The Global Trigger Tool: A Review of the Evidence

August 2013

Report for the Health Quality & Safety Commission New Zealand

Suggested citation:

Health Quality & Safety Commission. 2013. *Global Trigger Tools: A Review of the Evidence*. Wellington: Health Quality & Safety Commission.

ISBN: 978-0-478-38554-0 (online)

Published in 2013 by the Health Quality & Safety Commission.

The document is available online on the Health Quality & Safety Commission's website:

www.hqsc.govt.nz

Acknowledgements

This review was commissioned by Ms Carmela Petagna and staff at the Health Quality & Safety Commission New Zealand and was undertaken by Dr P Hider, Health Evidence Ltd.

Expert peer review and assistance with the development and completion of the work was provided by Ms Gillian Robb, GTT Clinical Lead, Health Quality & Safety Commission.

Disclaimer

Table of contents

Figures

Tables

List of abbreviations

ADE	adverse drug event
ADR	adverse drug reaction
AE	adverse event
AHRQ	Agency for Healthcare Research and Quality (US)
CI	confidence interval
CPOE	computerised physician order entry
DVT	deep vein thrombosis
ED	emergency department
ENT	ear, nose and throat
GTT	global trigger tool
ICD	International Classification of Diseases
ICU	intensive care unit
IHI	Institute for Healthcare Improvement
INR	international normalised ratio
MeSH	Medical Subject Headings
NCC MERP	National Coordinating Council for Medication Error Reporting and Prevention
NHS	National Health Service (UK)
NICU	neonatal intensive care unit
ns	not stated
NZ	New Zealand
PE	pulmonary embolism
PICU	paediatric intensive care unit
POA	present on admission
PPV	positive predictive value
ROT	return to the operating theatre
RR	relative risk
UK	United Kingdom
UROT	unplanned return to the operating theatre
US	United States
WHO	World Health Organization
WHO-ART	World Health Organization Adverse Reactions Terminology

Executive summary

Introduction: The recent focus on patient safety has driven the need for an efficient method to measure adverse events (AEs) at health care organisations. Trigger tools provide a stepped approach to the identification of these events and involve the application of various screening criteria to guide the medical record review process. Trigger tools potentially enable the review process to be more efficient. The Institute for Healthcare Improvement Global Trigger Tool (IHI GTT) was developed in 2000 as a low-resource option for identifying iatrogenic harm that does not require an organisation to operate a sophisticated computerised drug and patient management system. The tool brings additional advantages with its more structured methodology for case sampling, record review and statistical process control results presentation.

Aim: To review the literature associated with the development and use of trigger tools to determine rates of harm in health care settings with particular attention on the IHI GTT.

Methods: A systematic review methodology was employed with structured searches of MEDLINE and EMBASE under with various combinations of keywords. Additional searches of selected websites and reference lists also occurred. Data was extracted by a single reviewer using a dedicated template.

Results: Over 2800 potentially relevant studies were located by the searches. Some 124 studies were included in the review after exclusions were applied for non-English language or out-of-scope reports.

A substantial number of studies have now been published that have used trigger tools including the IHI GTT to measure AE rates in health care organisations. Using these tools it appears that AEs are common in hospitals occurring approximately 90 times per 1000 inpatients, 40 occasions per 100 admissions or among 30% of admissions. Most events are relatively minor and between 36% and 54% may be preventable. Adverse drug event (ADE) rates vary considerably when assessed by means of the tools but may be as high as 29 per 100 admissions or 38 per 1000 inpatient days.

As there is no true gold standard, the accuracy of trigger tools cannot be reliably ascertained. Using medical record review as the standard trigger tool appears to be an accurate method to detect iatrogenic harm with high sensitivity and specificity reported in some but not all studies. The tool also appears to be an efficient method to detect harm with high positive predictive values recorded in some studies. Assessments of the reliability of the tools suggest that there is moderate agreement amongst reviewers in their assessments of the occurrence of AEs. Limitations associated with this level of agreement may impact on the ability of the tool to reliably detect changes in patient outcomes at an organisation over time. Trigger tools are the best single method to detect harm and appear considerably more effective and cost-effective than voluntary reporting and pharmacist review to detect AEs. However, it seems likely that trigger tools also identify different types of harm compared with these methods and a comprehensive review of patient safety in an organisation should adopt multiple methods. Most experience with trigger tools has occurred in relation to ADEs, while experience is accumulating with intensive care and surgical patients.

Conclusions: Trigger tools, particularly the IHI GTT, assist organisations to measure and monitor harm. They appear to be the most accurate and efficient method to identify AEs. Further work is needed to confirm their reliability. Trigger tools are most effective when combined with other measures and patient safety interventions in the reduction of iatrogenic harm.

Background

Patient safety

Patient safety has been in the spotlight since the publication of studies documenting significant rates of adverse events (AEs) amongst hospital inpatients in many developed countries (Thomas et al 2000; Wilson et al 1996; Vincent et al 2001; Davis et al 2002; Baker et al 2004). An essential part of improving patient safety is the need to be able to monitor the level of safety so that areas can be prioritised and interventions mounted. Once underway, monitoring to assess impact is important. The main source of data to assess patient safety has been the medical record. The large resources required to evaluate the whole record using the methods developed by the original AE studies have led to an increasing interest in the use of triggers – prompts that direct the evaluation of the record and help screen for whether an AE is likely to have occurred. The increasing use of electronic medical records and the provision of electronic triggers have fuelled this interest.

Trigger tools

The term ‘trigger tool’ appears to have been first used by Jick (Jick 1974) to describe sentinel words that may identify AEs in the medical record. It has subsequently been adopted by Classen (Classen et al 1991) to describe a method to detect potential adverse drug events (ADEs). In Classen’s system, computer software linked to both the patient’s electronic record and hospital pharmacy system was used to identify key triggers (eg, antidotes or abnormal laboratory values) suggestive of medication-related error. In a trigger system, when a trigger flags a record, there is a method to further examine with a more detailed chart review to evaluate the presence of an AE. The original studies that documented the prevalence of AEs in hospitals in the United States (Thomas et al 2000; Brennan et al 1991), the United Kingdom (Vincent et al 2001), Australia (Wilson et al 1996), Canada (Baker et al 2004) and New Zealand (Davis et al 2002) all used a stepped approach to identify AEs that began with the application of various screening criteria. Trigger tools can be seen as an extension of this approach in which a series of prompts are used to more efficiently guide the record review process.

Methods used for this review

Key objective of the review

The main aim of this review is to describe the published literature associated with the development and use of trigger tools to determine rates of harm in health care settings. The review focuses on the use of the GTT and related versions developed by the IHI in the United States (Griffin and Resar 2009).

Main approach and key audience

The review was undertaken using a standard systematic review methodology. It included a structured search for all published studies that have considered the IHI GTT and its related forms. All relevant studies are summarised and key information is presented in tabular form. The key audience for the review is health professionals looking to use the IHI GTT to complement their other information sources about potential patient harm and to inform quality improvement projects. The key function of the review is to highlight the available literature. Only limited critical appraisal of the included studies is included. Instead general comments are made about the limitations of the IHI GTT approach to measuring patient safety.

Detailed scope and methods for the review

The review describes all published studies including reviews of published studies addressing the IHI GTT and related trigger tools including versions designed for paediatric care, surgery, intensive care, ADEs and ambulatory care.

A systematic search was undertaken of the following electronic databases: MEDLINE and EMBASE. The databases were searched using a range of text keywords or Medical Subject Headings (MeSH) alone and in various combinations (trigger tool$, adverse event$, adverse drug event$, medication error$, adverse effect, detection system, surveillance, detection system, evaluation, review, screening, chart review, record review, incident report, voluntary report, incident report).

The search was undertaken in December 2012 and updated in February 2013. It was conducted without any limitations by language and it included all years from 1990 onwards. Studies in languages other than English were identified but not translated and were excluded from the review.

Further ‘snowball’ searching was undertaken of the reference lists of published studies.

A limited search of ‘grey’ literature was conducted. The search included important conference abstracts and key literature from relevant websites such as that belonging to the IHI.

After screening the abstracts all potentially relevant full-text publications were evaluated. Studies were include if they considered the use of a trigger tool system to identify patient harm and presented numeric data. The review focused on the use of the IHI GTT and any of its derivatives (specific tools for specialty areas such as paediatrics, mental health etc). It excluded studies that have not used all the stages of the GTT methodology (that is, sampling followed by screening for triggers and an assessment of whether an AE occurred). Therefore, for example, text mining studies that solely identified potential AEs but did not determine whether such an event had occurred were excluded as were studies that just assessed medical records in order to identify the presence of AEs without reference to the use of any triggers.

A structured template was used to extract relevant information from the included studies. This information included details about the study setting, sample, important methods, key results (such as ADE rate per 1000 inpatients) and authors' conclusions.

Introduction

Aims of trigger tools

Trigger tools can function either as a counting system that aims to estimate the rate of harm at an organisation or as an alerting system that aims to highlight the occurrence of a potential AE so that it can be mitigated. Global trigger systems are non-actionable notifications that generate information at the systems level rather than the patient care level. Their intention is to provide information about rates of events at an institution and enable system changes to be evaluated. Such systems tend to be retrospective and generate information about events after patient care has been delivered, usually after the patient has been discharged. By contrast, an interventionist trigger system is one that provides actionable notifications that can be used at the time of patient care to prevent or mitigate an AE. Such interventionist systems are often specific trigger systems that identify accurately a particular event at the patient care level. These systems are often concurrent so that identification can occur in a timely manner to permit immediate action to improve care (Mangoni 2012). A number of studies have investigated the positive predictive value (PPV) of these interventionist triggers with a view to improving their accuracy. Most of these triggers have been drug related (Mull et al 2008).

