordings to become completely familiar with the data and comprehend its essential features. With the aim of the study in mind researchers read the text and made notes and headings in the margins to include all aspects of the content. The first author (KS) established categories based on citations. To ensure confirmability, coauthors discussed the categories, and changes were made until consensus was reached. To increase credibility according to Krueger and Casey,[20] at least three focus group interviews should be carried out. In this study, five focus groups were conducted. A careful description of the sampling procedure and data analyses was presented to ensure dependability. Each citation was given a number for data reporting purposes to show evidence of reporting across responses from the five teams (I–V).

Ethical considerations

All members of the GTT teams gave their individual informed consent to participate in the study.

RESULTS

The categories identified were: *Usefulness of the GTT, Application of the GTT, Triggers, Preventability of harm, Team design', Team tasks, Team members' knowledge development and Documentation.* Each category is presented by its strengths and weaknesses.

Usefulness of the GTT

Strengths

Team members found the method useful in identifying patient harm. An advantage was that the method, apart from being used in a random selection of records, could also be used by specific specialties, or for a subgroup of patients, for example, for deceased patients.

> It is a useful tool; it is relevant and it identifies harm. I feel that it is important. It feels that the tool can positively affect healthcare… (II)

Weaknesses

Even if all teams considered the method useful, they also mentioned that it was oriented mainly towards harm connected to actions undertaken by physicians. A nursing care perspective was missing and was requested.

> One wishes it (the GTT) was more care oriented...lack of care can also harm patients (IV)

All teams talked about the method's weakness in not capturing all failures, as omission is not part of the GTT.

> There is no trigger for omission. When reading notes, one sometimes wonders why no one has reacted (V)

Application of the GTT

Strengths

Most often RNs could easily make their assessments within 20 min, pointing out the time limit as a strong point of the method.

> You quickly find what you need when you have figured out how to review (I)

While some RNs felt it better to review medical records outside their specialties, others found it easier to examine records from their own specialty.

> It is easier to understand why certain things were done when you review your own area (IV)

Weaknesses

The teams gave careful reports on how they previously performed audits, and all mentioned that they used the method as described in the handbook. Still, all teams made changes in the review process. For example, in the manual it states that the two RNs in the team should review the same records separately and then reach a consensus. Instead, all RNs reported dividing the charts into two sets and reading them separately.

> It is good to sit together and talk, but we no longer review the same records (V)

The RNs sometimes thought it hard to restrict themselves to looking only for triggers and associated possible harm as indicated in the chart, when they found other striking things in patient records. This led them to use the method for purposes other than those originally intended. In such cases they chose to mark these findings on the GTT chart and discuss them with the team.

> When I find things that make me react I include them. They may be important in other contexts (I)

Even if the time limit was not considered a major problem, there were still situations when the time limit was a problem, for example, where patients had experienced a long period of care. RNs mentioned that in some cases the time limit was exceeded.

> We read the records very carefully; and 20 minutes was in no way enough! (IV)

When reviewing records from their own area there was a risk for the RNs becoming insensible, regarding patient harm as something that just happens.

> You are more forgiving in your own area, you become blind. Therefore, it is good that reviewers are from different clinics (III)

Triggers

Strengths

The teams found the triggers' intuitive in a sense that they were easy to keep in mind covering wide areas and facilitating reviewing.

> The triggers are good and useful (I)

Weaknesses

Even if all teams were satisfied with the GTT method, they found that some triggers were imprecise, and some never used.

> The trigger "treatment" is vague and can always or never be used (II)

The team members' affiliations influenced to some extent their statements about a need for additional triggers. All teams mentioned that a trigger for 'failure or measures not been performed as intended' was missing. The teams also said that several nursing care triggers were missing and should be added.

> Triggers evaluating harm in the areas of nutrition, elimination, pain and oral hygiene are all missing (IV)

One team mentioned that bad behaviour by healthcare providers was not brought up.

> The method is not designed for assessing personnel behaviour from a patient perspective (I)

Preventability of harm

Strengths

The teams mentioned that development of healthcare processes and the work of patient safety required new thinking and that, for Swedish conditions added 'judgement of preventability', helped them in achieving this view. The question "Could this have been done differently?" in the sentence "Could we have prevented this from happening?" provided an opportunity to consider preventability from more than one perspective.