The use of the IHI GTT either as a method to define rates of AEs in an organisation or as an alerting system with interventionist triggers both contrast with previous medical record review methodologies that have been used primarily just for research purposes.

Ideal features of a trigger system

A trigger system should exhibit a number of features regardless of its aim. Based on Shimada (Shimada et al 2008), the system should:

- identify AEs that are important; that is, they should be prevalent, associated with significant harm and preventable
- include triggers that ‘add value’; that is, they should provide a function that is not already well served by another tool
- generate information that is relevant and timely for their intended function; that is, if they are designed for concurrent patient care their information should be clinically meaningful and quickly delivered
- have a good signal-to-noise ratio and a good cost–benefit ratio; that is, they should be accurate and also cost-effective to implement
- be feasible in a variety of settings and locations; the system must be able to be adopted by healthcare facilities in different locations with varying resources.

The IHI GTT

The IHI GTT was developed as a low-technology and low-cost alternative for identifying iatrogenic harm that did not require an organisation to operate a sophisticated computerised drug and patient management system (Rozich et al 2003). The IHI GTT was developed by a

group of experts at the IHI and Premier in 2000. The IHI/Premier tool included 24 triggers and employed manual rather than computerised review procedures. The primary aim of the tool is to estimate the prevalence of AEs within a hospital setting by using high-yield triggers based in areas important to most hospitals, such as medication, post-operative care and the emergency department (Griffin and Resar 2009). The IHI GTT focuses on harm that is injurious to patients rather than error or failures in processes of care. The aim is to engage both clinicians and administrators and focus on systems that improve outcomes rather than blame individuals. The IHI GTT follows a definition of harm based on unintended physical injury resulting from or contributed to by medical care that requires additional monitoring, treatment or hospitalisation, or that results in death (Griffin and Resar 2009). The tool focuses on harm that occurs during the active delivery of care; issues related to substandard care are omitted. Thus the tool considers acts of commission and not omission. For example, a patient not appropriately treated for hypertension who sustained a stroke would not be included whereas the patient who was treated with anticoagulants who suffers an intra-cerebral bleed would be. To be included an AE must have occurred before and during and be detected during and/or after the index admission. Although preventability is important the IHI GTT does not include any assessment of the preventability of an event, merely the identification that it was an unintended consequence of medical care. The developers consider that preventability is rapidly changing with new innovation and it is therefore meaningless as the definition of included events would be constantly changing over time. The severity rating used in the IHI GTT is based on the classification system developed by the National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) Index for Categorizing Errors (MERP 2013). The IHI GTT only counts events where harm to the patient occurred. Category A–D events are omitted and only categories E and F (temporary harm), G (permanent harm), H (intervention to save life) and I (death) are included.

Methods to identify patient safety events and their advantages and disadvantages

The measurement of patient safety helps identify the magnitude of the problem in a system or hospital and can be used to compare organisations, change payments or monitor the impact of interventions (Suresh 2012). Measures of harm should be presented as a rate (rate of AEs per patient, inpatient day etc). However, obtaining such rates is challenging because many events are rare, most lack standardised definitions, few surveillance systems exist to identify numerator events of interest and systems may not be available to generate reliable denominator numbers. Problems with counting the number of events (numerators) are compounded by the need for some subjective judgement about whether an event was related to medical care or the underlying illness. Issues exist too with varying delays that may occur between treatment and the development of harm. Similarly, accurately measuring the denominator can be problematic as ideally the actual time at risk for each patient rather than just the number of patients would be assessed in relation to any particular event. In practice the number of AEs located by any method may simply reflect the resources spent looking for their occurrence. Finally, modern understanding about the causation of error and the importance of systems to prevent errors from leading to harm have led many experts to agree that attention should be directed at identifying and eliminating harm rather than focusing on error (Vincent 2010). Furthermore clinicians and administrators can unite in the pursuit of harm reduction whilst error identification is more problematic (Sharek and Classen 2006).

A number of methods exist to assess the extent of harm occurring within an institution. Conventional attempts to quantify harm include incident reports, chart reviews and observational data. All of these methods have various limitations. Incidents are notoriously under-reported by staff perhaps because many fear punishment. Chart reviews and observational studies are highly resource intensive. Indicators based on administrative data may lack clinical relevance while cases identified from malpractice claims, autopsy series or complaints may not be representative. Trigger tools have emerged as a strategy to avoid many of these limitations. The IHI GTT has been promoted as the best available single method to determine rates of harm at healthcare settings (Parry et al 2012), although a variety of methods may be necessary in order to obtain a comprehensive picture of patient safety within an organisation (Hogan et al 2008).

Table 1: Review of methods to detect harm in health care settings

Method	**Advantages**	**Disadvantages**
Chart review	Easy to assess, especially if electronic records.	• Expensive process • Needs trained reviewers • Difficulty with standardising judgement • Unable to detect AEs not documented in record.
Automated trigger tools	• Can search large volume easily • Can generate periodic reports automatically • Can be real time.	• Unable to detect all events • Resources needed to set it up • Still needs chart review to confirm AEs.
Administrative data	• Data readily available • Easy to analyse.	• Coding vagaries • Incomplete data • Data divorced from clinical context.
Malpractice claims	Multiple perspectives (legal system).	• Bias from hindsight and reporting • Non-standardised source of data.
Observation	Potentially accurate and able to detect errors in real time.	• Expensive • Need trained observers • Hawthorn effect • Threaten staff or patient confidentiality • Hindsight bias • Large amount of info.
Autopsy series	Familiar to providers.	• Infrequent and non-random selection • Hindsight bias • Reporting bias • Focused on diagnostic error.
Mortality and morbidity conferences	• Familiar to providers • Cases selected more likely to have errors.	• Error may not be acknowledged easily • Hindsight bias • Reporting bias.
Complaints	• Multiple perspectives • Few resources.	• Reporting and hindsight bias • Need process to reliably investigate events.

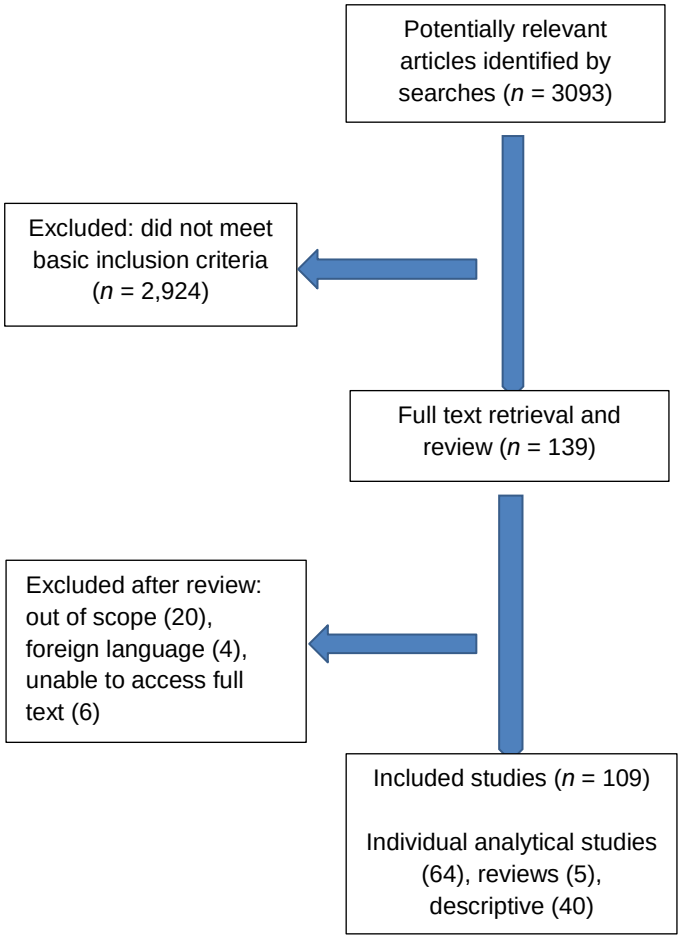
Based on Thomas and Peterson (2003) and Suresh (2012)

Results

Most studies (94) were identified from the MEDLINE and EMBASE databases. A further 42 studies were located from the reference lists of the identified studies and three were found on key websites. Some 30 studies were excluded. Most of the excluded studies were deleted because they were out of scope, not English language or were not available in full text. In addition studies that have used trigger tools as an outcome assessment tool were also excluded. That is, studies that used triggers as a method to assess the impact on patient safety from other interventions, eg, the introduction of a protocol by de Boer (de Boer et al 2011) were excluded. Among the 109 studies included in the review, 64 presented data related to the use of the GTT, 5 were reviews of other studies and 40 primarily described aspects of GTT development or process.

Appendix one presents descriptions of the key methods and results in tabular form for the included individual studies that present data related to the use of the GTT.