To gain improvement you must consider many findings as avoidable. It may be avoidable if things are done differently (III)

Weaknesses

When 'judgement of preventability' came up for discussion, all teams considered the concept as too subjective. They felt there would be as many answers as the number of people asked.

Determining harm to the patient is easy but the judgment of preventability is difficult (I)

Team composition

Strengths

The teams found strength in their interdisciplinary representation of healthcare specialties. Several personnel categories to consider patient care from different perspectives were considered a prerequisite for the implementation of healthcare improvements.

To gain good results, everyone's efforts are necessary and important; we see things differently (II)

Weaknesses

A sufficient number of reviewers were considered important for avoiding problems in the event of reviewers dropping out. Too few reviewers led to team vulnerability due to the risk of team members being unavailable.

At least four reviewers are required (I)

It was also important for team members to continue reviewing for long periods of time.

You probably need, in any case, at least some years' of experience of reviewing (IV)

However, the team also saw a danger in remaining too long as a reviewer, due to possible increasing tolerance toward AEs or substandard care.

What I'm wondering about is that there is a risk of becoming less careful, to put things aside thinking that this is nothing (II)

Tasks

Strengths

The interdisciplinary composition of the teams made it easier to take up reviewed events for discussion, not only within the team but also with others. Physicians discussed with other physicians, while RNs discussed with colleagues from their own departments.

After reviewing I discussed findings with my colleagues (IV)

Weaknesses

Teams sought intensified discussions to increase the work of quality improvement of patient safety. They perceived their most important task was to convey results back to their clinics. This was difficult when randomly reviewing records from an entire hospital.

Our ambition is to go to the different clinics and tell them directly of our concerns for their specific problems (I)

Even if the teams would have liked to provide feedback to the clinics, they were concerned about how their findings would be received.

We do not want them to look at us as police officers, but wish to review and return to the clinics with our findings with the aim of improving things (I)

Team member's knowledge development

Strengths

All teams had used the GTT on a monthly basis from 3 to 5 years for their hospital's patient safety work and were accustomed in reviewing patient records from the perspective of various medical specialties. Team members believed they had developed their skills gradually. They had also gained a greater understanding of the healthcare system during their time on the GTT team. They had increased their skills of how to and how not to document through the reading of patient records.

You learn a lot by reading others' notes on how to better carry out documentation (IV)

They also mentioned they could better observe care and harm from the patient's perspective.

It (GTT) requires time to be able to say that from the patient's perspective harm has occurred (II)

Documentation

Strengths

The teams considered RNs more accurate and precise in their documentation than physicians; and it was the RNs' notes that revealed the greatest number of AEs and helped the most to determine levels of harm. Nursing care documentation was richer in its comments about the patients' conditions during hospital stay.

RNs' notes are much more detailed, concerning, for example, urinary infections and how the patient really feels (V)

Weaknesses

Although the teams felt they had identified patient harm through the RNs' documentation, they considered documentation generally poor. They described sparse and duplicated documentation. The teams mentioned

that minor incidents during hospital stay, causing minor patient harm, were not mentioned at all in the discharge letter.

> There is a considerable amount of duplicated documentation (I)

> Medical records summaries are often written after the care episode and are very sparse, small incidents aren't mentioned at all (V)

The teams felt they had to look through a large body of text that sometimes made it difficult to orient themselves in the patients' medical records. They also felt that some of the notes had no real impact on patient care, which influenced their overview of the patients' medical records. Sometimes it was hard for team members to understand how patient care had actually been conducted, making review more difficult.

> There is so much text that it becomes unmanageable, you get no clear picture of the patient's condition (III)

> One sometimes wonders how care has been carried out (I)

DISCUSSION

To the best of our knowledge, this study is the first investigation of how the GTT method is experienced and implemented from the reviewing team's perspective. The teams found the GTT method useful whereby it identified patient harm and could be used for different specialties. They had gradually modified the original review method to suit a personal context, for example, time dedicated for reviewing. The method was found subjective in its estimation of the, for Swedish conditions, added judgement of preventability. A nursing care perspective was missed and should be added, and insufficient documentation was a barrier to medical patient record reviewing.