Figure 1: Flow chart of the review



Descriptive studies outlining the IHI GTT

Specific components of the IHI GTT approach

Key components of the IHI GTT have been outlined (Griffin and Resar 2009; Adler et al 2008) and include the following:

1. Random sampling of small number of hospital charts (typically 10 records every second week or 20 records every month per hospital).
2. First evaluation of the charts independently by two trained reviewers (usually nurses but can be other health professionals) using a predetermined list of 54 defined triggers.
    a. A small number of high-yield triggers are used that are closely linked to important AEs.
    b. The review of each record is undertaken in a structured manner within a defined period (usually 20 minutes).
3. A positive trigger initiates a more comprehensive review of the relevant part of the medical record to determine whether or not harm occurred. The team then discusses the findings together.
4. Second-stage review of trigger positive charts for AEs by physicians/pharmacists. The physician does not generally review the record but does authenticate the consensus findings of the previous reviewers in relation to the AEs and the severity rating, and answers questions from reviewers in the previous stage. The physician remains the final arbitrator and does not have any time limit for his/her determination.
5. Members of the review team are trained and use standardised definitions.
6. Results are presented using statistical process control.

Descriptive information outlining the IHI GTT process

A number of publications have outlined the development of the IHI GTT and its application at a health care organisation, how it may be utilised by a host organisation or key aspects of the refinement of the tool.

Table 2: Published information outlining the GTT process

IHI Website (www.ihi.org)	The website provides copies of simple GTT templates, description of the methodology for their use and instructions for reviewers in how to undertake retrospective reviews of inpatient records using triggers to identify potential AEs. Instructions and documentations are provided for collecting the data to track: AEs per 1000 patient days, AEs per 100 admissions, percentage of admissions with an AE.
Cymru NHS Wales (Anonymous 2010b)	Outlines the use of the United Kingdom (UK) GTT for hospital use and primary care. The document presents information about the tool and its use in the United States as well as more details about its application to improve care at Glan Clwyd Hospital in Wales. The document presents in some detail a step-by-step guide to the use of the UK GTT in hospitals and primary care. Additional material is provided about how to present the results from the tool. A variety of resources are included such as various GTT forms, definitions and the answers to common questions.
Griffin et al (2009)	The white paper provides more detailed information for identifying AEs and measuring the rate of AEs over time. The generic process is based on the preceding Trigger Tool for Measuring Adverse Drug Events developed in 2000 by health professionals in the United States. The authors outline the merit in undertaking an ongoing measure of harm and suggest that random sampling is a pragmatic approach to guide patient safety improvement in hospitals. This reference document for the IHI GTT describes the background to its development and outlines the methods needed for their implementation. All of the triggers from each set are defined and specified. Requirements for training are documented along with tips to assist organisations introducing the tools. A series of stories are presented from experienced organisations as case histories. Appendices give answers to frequently asked questions along with worksheets for the application of tools and answer sheets for the training records.
Resar et al (2003)	This descriptive article outlines the nature of harm and distinguishes between errors and AEs. Methodologies for measuring harm are described and the trigger tool methodology is detailed. The article outlines the history, application and selected impacts from the use of the trigger tools and suggests future research directions as well. Benefits of the tool are listed as the generic approach to measuring harm and flexibility with suiting low tech assessment as well as computerised clinical environments. An appendix describes an intensive care unit (ICU) trigger tool checklist and outlines the rationale for the ICU trigger tool.
Classen et al (2008)	Describes the development and evaluation of the GTT methodology. They used a two-stage record review process based on a refinement of the Harvard Study methodology to review 15 training records. In preparation, reviewers read the IHI GTT white paper which outlines the methodology and completed the training records. The authors then introduced a 2-hour formal training session in the interpretation and use of the IHI GTT. Reviewers then each reviewed an additional 50 records using the same methods. Statistically significant levels of improvement in inter-rater reliability were demonstrated. Initially agreement ranged from 38.5–76.9% with kappa ranges -0.077–0.512. After training, agreement with test records ranged from 66.7–93.9% with kappa ranges 0.164–0.703.
Adler (2008)	This article provides a step-by-step guide to obtaining leadership agreement, team training and implementing the GTT based on the author's experience at a 2000-plus bed set of eight Florida hospitals with more than 105,000 admissions per year. Seven key steps are outlined: getting started, developing a team, training, review of 10 records, development of processes, briefing leadership, implementing formal programme, setting up organisational flow. Resources are provided and implementation data such as costs and reproducibility data are provided.
Rozich (2003)	This report describes the trigger tool in detail: its characteristics and utility, the way in which it was tested, and the results of the tests. The paper outlines the feasibility of training individuals to use the trigger tool methodology efficiently, the training requirements, and describes the extent and scope of the ADEs identified in different inpatient organisations. Limitations of the tool are outlined and the appendices describe the chart review sheet, the chart review procedure and the process of investigation of a positive trigger.
Handler (2010)	Outline of expansion of trigger tools set to identify ADEs in the nursing home setting. Outline of trigger tool process to identify ADEs for use in nursing homes.
Kaafarani (2010)	Description of process to determine set of trigger tools for surveillance of AEs in outpatient surgery. The process involved a systematic review that identified 745 available trigger algorithms, followed by focus group discussions about key features of a trigger. A preliminary set of triggers was refined by a Delphi panel process down to a final set of five. The set were: same day surgery and subsequent emergency department (ED) visit, same-day surgery and unscheduled readmission, same-day surgery and unscheduled procedure or reoperation within 30 days, scheduled same-day surgery and hospital length of stay > 24 hours, same-day surgery and postoperative lower extremity Doppler with International Classification of Diseases (ICD) code for deep vein thrombosis (DVT) or pulmonary embolism (PE) within 30 days.

De Wet (2011)	Outline of trigger tool approach to screening electronic health records in primary care. The article covers: the sampling of medical records, methods of how and why the tool may be applied in primary care and provides a description of what action should be initiated from the review.
Mull (2011)	Description of modified Delphi process used to establish consensus on the definition of 6 outpatient trigger tools to determine ADEs.
Matlow (2005)	Outline of development of paediatric version of the GTT to identify AEs in Canadian paediatric hospitals.

Developing experience with the use of the GTT

The IHI GTT has now been used in numerous countries in North America (Sharek 2009; Matlow et al 2012), the United Kingdom (Anonymous 2010; Franklin et al 2010), Europe (Anonymous. 2011; Von Plessen et al 2012), Scandinavia (Danish Safer Hospital Programme 2012), Asia (Rajesh et al 2012; Asavaroengchai et al 2009), Africa (Fayed et al 2009) and Australia/New Zealand (Seddon et al 2013). Variations of the GTT have been produced for use with surgery (Griffin and Klassen 2008), oncology (Lipczak et al 2011a), primary care (De Wet and Bowie 2009), medication safety (Rozich et al 2003), paediatrics (Agarwal et al 2010), nursing homes (Handler and Hanlon 2010), intensive care (Resar et al 2006) and neonatal care (Sharek et al 2006). A number of large health care organisations, such as Kaiser Permanente, have now amassed considerable experience with the tool (Lau and Litman 2011).

Support for the use of GTTs

With the proliferation of the types of tools available for use and the increasing number of countries employing these tools it is apparent that the trigger tools have received widespread approval. Supportive opinion articles by leading professionals at key organisations (Suresh 2012; Lau and Litman 2011; Beyea 2005; Leape 2007) and editorials in prominent journals have added their endorsement (Stockwell 2010; Mack and Brilli 2007). Key cited advantages for the use of trigger tools in general and more specifically the GTT in particular (Griffin and Resar 2008; Resar et al 2003) include:

- the inclusion of a sampling strategy that can help ensure that a representative assessment of harm within an organisation can be captured and enables results to be more readily generalised across an organisation
- the guided decision-making process that helps to more consistently identify harm
- the use of a ‘low tech’ approach to sampling and event monitoring – a sophisticated electronic patient record system is not needed
- a focus on high-risk areas such as medication and post-operative care where events are most likely to occur
- a pragmatic approach to record review that enables reviews to be undertaken in a short amount of time (up to 20 minutes)
- the surveillance of events that are tightly linked to enable a more powerful strategy to reduce injury
- a tiered approach that may increase the likelihood that harm will be accurately detected

- the inclusion of process measures that may be ideal pointers to adverse outcomes – such as abnormal international normalised ratio (INR) measures for anticoagulation therapy
- a focus on training and standardised procedures to help increase the reliability of any determinations
- the presentation of results as a rate which can be graphed with a control chart to readily present trends over time that may be readily understood by a wide audience.

Thus the GTT aims to provide consistent, reliable, relevant and accurate information about the occurrence of harm at low cost.

General limitations of trigger tools and the studies that have examined them

A number of limitations have been identified with the use of the IHI GTT. One problem is that the tool has often only been applied retrospectively after care has been provided rather than concurrently. Thus the tool may exaggerate the frequency of events that may not be clinically important, so more events may be recorded than would potentially be identified if the key issue was whether some change needs to immediately occur to patient care as a result of the notification. Another issue is that the determination of an event can be made by staff remote from the care of the patient who may not always be, at least for the first stage of the tool, clinicians. In addition, even with the provision of structured criteria, the determination of whether an event has occurred still requires some subjective assessment. The subjectivity of the assessment means that reviewers may be unlikely to make the same assessments consistently over time or that different reviewers may vary in their judgements about whether an event has occurred. The tool methodology involves the assessment of only a small number of case notes per month and the ability of such a small sample to give an accurate estimate of the safety of care over a large organisation is unclear (Lessing et al 2010). Furthermore, the triggers are limited in number and scope – not every facet of patient care can be evaluated by them. The process of limiting case note review to just 20 minutes and the total reliance on just the medical record to ascertain whether an AE has occurred are all potential limitations. Medical record review has become the gold standard for the determination of the frequency of AEs but it is an imperfect gold standard. The medical record does not contain all the information about what happens to a patient. The medical record is entirely dependent on the awareness and willingness of the treating clinicians to accurately and completely identify and document patient management. Similarly, it is largely limited to inpatient care and does not include information about events that become apparent after discharge except for readmission. Many studies have been conducted without the assessment of AEs by the gold standard. Thus the assumption is made that the AEs identified in the study were the sum total of all the events. In addition, many of the studies have been undertaken at major tertiary hospitals and it is unclear how representative the results may be to hospitals in other countries or other types of facilities, although it should be noted that increasing experience with the tool across a range of settings is mitigating this concern. The results from several studies suggest that different AE identification tools may actually locate different types of AEs. For example, research at one United Kingdom hospital illustrates that there was relatively little overlap using seven different methods to identify AEs.[19] Thus the GTT may not be the best tool at identifying all types of AEs but instead

may be the most proficient at locating a certain range of patient harms. In addition, critics have suggested that by including AEs associated with temporary harm and all events regardless of their preventability the results from the tool may over-inflate estimates of iatrogenic illness. Comparative studies suggest that trigger tools locate the highest proportion of AEs compared with other methods (incident reporting, patient complaints, clinical indicators). Therefore the tool may have high sensitivity – however, it is unlikely to be high enough that a negative result would effectively rule out an AE. Many studies provide estimates of the PPV of the individual or collective triggers. The PPV of a tool is an important measure of performance. It describes the probability that a positive trigger accurately represents a true event. Such a measure of AE yield of triggered events is largely an assessment of efficiency. There are, however, two main problems with only presenting information about the PPVs of a tool (Nebeker 2008). Firstly, the assessment does not provide any measure on how many events the trigger succeeds or fails to flag but instead only bases its estimate on the rate of positively identified flags. Secondly, the PPV is highly influenced by the prevalence of AEs. Therefore a low PPV may be due to poor trigger performance, low event rates or a combination of both. The PPV changes markedly with different prevalence rates, especially at the low event rates common for AEs (Hougland et al 2006). Other criticisms of the trigger tools include their sole focus on errors of commission while ignoring errors of omission such as diagnostic errors or failures to provide better alternative forms of management. Finally, even though the tool is faster than medical record review it still requires manual chart review and remains relatively labour intensive and needs resources to be made available from any organisation wishing to undertake the work.

Limitations of this review

Although a number of different terms were used to search for relevant studies it is possible that some were not identified. The absence of a MeSH directly related to trigger tools made searching in Medline more difficult. A small number of studies were located but could not be retrieved. Studies not in English were excluded. Searching of the grey literature was limited and only a small number of websites were examined. Individual studies were described but critical appraisal of methodological quality was mainly presented in relation to the collective of included studies. Readers are referred to the individual studies in order to base their decision-making.

Literature describing use of the IHI GTT at hospitals

Reviews of the literature related to trigger tools

Five reviews have assessed the literature broadly related to the use of trigger tools to measure AEs. The three most relevant reviews were specifically focused on the use of trigger tools while another considered the use of trigger tools as one of four methods to determine ADEs and the fifth reviewed a range of specific pharmacy and laboratory signals to detect ADEs. Among the reviews that examined trigger tools, the study undertaken by The Health Foundation in 2010 (Anonymous 2010a) provided a structured search of a number of relevant databases, although the authors made it clear it did not satisfy the requirements of a systematic review. No critical appraisal of individual studies was provided, although abstracts describing key studies were presented. The review by Mull et al (2008) focused on the use of trigger tools to estimate rates of AEs and some 45 studies were cited – however, no information was provided about the search parameters and review methodology. Finally, Doupi (2012) examined the use of trigger tools only in the context of electronic health records and only nine studies were considered with such a narrowly focused review.

Table 3: Reviews assessing the use of trigger tools to measure AEs

Author	Aims	Methods	Results	Key conclusions by authors
Anonymous (2010a)	Rapid collation of empirical research on the topic	Stated as not being a systematic review. A search of MEDLINE, EMBASE, ERIC, Cochrane Library, Controlled Trials Register, IHI and Health Management Consortium was undertaken along with reference lists and websites at April 2010.	27 studies identified	There is a surprising lack of evidence about the effectiveness and utility of the tools but a lack of evidence does not mean a lack of effectiveness. The published evidence describes the tools and outlines their application. Studies of utility were based on relatively large samples and multiple hospitals in the United States. The literature generally describes the use of tools to generate rates of AEs for a large population rather than documenting small-scale use at an individual organisation, which is how the tool tends to be applied in the United Kingdom.
Doupi (2012)	Review trigger tool literature and literature related to ADEs for electronic health records	Staged review searches made of websites and PubMed using 'triggers' and 'patient safety'. Snowball searching of references.	9 studies included	The trigger tool is important for identifying events that would not have been noticed by standard methods (incident reports, pharmacy interventions). Controversy exists over the reliability of the tool due to limited validation. The tool has been used in a series of local variants and inter-rater reliability and the use of the tool for benchmarking between organisations may be limited.

Author	Aims	Methods	Results	Key conclusions by authors
Mull et al (2008)	Review trigger literature and gaps	Limited information about methods. Search period stated to be 'up to end 2007'. No information provided about which databases were searched and no information about data extraction methods. No assessment was undertaken about the quality of the information obtained.	45 studies identified	Specifies the development of accounting trigger systems (ie, ones to estimate rates of AEs). Reviews literature related to specific AEs. Most specific triggers relate to medications (*n* = 364). 23 ADEs had > 5 triggers. Triggers varied in the amount of detail or type of data used to detect an AE. Specific triggers related to medical mismanagement are specified. A list of surgical AEs targeted by triggers is also provided. Gaps for future research are outlined.
Meyer-Massetti et al (2011)	Compare accuracy, efficiency and efficacy of four main methods to determine ADEs.	PubMed, EMBASE and Scopus databases searched 2000–2009 with combinations of text terms. No language restriction. Reference lists were checks and selected websites. Data extraction undertaken by 2 reviewers.	28 studies were included. 5 studies compared trigger tool with chart review and 2 incident reporting with trigger tool. Incident reports identified least number of drug-related problems. Among the studies comparing chart review and trigger tools 2 report higher drug-related problem rates with triggers and 3 the reverse. The number of drug-related problems detected by trigger tools compared with chart review related to the specificity of the triggers. There was little overlap in the drug-related problems found by the different methods. The overlap between trigger tool and incident reported events was 0.5–10%. Incident reporting was less sensitive than trigger tools. Using trigger tools was the most time-efficient method of the 4 when the trigger used had been already validated. The start-up costs were high for trigger tools but they were less expensive than chart review.	All four methods have different strengths and weaknesses. Overlap between the methods in identifying drug-related problems is minimal. Using trigger tools was the most effective and labour-efficient method. Incident reporting identified the most severe events.
Handler and Hanlon (2007)	Review pharmacy and laboratory signals used by clinical event monitor systems to detect ADEs in adult hospitals.	Search of MEDLINE, CINAHL, EMBASE 1985–2006. Two reviewers assessed studies using standardised forms. Pooled PPVs calculated if no significant heterogeneity. However, no examination was undertaken of quality of included studies and some information (eg, administrative data was excluded).	12 studies included. PPVs ranged from 0.03 (0.03–0.03) for hyperkalaemia to 0.50 (0.39–0.61) for low levels of quinidine. Medication levels (range 0.03–0.5) and abnormal laboratory values (range 0.03–0.27) had generally higher PPV values than antidotes (range 0.09–0.11).	Findings useful for clinical information systems and decision-support tools to develop or improve clinical event monitors to detect ADEs by prioritising signals with highest PPVs.

Estimates of AE event rates based on the IHI GTT and associated trigger tools

Twenty-seven studies have described the application of the IHI GTT or related trigger tools to assess the rate of AEs among patients in hospitals (26 studies) and outpatient settings. Almost all of the studies have either used the IHI GTT or have employed a modified version of it. That is, they have cited the IHI GTT or a publication that has employed it as part of their methodology. Most (16) were conducted in the United States, three were based in Canada, two in each of Denmark and Sweden. Two locations were not identified and single studies were conducted in other localities (Thailand, Scotland). Study sample sizes varied widely and included between 10,000 and 16,172 patients.

Resar (2009) has suggested that the average AE rate identified by trigger tools was 90 per 1000 inpatient days, 40 per 100 admissions, and 30% of admissions were associated with at least one AE. The results from the 26 hospital studies are broadly consistent with these three estimates in relation to general but not intensive care inpatients. The rate of AEs ranged from 27 to 99 per 1000 inpatient days for general inpatients but was considerably higher in two studies undertaken at intensive care units (ICUs) (Agarwal et al; Larsen et al 2008). The average AE rate across all studies was 115 per 1000 inpatient days – however, the average when restricted to just general inpatients was 44. The number of AEs per 100 admissions among general inpatients was 17–49 – however, it was higher (74) for one group of intensive care patients (Sharek et al 2006) and it was lower (6.4 per 100 admissions) when just assessed among patients with a short length of stay (< 3 days) (Kenneley et al 2013). The average number of AEs per 100 admissions across all studies was 38. The percentage of admissions with an AE was the most variable measure. The percentage ranged from 6% to 74% among inpatients and the average was 33%. Once again the result was generally higher when assessed among intensive care patients.