The teams made small and gradual changes in the methodology, which may have contributed to large differences in the assessments over time. They mentioned having used the method as described in the Swedish manual,[9] based on the original IHI method.[6] However, interviews revealed how teams had deviate from the original method by using it somewhat differently, occasionally even for other purposes. For example, the RNs did not review the same charts and reach a consensus about their findings as intended in the manual. Instead, they divided the charts into two parts and reviewed half of them each. Some teams, besides random admissions, also reviewed the records of all hospital deaths. Another example was how one team sometimes interpreted the presence of triggers as equal to the occurrence of harm. From their point of view the triggers indicated substandard care. If treatment had been given according to standard, the patient would not have needed, for example, to be readmitted within 30 days. All teams missed a trigger for 'measures not performed according to standard care', indicating that the teams could have mixed quality measures and harm to patients. Another possibility is that the teams have had the Swedish National Board of Health and Welfare's definition of AEs in their mind instead of the definition of harm from the GTT manual, restricting harm to physical injuries. By the Swedish National Board of Health and Welfare harm is defined as "Any suffering, discomfort, bodily or mental injury, illness or death caused by healthcare and which is not an inevitable consequence of the patient's condition or an expected effect of the treatment received by the patient because of her/his condition."[24]

The teams had also made another deviation from the original method; the time limit of 20 min was sometimes exceeded, as the records were read carefully when the RNs did not always restrict themselves to look for triggers alone. A recent Danish study showed that the method was initially interpreted differently at different hospitals.[25] Small changes in methodology and a gradual development of skills could explain inter-rater disagreement between teams. Another reason for disagreement between teams could be the teams obtaining of a patient's perspective along with different team composition regarding interdisciplinary representation. However, it is important to remember that several studies have shown the superiority of record reviews in detecting and categorising AEs.[1 7 8 26–29] This was also the team members' view, considering that the most important function of the GTT method was the identification of AEs. In line with Brandrud *et al*,[30] they developed new knowledge that appeared to be useful locally.

Even if it was also considered important to provide feedback from GTT reviews to RNs and physicians on the wards, no team wanted to be regarded as controllers. They would rather look at harm on a systematic level as did Resar *et al*,[31] who mentioned that focus on AEs targets the system rather than the individual, and can lead to the exploration of methodology to improve or enhance clinical outcomes. The advantages of the GTT method were that it could be adapted to specialties, subgroups of patients or healthcare processes. Safety problems can be identified by medical record review leading to specific measures to prevent future harm.

The teams had views concerning triggers that should be added or removed. They mentioned that the method included triggers never used, and those that could be used at all times, without necessarily identifying AEs. The teams missed nursing care triggers, such as patient pain, nutrition and elimination. They also saw a need for clinic-specific triggers. There are activity-specific trigger versions for ambulatory care,[32] intensive care,[31] surgical care,[33] neonatal care[29] and primary care.[26] Creating new triggers requires reflection. We agree with Kaafarani *et al*,[32] on the importance of considering clinical relevance, utility and feasibility of implementation when designing triggers.

The teams found remarkably insufficient documentation and mentioned this as a potential problem for becoming more proactive in patient safety work. As the GTT method is based solely on findings in patient records the teams highlighted the lack of documentation, which casted doubt on how care had been conducted. This is in line with other studies.[34–36] Weingart *et al*,[36] mentioned that many AEs are not recorded in the medical charts, attributable to variable standards of documentation, clinical unawareness or oversight, and concern about liability exposure. In our study team members mentioned finding it difficult to gain an overview of the patient's illness. The same result has been shown by Stevenson and Nilsson.[35] They found that essential information such as vital signs was difficult to enter and locate in electronic records. It was also unclear where specific information such as blood pressure and pulse should be documented.

The team members gained greater understanding of the structure of healthcare, and also looked on AEs more from a patient's perspective. The team's belief that their most important task was to bring back results to the clinics can point at an unclear structure for the team's task. It is of importance to have a clear structure for the patient safety work at the hospital level including clarity for the teams, putting their work into the right perspective, for example, that increased patient safety work is a lengthy process.

Being a GTT team member allows the participant new possibilities for collaboration, and a better chance to be included in a patient safety context with accountability for tracking and making changes in healthcare. This is in accordance with Brandrud *et al*,[30] who found three success factors for continuous quality improvement; continuous and reliable information, involvement by all, and an infrastructure based on improvements in knowledge.

Focus group interviews were chosen as they encourage interaction between participants. The team members in our study had worked together as teams for several years. Since we wanted to achieve a picture of positive and negative experiences of working with the GTT in clinical practice, we tried to create an interview situation as comfortable as possible for the team members. For this reason, we chose to keep the teams in their initial compositions in the focus groups.

One limitation may be that this study was carried out in Sweden where preventability has been added to the GTT method. Judgement about this was considered as subjective by the teams. However, apart from the addition of preventability, the Swedish GTT method does not differ from the original GTT method used in many other countries. The analysis is based on propositions from the teams, but we do not know how well they reflect other teams' opinions from other hospitals or under other circumstances. Other GTT teams will probably have similar experiences and be able to refer to these interpretations, and hopefully develop their work of patient's safety. Further limitations are that our study lacks questions related to how the core measurements of the GTT protocol, for example, number of AEs per 100 admissions were analysed and presented and in addition we did not pick up the aspect with statistical process control in the interviews.

CONCLUSIONS

The GTT method was found useful and important with triggers facilitating reviewing, although documentation was generally poor. The most important and difficult task as a review team was to report AEs to the involved clinics. The teams gradually made small changes in the methodology, which together with gradual expertise and different team composition may contribute to differences between teams in the assessments over time. Despite this, we conclude that the GTT method has the strength to identify AEs and could be used in subgroups of patients or in different specialties.

Acknowledgements The authors would like to thank the teams for their efforts.

Contributors KS and GN were responsible for study design, acquisition of data, analysis, interpretation of data, drafting the article and intellectual content. LN and KÅ were responsible for study design, analysis, interpretation of data, drafting the article and intellectual content. JP contributed to study design and intellectual content.

Funding This work was supported by FORSS—the Research Council of Southeast Sweden and the County Council of Kalmar for KS.

Competing interests None.

Ethics approval Ethical permission was obtained from the Regional Ethical Board of Linköping University, Sweden (study number 2010/399-32).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. Landrigan C, Parry G, Bones C, *et al*. Temporal trends in rates of patient harm resulting from medical care. *N Engl J Med* 2010;363:2124–34.
2. Schioler T, Lipczak H, Pedersen BL, *et al*. Incidence of adverse events in hospitals. A retrospective study of medical records. *Ugeskr Laeger* 2001;163:5370–8.
3. 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;21:285–91.
4. Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001;322:517–19.
5. Wilson R, Runciman W, Gibberd R, *et al*. The quality in Australian health care study. *Med J Aust* 1995;163:458–71.
6. Griffin F, Resar R. *IHI Global Trigger Tool for measuring adverse events*. 2nd edn. Cambridge, Massachusetts: Institute for Healthcare Improvement, 2009.
7. Classen D, 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* 2011;30:581–9.
8. Good V, Saldana M, Gilder R, *et al*. Large-scale deployment of the Global Trigger Tool across a large hospital system: refinements for the characterisation of adverse events to support patient safety learning opportunities. *Qual Saf Health Care* 2011;20:25–30.