Based on the results from the four general inpatients studies (Von Plessen et al 2012; Asavaroengchai et al 2009; Griffin and Classen 2008; Classen et al 2011) that assessed severity, most AEs were minor and relatively few (< 16%) were associated with permanent harm, required life-saving treatment or had been fatal. Two of the three ICU studies that considered severity reported higher rates of severe harm (10–29%) (Agarwal et al 2010; Resar et al 2006). One study based on elderly patients in primary care in the United States reported that 17% of patient charts were associated with a severe AE (Singh et al 2009).

Most (four) assessments of the preventability of AEs were conducted in the intensive care setting. Between 36% and 54% of events were judged preventable. Two studies of general inpatients observed that between 58% and 72% of AEs were preventable and a study of elderly primary care patients concluded that 30% of AEs were preventable.

One study identified that 40% of AEs were present on admission and between 5% and 12% of AEs related to care that was not provided (Kenneley et al 2013). The same study determined that 91% of AEs occurred among patients admitted for at least three days.

Table 4: Estimates of the rate of AEs using the GTT and related trigger tools

Reference	Setting	Sample	IHI GTT	AE per 100 admissions	% of admissions with an AE	AE per 1000 inpatient days	Other results
Kenneley et al (2013)	8 US hospitals	16,172	Yes	6.4–27.1			72% preventable, 40% present on admission
Classen et al (2011)	3 US hospitals	795	Yes	49	33	91	7% severe
Lipczak et al (2011a)	5 Danish hospitals	572	Yes (variant)		45		
Huddleston et al (2011)	1 US hospital	1711	Yes		38		
Landrigan et al (2010)	10 US hospitals	2341	Yes	25	18	57	13.8% severe
Kandpal et al (2012)	Unidentified hospital	260	Yes		74		
Von Plessen et al (2012)	5 Danish hospitals		Yes (variant)		25	60	4% severe
Good et al (2011)	12 US hospitals	2369	Yes (variant)	51	40	68	13.4% severe
Zimmerman et al (2011)	1 Canadian hospital	1817 deaths	Yes (variant)		14		
Asavaroengchai et al (2009)	1 Thailand hospital	576	Yes (variant)	41	–	50	4% severe, 58% preventable
Sharek (2009)	10 US hospitals		Yes	17.2–36.6		–	
Levinson (2010)	Various US hospitals	278	Yes	33.5			
Schildmeijer et al (2012)	5 Swedish hospitals	50	Yes (variant)		–	27–99	
Szekendi et al (2006)	1 US hospital	327	No		74		15% severe
Naessens et al (2009)	1 US hospital	235	Yes	27.7		–	
Surgical							
Griffin and Classen (2008)	11 US hospitals	854	Yes	16	14.6		8.7% severe
Paediatric							
Matlow et al (2011)	6 Canadian hospitals	591	Yes (variant)		15.1		
Matlow et al (2012)	22 Canadian hospitals	3669	Yes (variant)		9.2		
Kirkendall et al (2012)	1 US hospital	240	Yes	36.7	25.8	76	
Lander et al (2010)	1 US hospital	553	Yes (variant)		6.1		

ICU							
Sharek et al (2006)	15 US neonatal intensive care units (NICUs)	749	Yes (variant)	74	–	32	54% preventable
Resar et al (2006)	54 US hospitals	12,074	Yes		54	113	2% severe
Nilsson et al (2012)	1 Swedish hospital	128	Yes (variant)	32	19.5	–	54% preventable
Agarwal et al (2010)	15 US paediatric intensive care units (PICUs)	734	Yes		62	286	10% severe, 45% preventable
Larsen et al (2008)	1 US PICU	259	Yes (variant)		59	530	3% serious, 36% preventable
Pravinkumar et al (2009)	1 unidentified hospital	10	Yes		30	–	
Outpatient/General practice							
De Wet and Bowie (2009)	5 practices Scotland	2251 consultations	Yes		2 per 100 consultations	–	4% events severe

Additional information about AEs identified from trigger tools studies

Several studies have reported that inpatient AEs frequently have occurred soon after admission and older patients and those with more co-morbidities are generally at greater risk (Kenneley et al 2013; Classen et al 2011; Huddleston et al 2011). Patients identified with an AE by the trigger tools in one large study of adult inpatients were older, had higher mortality and longer length of stay (Classen et al 2011). Patient care processes, surgery and medication were common areas associated with high rates of AEs located by trigger tools among inpatients (Asavaroengchai et al 2009). For outpatients, prescribing was considered to be most important area related to harmful events (De Wet and Bowie 2009). Health care associated infections, hypoglycaemia and pressure sores were the most common harmful events identified in one study related to paediatric intensive care units in the United States (Sharek et al 2006).

Assessments of adverse drug event (ADE) rates and adverse drug reaction (ADR) rates based on the GTT and related tools

A number of studies have used trigger tools (the IHI GTT and related variants) to estimate the rate of adverse drug events (ADEs) and adverse drug reactions (ADRs) at hospitals and outpatient clinics. An ADR is an adverse outcome that can be attributed to some action of a drug; an ADE is an adverse outcome that occurs while a patient is taking a drug, but is not necessarily attributable to it (Schade et al 2006). Thus, ADEs can be regarded as the larger grouping and ADRs are the subset of ADEs with a causal link to a drug. ADRs likely contribute substantially to the incidence of ADEs and their reporting is closely linked (Schade et al 2006). Among the 20 studies that have estimated the rate of ADEs, most (13) have been located in the United States. The other studies were located in a variety of countries

including New Zealand (two studies). Study samples have varied considerably – between 20,000 and 36,653 patients have been included depending at least in part on whether manual or automated methods were used to identify events. There is wide variation in the rate of ADEs presented in the studies regardless of which measure is considered (ADEs per 100 admissions, percentage of admissions with an ADE, or ADE rate per 1000 inpatient days). Between 2 and 28.9 ADEs per 100 admissions have been recorded among adult inpatients, 3–31% of admissions have been associated with an ADE and 1–38 ADEs occur per 1000 inpatient days. A New Zealand study (Seddon et al 2013) has described the highest rates of ADEs among adult inpatients. The authors noted the result was higher than previously reported and suggested that it may relate to their inclusion of ADEs regardless of whether they occurred during hospitalisation or were present on admission.

There is some variability in the results presented by paediatric studies too. Between 1.8 and 25 ADEs have been recorded per 100 admissions and 1.6–22.3 ADEs have been noted per 1000 inpatient days. ADE rates in the intensive care setting are similar to those noted among adult and paediatric inpatients, although one study identified a very high rate (173 per 1000 inpatient days) at one hospital based on a small number of patients. ADEs appear to be relatively frequent in the outpatient setting. One study observed 60 ADEs per 100 charts at 6 ambulatory care practices serving elderly patients in New York.

The results from studies conducted among hospitalised patients suggest that most ADEs are not severe; with the exception of the study by Jha et al (1998), more than 80% of cases were relatively minor. The preventability of inpatient ADEs does not appear to be high. Less than 30% of ADEs were considered to be preventable in the five hospital studies that considered the issue. Among outpatients, ADEs may be more severe (approximately 30%) but also more preventable (40%).

A number of studies have suggested that opiates and other analgesics, anticoagulants and antibiotics were medications commonly associated with ADEs (Classen et al 1991; Seddon et al 2013; Resar et al 2006; Schade et al 2006; Zolezzi et al 2007; Ferranti et al 2008; Takata et al 2008b).

One study (Cohen 2004) presented the results from the introduction of an intensive ADE surveillance procedure using the IHI GTT at a hospital (Cohen et al 2005). The provision of ADE monitoring was associated with a three-fold reduction in medication events at the hospital.

Table 5: Assessments of ADE rates using the GTT and related trigger tools

Reference	Setting	Sample	IHI GTT	ADEs per 100 admissions	Percentage of admissions with ADE	ADE rate per 1000 inpatient days	Other results
Seddon et al (2013)	3+ New Zealand (NZ) hospitals	1210	Yes	28.9	–	38	Most ADEs minor but 18 (5.5%) severe. Morphine, warfarin and tramadol were most frequently associated with an ADE.

Reference	Setting	Sample	IHI GTT	ADEs per 100 admissions	Percentage of admissions with ADE	ADE rate per 1000 inpatient days	Other results
Kilbridge et al (2006)	2 US hospitals	900	No	4.4–6.2	3.6–4.9	6.1–7.3	
Jha et al (1998)	1 US hospital	not stated (ns)	No	–	–	9.6	50% severe, 25% preventable
Cohen et al (2004)	1 US hospital	ns	Yes	–	31 to 10	5.07–1.3	Median ADEs per 1000 doses of medication declined from 2.04–0.65 ($p < 0.001$).
Franklin et al (2010)	1 UK hospital	207	Yes (variant)	–	3.4	7	29% preventable
Yeesoonpan et al (2011a)	11 Thailand hospital	136	Yes	12.5			
Schade et al (2006)	1 US hospital	3572	No	3			27% preventable, anticoagulant, hypoglycaemic and analgesia commonly associated
Classen et al (1991; 2005)	1 US hospital	36,653	No	2.0	1.8		Analgesics, anti-infectives, cardiac drugs common
Zolezzi et al (2007)	1 NZ hospital	286	No		8.5		Morphine, anticoagulants and benzodiazepines common
Paediatric							
Ferranti et al (2008)	1 US	4711	No	1.8	–	1.6	5% severe, nephrotoxins, narcotics and benzodiazepines were commonly associated
Takata et al (2008a)	12 US hospitals	80	Yes	9.3	–	13.1	22% preventable, 3% severe, opioid analgesics and antibiotics common
Yeesoonpan et al (2011b)	1 Thai hospital	20	Yes	25	15		
Takata et al (2008b)	5 US hospitals	ns	Yes	11.2	9.1	22.3	Analgesics common, 7.6% preventable, 6.3% severe
ICU							
Resar et al (2006)	54 US hospitals	12,074	Yes			20	17% severe, narcotics, antibiotics common
Seynaeve et al (2010; 2011)	1 Belgium ICU	79	Yes (variant)			173	4% severe
Fayed et al (2009)	1 Egypt ICU	240	ns	8.8		–	5% were severe

Reference	Setting	Sample	IHI GTT	ADEs per 100 admissions	Percentage of admissions with ADE	ADE rate per 1000 inpatient days	Other results
Agarwal et al (2010)	15 US PICUs	734		4.9			
Primary care/Outpatients							
Singh et al (2009)	6 US practices	383	No	60 charts			30% severe 40% preventable
Gurwitz et al (2003)	US	30,397 person years	No			50 per 1000 person years	38% severe 42% preventable
Brenner et al (2012)	1 US clinic	516	No	17.6			54% of these ADEs occurred during medication monitoring and 45% during patient self-administration.