9. Institute for Healthcare Improvement. *Structured record review to identify and measure the incidence of injuries in health care by the method described Global Trigger Tool: guide to patient safety work [Swe: Strukturerad journalgranskning för att identifiera och mäta förekomst av skador i vården enligt metoden Global Trigger Tool: handbook för patientsäkerhetsarbete]. Institute for Healthcare Improvement Innovation 2007 (a Swedish translated and adjusted version 2008)*. Stockholm: Kommentus, 2008.
10. The Danish Safer Hospital Programme. *Safety DsfP*. 2nd edn. 2012. http://www.sikkerpatient.dk/in-english/the-danish-safer-hospital-programme.aspx (accessed 18 Jun 2013).
11. Canadian Patient Safety Institute, 2012. http://www.patientsafetyinstitute.ca/English/Pages/default.aspx (accessed 18 Jun 2013).
12. NMOH. *In safe hands. Norwegian Knowledge Centre for the Health Services*, 2012. http://www.pasientsikkerhetskampanjen.no/no/I+trygge+hender/In+English/In+Safe+Hands%3A+the+Norwegian+patient+safety+campaign+2011+-+2013.451.cms (accessed 18 Jun 2013).
13. *Patient safety first*. 2nd Edn. Agency NPS, 2012. http://www.patientsafetyfirst.nhs.uk/Content.aspx?path=/About-the-campaign/ (accessed 18 Jun 2013).
14. Naessens JM, Campbell C, Huddleston J, *et al*. A comparison of hospital adverse events identified by three widely used detection methods. *Int J Qual Health Care* 2009;21:301–7.
15. Forster AJ, O'Rourke K, Shojania KG, *et al*. Combining ratings from multiple physician reviewers helped to overcome the uncertainty associated with adverse event classification. *J Clin Epidemiol* 2007;60:892–901.
16. Hofer TP, Bernstein SJ, DeMonner S, *et al*. Discussion between reviewers does not improve reliability of peer review of hospital quality. *Med Care* 2000;38:152–61.
17. Mattsson TO, Knudsen JL, Lauritsen J, *et al*. Assessment of the Global Trigger Tool to measure, monitor and evaluate patient safety in cancer patients: reliability concerns are raised. *Qual Saf Health Care* 2013;7:571–9.
18. 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;21:307–14.
19. 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;46:654–78.
20. Krueger RA, Casey MA. *Focus groups—a practical guide for applied research*. 4th edn. Washington: SAGE, 2009.
21. Classen DC, Lloyd RC, Provost L, *et al*. Development and evaluation of the Institute for Healthcare Improvement Global Trigger Tool. *J Patient Saf* 2008;4:169.
22. Resar R, Rozich J, Classen D. Methodology and rationale for the measurement of harm with trigger tools. *Qual Saf Health Care* 2003;12(Suppl 2):39–45.
23. Patton MQ. *Qualitative research & evaluation methods*. 3rd edn. Washington: SAGE Publications Inc, 2001.
24. National Board of Health and Welfare. *Patient safety and patient safety improvement. An overview [Patientsäkerhet och patientsäkerhetsarbete. En översikt]*. Stockholm: National Board of Health and Welfare, 2004.
25. Von Plessen C, Kodal A, Anhoj J. Experiences with Global Trigger Tool reviews in five Danish hospitals: an implementation study. *BMJ Open* 2012;2:1–8.
26. 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;85:176–80.
27. Nilsson L, Pihl A, Tagsjo M, *et al*. Adverse events are common on the intensive care unit: results from a structured record review. *Acta Anaesthesiol Scand* 2012;56:959–65.
28. Rozich JD, Haraden CR, Resar RK. Adverse drug event trigger tool: a practical methodology for measuring medication related harm. *Qual Saf Health Care* 2003;12:194–200.
29. Sharek P, Horbar J, Mason W, *et al*. Adverse events in the neonatal intensive care unit: development, testing, and findings of an NICU-focused trigger tool to identify harm in North American NICUs. *Pediatrics* 2006;118:1332–40.
30. Brandrud A, Schreiner A, Hjortdahl P, *et al*. Three success factors for continual improvement in healthcare: an analysis of the reports of improvement team members. *BMJ Qual Saf* 2011;20:251–9.
31. Resar R, Rozich J, Simmonds T, *et al*. A trigger tool to identify adverse events in the intensive care unit. *Jt Comm J Qual Patient Saf* 2006;32:585–90.
32. Kaafarani H, Rosen A, Nebeker J, *et al*. Development of trigger tools for surveillance of adverse events in ambulatory surgery. *Qual Saf Health Care* 2010;19:425–9.
33. Griffin FA, Classen DC. Detection of adverse events in surgical patients using the Trigger Tool approach. *Qual Saf Health Care* 2008;17:253–8.
34. Cheevakasemsook A, Chapman Y, Francis K, *et al*. The study of nursing documentation complexities. *Int J Nurs Pract* 2006;12:366–74.
35. Stevenson J, Nilsson G. Nurses' perceptions of an electronic patient record from a patient safety perspective: a qualitative study. *J Adv Nurs* 2012;68:667–76.
36. Weingart S, Pagovich O, Sands D, *et al*. What can hospitalized patients tell us about adverse events? Learning from patient-reported incidents. *J Gen Intern Med* 2005;20:830–6.