Assessments of the rate of ADRs identified by trigger tools

Trigger tools have been used to identify ADRs although it should be noted that the IHI GTT identifies harm (ADEs). A number of studies largely based at one hospital in Germany have reported on the use of trigger tools to locate ADRs among inpatients. ADR rates among inpatients appear common and may be as high as nearly half of admissions. Between 7% and 17% of the reactions were determined to be severe. Rates of ADRs are lower when assessed with paediatric populations.

Table 6: Assessments of the rate of ADRs using trigger tools

Reference	Setting	Sample	IHI GTT	Percentage of admissions with ADR	Serious ADRs
Adult inpatients					
Levy et al (1999)	Single hospital Israel	40	No	20%	14% severe
Tegeder et al (1999)	Single hospital Germany	98	No	18%	17% severe
Dormann et al (2000)	Single hospital Germany	379	No	8.9%	7% severe
Thuermann et al (2002)	Single hospital Germany	600	No	18%	
Egger et al (2003)	Single hospital Germany	163	No	48%	–
Dormann et al (2004)	Single hospital Germany	474	No	22.9%	–
Paediatric					
Haffner et al (2005)	Single hospital Germany	703	No	5.7%	–
Neubert et al (2006)	Single hospital Germany	439	No	6.2%	–

Assessments of the accuracy of the GTT and related tools

Twenty studies have considered the validity of the GTT or related trigger tools in relation to whether the trigger tool accurately identifies the occurrence of AEs. As there is no true gold standard for detecting AEs the accuracy of the GTT remains unknown. However, for the purposes of this review, full medical record review is considered to be the gold standard. Thus the results from the trigger tool have been assessed against those provided by a medical record review process. Most studies that have undertaken these analyses have only assessed the PPV of the tool (or individual triggers). That is, they have sought to confirm whether (or not) an AE generated from a positive trigger actually represents an episode of patient injury. Not all of the studies have conducted a full record review. Regardless of the extent of the record review, the importance of the information gained from an assessment of the PPV of the triggers is somewhat limited as the PPV of the tool is strongly influenced by the prevalence of AEs at the organisation. In order to examine the sensitivity and specificity of the tool, records without any trigger event must also be assessed in order to estimate whether negative events truly represent hospitalisations where there was no harm. Sometimes instead of a full record review of both positive and negative cases authors have attempted to ascertain sensitivity and specificity by comparison with some other method for determining AEs such as the results from pharmacist review rather than the gold standard. Relatively few studies have formally reported the accuracy of the tool with a full medical record review based on a sample of positive and negative cases. When this has occurred the number of cases considered has often been relatively small.

Sensitivity and specificity

Two main studies (Classen et al 2011; Matlow et al 2011) have examined the accuracy of the IHI GTT with full medical record review and have also included a sample of negative cases. Both were conducted in North America. The results from these studies suggest that the IHI GTT has very high sensitivity (95%) and specificity (100%) when applied to adult inpatients (Classen 2011) and relatively high sensitivity (85%) but lower specificity (44%) (Matlow et al 2011) when employed with paediatric patients. However, another study by Sharek et al (2011) reported a considerably lower sensitivity when the IHI GTT was used with adult inpatients. The study, however, did not assess the accuracy of the tool against full record review but rather only compared the use of the tool by review groups against the findings from another expert group. Two other studies have assessed the sensitivity and specificity of trigger tools to identify AEs among paediatric inpatients (Lander et al 2010; Neubert et al 2006). The studies have reported discordant results. One study was consistent with the findings of Matlow and indicated that the tool was associated with a high sensitivity (90%) but much lower specificity (20%) (Neubert et al 2006) while the other observed that among children admitted for ear, nose and throat (ENT) surgery the sensitivity of the tool was very low (17%) but the specificity was higher (82%) (Lander et al 2010).

Four other studies have considered the accuracy of trigger tools in European settings in relation to ADEs or ADRs (Franklin et al 2010; Dormann et al 2000; Thuermann et al 2002; Egger et al 2003). The single study among them that examined the accuracy of an IHI-derived tool focused only on preventable events (Franklin et al 2010). The study recorded only modest (0.40) sensitivity related to the tool. The other (non-IHI) tools (sometimes automated) recorded moderate sensitivity and specificity.

Positive predictive value

The overall PPV of the IHI GTT for adults was reported by Kenneley et al (2013) in a large study that involved over 16,000 patients. The overall PPV of the tool was recorded to be 17%. The overall PPV of the IHI GTT to identify paediatric ADEs was recorded by Takata et al (2008a) as 4%. The overall PPV of other trigger tools has been assessed and found to be 4% for preventable ADEs among adults (Franklin et al 2010), 13% for adult ADRs (Dormann et al 2000), 18% for paediatric ADRs (Haffner et al 2005), 17% for adult ADEs (Jha et al 1998) and 39% for paediatric ENT patients (Lander et al 2010).

All studies regardless of their setting or patient population have observed that there is a wide variation in the PPVs for individual triggers. Kenneley et al (2013) noted that the PPVs for the individual triggers ranged from 0 to 100%. Likewise the PPVs for individual adult triggers was found by Naessens et al (2011) to vary from 26% to 80%. The PPVs of the individual paediatric GTT triggers was recorded by Matlow et al (2011) to be from 0 to 88%. In smaller studies, PPVs for individual triggers have varied from 0 to 100% for adult ADEs (Franklin et al 2009), 0–100% for adult ADRs (Thuermann et al 2002), 7–100% (Singh et al 2009) and 12–96% (Brenner et al 2012) for adult outpatient ADEs, 6–62% for outpatient AEs (Rosen et al 2010) and 15–93% for paediatric ADEs (Lemon and Stockwell 2012).

The PPV of trigger tools, however, remains of only limited importance as it is dependent on the prevalence of AEs at each hospital.

Table 7: Studies assessing the accuracy of trigger tools compared with medical record review

Reference	Setting	Sample	IHI GTT	Accuracy
Adult AE				
Kenneley et al (2013)	8 general US hospitals	16,172	Yes (variant)	Trigger yield varied between 0 (4 triggers) and 100% (4 triggers). Overall trigger yield was 17.1% and surgical and medication modules provided most positive yields. Some triggers had lower PPVs than other reports suggesting some organisational refinement of the triggers is indicated (eg, mechanical ventilation had PPV = 7% in this study but 82% in the study by Naessens, 2010). Not full record review.
Classen et al (2011)	3 large unnamed US hospitals	300*	Yes	GTT was associated with 95% sensitivity and 100% specificity.
Naessens et al (2011)	4 US hospitals	1138	Yes	PPVs for triggers varied between 80% (return to surgery) and 26% (intra-op X-ray). Cases with AEs had more triggers than those without (average 4.7 vs 1.8 $p < 0.001$).
Sharek et al (2011)	10 North Carolina hospitals	202	Yes	The internal review team had higher sensitivity (49% vs 34%) and specificity (94% vs 93%) compared with the external team. No full record review.
ICU				
Sharek et al (2006)	3 NICUs	749	Yes (variant)	The mean PPV for the triggers was 0.38.

Reference	Setting	Sample	IHI GTT	Accuracy
Adult ADE/ADR				
Franklin et al (2010)	Single hospital in London	207	Yes (variant)	Overall PPV = 0.04 and 0.01 for preventable ADEs. PPVs for individual triggers varied widely from 0–100%. Sensitivity of locating preventable ADEs was 0.4.
Dormann et al (2000)	Single German hospital	379	No	Computer triggers had 74% relative sensitivity and 75% relative specificity. All 3 serious ADRs were noted by computer monitoring. The PPV of the alerts was 13%. No full record review.
Egger et al (2003)	Geriatric rehabilitation ward at German hospital	163	No	Sensitivity = 58% and specificity = 1.4%. Limited record review.
Thuermann et al (2002)	Neurology hospital in Wuppertal, Germany	600	No	PPV for the triggers ranged from 0–100%. The highest were for high INR or increased serum concentrations. Sensitivity = 45.1% and specificity = 78.9%. No full record review.
Jha et al (1998)	1 US hospital	ns	No	The PPV of the rules was 17%. The PPV of the individual rules varied from 9–28%.
Outpatient ADE				
Singh et al (2009)	6 US primary care practices	1289	No	The top nine triggers identified 94% of the AEs. The PPV of the triggers varied from 6.7–100%.
Brenner et al (2012)	1 US outpatient clinic	516	No	The PPV for abnormal values of INR was 96% but PPVs were 12% or less for the other triggers.
Outpatient AE				
Rosen et al (2010)	Outpatient US clinics	Up to 150 cases out of 17,498	No	There was a wide range in PPVs for the triggers (6–62%). Not full record review.
Paediatric				
Matlow et al (2011)	6 paediatric hospitals	591	Yes (variant)	The sensitivity and specificity were 0.88 and 0.44, respectively. The PPV for each trigger ranged from 0 to 88.3%.
Neubert et al (2006)	Single hospital Germany	439	No	Sensitivity = 90% and specificity = 20%.
Lander et al (2010)	ENT Service, Boston hospital	50	No	The trigger tool had 17% (14–20%) sensitivity, 82% (79–84%) specificity, 39% (33–46%) PPV and 59% (56–62%) negative predictive value.
Lemon and Stockwell (2012)	1 US hospital	ns	No	The individual triggers ranged in PPV from 15–92.5%.
Paediatric ADE/ADR				
Takata et al (2008b)	5 US hospitals	ns but 25,763 to 41,831 bed days per hospital	Yes	Triggers had a PPV of 16.8%

Reference	Setting	Sample	IHI GTT	Accuracy
Takata et al (2008a)	12 US hospitals	900	Yes	The PPV of the triggers was 3.7% for ADEs
Haffner et al (2005) ADR	Single German hospital	ns	No	The mean PPV of the triggers was 18.6%.

* Exact number is not stated – 795 were included from 3 hospitals but the accuracy assessment was conducted only at the single largest.

Assessments of the reliability of the GTT

Ten studies have assessed the inter-rater reliability of the GTT by comparing the results from the application of the tool by either one reviewer or evaluation team with that obtained by another. Seven studies have addressed reliability in relation to adult inpatients and three with respect to children. Seven of the ten studies were conducted in the United States and three of them included a large number (> 1000) of participants. The largest studies included study sizes of 2341 and 2008 (Landrigan et al 2010; Sharek et al 2011) patients but likely included many of the same participants. One study (Classen et al 2008) that described some of the development of the tool assessed its reliability in relation to a set of training records that included a predetermined number of AEs. The study concluded that training generated a statistically significant improvement in the ability of the assessors to reliably identify the events. The study by Naessens et al (2011) assessed the reliability of the GTT as its primary objective.

The agreement between teams in relation to their assessments of whether or not an AE had occurred has usually been described with a kappa statistic where 1 signifies complete agreement and 0 no overlap. The teams have usually assessed the same medical records at an institution, although some reports have also been conducted with external teams invited from other locations to assess the records at the hospital and compare their findings with local reviewers. Inter-rater reliability assessments between members of internal review teams working within an organisation range from moderate to very high (0.32–0.9). A similar range of agreement was also recorded with the use of the paediatric version of the tool (0.3–0.9), although two studies recorded moderate agreement (kappa 0.6). The agreement between internal and external review team members (reviewers from outside of the organisation) likewise ranged from moderate to high (0.4–0.9) in the studies. Recorded agreement between nurse reviewers and physicians in relation to the assessment of AEs was high (0.65–0.86). Agreement between nurse reviewers in relation to individual triggers was more variable and was sometimes low (0.02–0.22) particularly for triggers that required more subjective assessment (such as the determination of over-sedation) rather than objective evaluation (such as INR result > 6) (kappa = 0.76–1.0). All studies have highlighted the need for substantial training to be provided to team members and pointed to the availability of training resources on the IHI website. Despite the provision of criteria for the determination of triggers and AEs, considerable variation can occur among the judgements made by reviewers. Such variation is lessened when the same team(s) is making the assessments at one organisation but is likely to be highly problematic if the GTT is being used for making comparisons between hospitals when the results will be based on the judgements of different teams and changing team members over varying periods of time.

Another critical issue yet untested is the impact of inter-rater variation on the ability of the GTT to measure and identify variation in AE rates over time at a single institution. This issue is important because while some triggers are highly specific (eg, INR > 6) and lead to clear parts of the medical record to confirm their occurrence, other triggers are more vague and require more time and skill to identify. Thus Schildmeijer et al (2012) observed that only 7% of all AEs were located by all five reviewing teams. Another area of possible disagreement whose impact is not clear is the determination of the severity of the AE. Finally, issues may arise with the use and interpretation of the statistical control charts used to plot results.

Table 8: Assessments of the reliability of the GTT and related trigger tools

Reference	Setting	Sample	Hospitals	IHI GTT	Key results related to inter-rater reliability
Kenneley et al (2013)	United States	94	8	Yes	Moderate (kappa 0.62) for reviewer comparison in relation to AE or not assessment.
Sharek et al (2011)	United States	2008	10	Yes	Moderate (kappa 0.64) to almost perfect (kappa 0.93) agreement between internal reviewers and external reviewer team.
Landrigan et al (2010)	United States	2341	10	Yes	Kappa was 0.64–0.93 for internal review teams and 0.40–0.72 for external teams. Internal vs external reviewers kappa = 0.49. Likely to be overlap with above study.
Classen et al (2008)	United States	65	Training records	Yes	Kappa significantly improved from a range of -0.077–0.512 before training to 0.164–0.703 after training.
Naessens et al (2011)	United States	1138	3	Yes	Kappa for the triggers = 0.53–0.73 and 0.4–0.6 for AEs. The agreement between nurses and physicians for AEs was 0.65–0.77. Agreement between nurses on individual triggers varied with lower levels with more subjective measures such as over-sedation kappa = 0.11 (0.02–0.22) compared with more objective triggers such as INR > 6 kappa = 0.9 (0.76–1.0).
Asavaroengchai et al (2009)	Thailand	576	1	Yes (variant)	Kappa for the triggers was = 0.86.
Schildmeijer et al (2012)	Sweden	50	5	Yes (variant)	Weighted kappa for number of triggers team by team was 0.32–0.6. Weighted kappa for AE detection was 0.26–0.77.
Paediatric version					
Kirkendall et al (2012)	United States	240	1	Yes	Agreement between the 2 nurses for AEs was 0.63.
Lander et al (2010)	United States	50	1	No	Agreement was 0.35–0.90 for trigger categories.
Matlow et al (2011)	Canada	591	3	Yes	Agreement was 0.62 between nurses and 0.57 between nurses and doctors.

Comparisons of trigger tools with other methods to find harm

Aside from comparisons with the ‘gold standard’ (full medical record review) the relative effectiveness of trigger tools (IHI or related versions) to identify harm in health care organisations has been compared with other methods in one systematic review and 22 individual studies. The alternative methods primarily include voluntary reporting and pharmacist review although comparisons with administrative indicators and physician surveillance have also been reported. The assessment of the comparative performance of

trigger tools in relation to medical record review is considered in the section 'Assessments of the accuracy of the GTT and related tools'.

The relative ability of trigger tools to identify AEs among adults in comparison with voluntary reporting has been considered in relation to both adults (five studies) and children (two studies). Two of the adult studies also compared the return from the use of clinical indicators based on administrative data. All of the studies except two were conducted in the United States and all seven studies employed samples of less than 800 patients. Trigger tools were consistently identified by all studies as the method which identified the most patient harm. This suggests that trigger tools may have high sensitivity – however, as there is no true gold standard this cannot be confirmed. In most of the studies (six out of seven) trigger tools identified more than five times the number of voluntarily reported events and in four studies the return was over 10 times higher. Notably the two studies that employed the IHI version of the tool both consistently report that the use of triggers was markedly better than voluntary reporting (Seddon et al 2013; Nilsson et al 2012). Trigger tools also usually generated higher AE rates than indicators by a factor of at least 10. It should be noted, however, that only one of these seven studies included an assessment of the 'true' rate of AEs by means of a full medical record analysis. In the single study that also included full medical record review, trigger tools located 90% of the AEs while indicators identified 10% and only 1% were reported voluntarily (Classen et al 2011).

ADE/ADRs

One systematic review and 15 individual studies have considered the effectiveness of trigger tools in comparison with other methods apart from medical record review to detect ADEs/ADRs.

Systematic review

A systematic review by Meyer-Masetti et al (2011) has compared the accuracy and efficiency of different methods to detect ADEs. The review examined 28 studies published from 2000–2009 (see: Reviews of the literature related to trigger tools). Two studies were identified that compared trigger tools with incident reporting and the authors concluded from these studies that trigger tools identified more ADEs than reporting. In addition, the overlap in the ADEs identified from both methods was very low (5–10%) suggesting that both methods identified different types of ADEs. Trigger tools were also noted to be the most cost-effective method, although start-up expenses could be relatively high.

Individual studies

Trigger tools have been compared with other methods to detect ADEs/ADRs in adult (11 studies) and paediatric (4 studies) populations. Thirteen of the studies were located in hospital settings and one was restricted to surgical inpatients. Eight of the studies were located in the United States, four in Germany and one in the United Kingdom, Sweden and New Zealand. It should be noted that the assessment of the performance of the tool in comparison with other methods was not necessarily the primary objective of all of these studies. Among the eight hospital studies that have compared the use of trigger tools with voluntary reporting, only one study (Ferranti et al 2008) concluded that voluntary reporting identified more ADEs, even when the reporting was actively encouraged. The comparison between pharmacist review and trigger tools is more mixed. Two of four studies have observed that triggers detect more ADEs. However, the results from one study suggest that

pharmacist review may detect a considerably higher rate of ADEs compared with triggers (Franklin et al 2009). One outpatient comparison, based on large numbers of visits to New York clinics, reported that trigger tools identified more harm (Hope et al 2013). The results from two studies that both considered the return from physician surveillance with paediatric admissions were mixed (Haffner et al 2005; Neubert et al 2006), while a single study concluded that free text searching was superior to trigger tools (Gurwitz et al 2003).

Finally, a number of studies based on either adult or paediatric populations have observed that there was relatively little overlap among the ADEs identified by the different methods (Naessens et al 2009; Ferranti et al 2008; Takata et al 2008b; Jha et al 1998; Franklin et al 2009). Such a conclusion is important as it suggests that in order to undertake a comprehensive assessment of patient safety an organisation would need to employ several methods to reliably estimate the full occurrence of harm at their facility.

Table 9: Comparisons of trigger tools with other methods to detect harm

Reference	Setting	Sample	Outcome	Trigger vs	Key result: Method identifying most AEs or ADEs/ADRs
Adult inpatients AE					
Von Plessen et al (2012)	5 Danish hospitals	ns	AE	• Voluntary reporting	IHI GTT – Reported incidents varied from 3–12 per 1000 patient days and the average GTT harm rates were 60 per 1000 patient days.
Classen et al (2011)	3 large US hospitals	795	AE	• Agency for Healthcare Research and Quality (AHRQ) indicator • Voluntary reporting	IHI GTT – The GTT identified 90% of AEs. Incident reporting identified 1% and indicators 9%.
Naessens et al (2009)	US hospital	239	AE	• AHRQ indicator • Voluntary reporting	IHI GTT identified 65 AEs vs 9 reporting and 2 by indicators.
Levinson (2010)	Hospitals in 2 US counties	278	AE	• Interview of patients/family • Incident reports • Use of present on admission (POA) coding • AHRQ indicators	IHI GTT identified 90/120 AEs and POA analysis 60/120.
Adult ICU AE					
Nilsson et al (2012)	1 Swedish ICU	128	AE	• Voluntary reporting	IHI GTT found 41 AEs vs 3 voluntarily reported.
Paediatric AE					
Sharek et al (2006)	15 NICUs US	749	AE	• Voluntary reporting	Triggers identified 554 AEs and reporting 85.
Lemon and Stockwell (2012)	1 US hospital	ns	AE	• Voluntary reporting	Triggers identified 10 times more AEs.
Adult inpatients ADE/ADR					
Dormann et al (2000)	1 German hospital	ns	ADR	• Stimulated voluntary reporting	Triggers identified 2 times more ADRs.
Ferranti et al (2008)	1 US hospital	ns	ADE	• Voluntary reporting	Voluntary reporting identified 93 vs 78 ADEs.
Jha et al (1998)	1 US hospital	ns	ADE	• Pharmacist review • Stimulated voluntary reporting	GTT identified 139 ADEs vs 23 for reporting.
Kilbridge et al (2006)	1 US hospital	900	ADE	• Voluntary reporting	Triggers identified 3.6–12.3 times more ADEs.
Seddon et al (2013)	3 NZ hospitals	400	ADE	• Voluntary reporting	IHI GTT identified 128 ADEs and reporting none.
Thuermann et al (2002)	1 German hospital	231	ADR	• Pharmacist surveillance	Pharmacist surveillance detected 2 times more ADRs.
Meuthing et al (2010)	1 US hospital	ns	ADE	• Voluntary reporting	Triggers identified 65 hypoglycaemic or opiate associated events compared with 5 (7.8%) reported.
Zolessi et al (2007)	1 NZ hospital	528	ADE	• Voluntary reporting	Triggers identify 8.5% of patients with an ADE compared with 0.07% voluntarily reported.
Surgical patients ADE/ADR					
Franklin et al (2009)	1 UK hospital	93	ADE	• Ward pharmacist • Record review • Voluntary reporting	Pharmacist found 78 ADEs, with triggers and reporting 2 each.

Reference	Setting	Sample	Outcome	Trigger vs	Key result: Method identifying most AEs or ADEs/ADRs
Primary care/outpatients ADE/ADR					
Gurwitz et al (2003)	Single US practice	30,397 consul-tations	ADE	• Voluntary reporting of incidents • Free text	Free text – 37% free text search, 28.7% of ADEs identified by triggers, 11% by incident reports, 11% by discharge summaries, 12% by ED notes review.
Hope et al (2003)	33 clinics US	93,000 visits	ADE	• Pharmacist	Triggers identified more ADEs and at less cost.
Paediatrics ADE/ADR					
Haffner et al (2005)	1 German hospital	ns	ADR	• Physician surveillance	Physicians identified 101 vs 45 ADRs.
Neubert et al (2006)	1 German hospital	439	ADR	• Treating physician	Triggers identified 31 vs 23 ADRs.
Takata et al (2008b)	5 US hospitals	80	ADE	• Pharmacist • Voluntary reporting	Triggers identified 10 times more ADEs – identified different ADEs.
Takata et al (2008a)	12 US hospitals	960	ADE	• Voluntary reporting	Triggers identified 107 ADEs vs 4 for reporting.

Use of trigger tools to detect ADEs

The largest experience with trigger tools has been in the context of monitoring clinical records for the occurrence of ADEs and ADRs. This monitoring has been undertaken either by electronic or manual methods. The use of electronic methods pre-dates the IHI version of the trigger tool and relates back to key work by Classen et al (1991). One of the reported advantages for the IHI version of trigger tools has been the widened availability of the methodology to low-resource hospitals and settings where electronic records do not exist and electronic monitoring for ADEs has not yet been possible (Adler et al 2008). Twenty studies have examined the use of trigger tools to determine the rate of ADEs among adult inpatients (9 studies), hospitalised children (4 studies), intensive care patients (4 studies) or outpatients (3 studies). A further six studies have focused on the use of trigger tools to measure ADRs among inpatients while two studies have examined outpatients (2 studies). The accuracy of trigger tools has been considered by seven inpatient studies (six adult studies and one paediatric) and two outpatient studies. Fourteen studies have compared trigger tools with other methods to determine patient harm. Most (12/14) of these studies have been based on inpatient populations (eight adult and four paediatric).

Table 10: Use of trigger tools in relation to ADEs/ADRs

Use of trigger tools to determine rate of ADEs	Use of trigger tools to determine rates of ADRs	Studies assessing the accuracy of trigger tools	Comparison with other methods to determine harm
Adult inpatient	**Adult inpatients**	**Adult inpatients**	**Adult inpatients**
Seddon et al (2013) IHI manual	Levy et al (1999)	Franklin et al (2010) IHI manual	Dormann et al (2000)
Kilbridge et al (2006)	Tegeder et al (1999)	Dormann et al (2000)	Ferranti et al (2008)
Jha et al (1998)	Dormann et al (2000)	Haffner et al (2005)	Jha et al (1998)
Cohen et al (2004)	Thuermann et al (2002)	Egger et al (2003)	Meuthing et al (2010)
Franklin et al (2010) IHI manual	Egger et al (2003)	Thuermann et al (2002)	Kilbridge et al (2006)

Use of trigger tools to determine rate of ADEs	Use of trigger tools to determine rates of ADRs	Studies assessing the accuracy of trigger tools	Comparison with other methods to determine harm
Yeesoonpan et al (2011a) IHI manual	Dormann et al (2004)	Jha et al (1998)	Seddon et al (2013)
Schade et al (2006)	**Paediatric inpatients**	**Outpatients**	Thuermann et al (2002)
Classen et al (1991; 2005)	Haffner et al (2005)	Singh et al (2009)	Franklin (2009)
Zolezzi et al (2007)	Neubert et al (2006)	Brenner (2012)	**Outpatients**
Paediatric inpatient		**Paediatric inpatient**	Gurvitz et al (2003)
Ferranti et al (2008)		Takata et al (2008a)	Hope et al (2003)
Takata et al (2008a)		Takata et al (2008b)	**Paediatric inpatients**
Yeesoonpan et al (2011b) IHI manual			Haffner et al (2005)
Takata et al (2008b)			Neubert et al (2006)
ICU			Takata et al (2008b)
Resar et al (2006)			Takata et al (2008a)
Seynaeve et al (2010; 2011)			
Fayed et al (2009)			
Agarwal et al (2010)			
Primary care/Outpatients			
Singh et al (2009)			
Gurwitz et al (2003)			
Brenner et al (2012)			

Use of paediatric versions of trigger tools

Paediatric applications of the use of trigger tools, including the IHI version, to measure harm have been well described. The development and application of the Canadian form of the paediatric IHI GTT has been well documented (Matlow et al 2005; 2001) and a study outlining the considerable experience with its use (3669 cases) across 22 hospitals has been recently published (Matlow et al 2012). Three studies have assessed the rate of AEs at paediatric or neonatal ICUs (Agarwal et al 2010; Sharek et al 2006; Larsen et al 2008). Seven studies have measured the rate of ADEs (five) or ADRs (two) among hospitalised children. One of these studies included a large sample of over 4700 patients although the trigger tool was not the IHI version (Ferranti et al 2008). Six studies have assessed the comparative accuracy of trigger tools in comparison with medical record review, while the same number have reported the accuracy of the tools in relation to other methods for detecting harm. Three studies have considered the reliability of the use of trigger tools among paediatric populations.

Table 11: Use of trigger tools with paediatric patients

Use of trigger tools to determine paediatric AE rate	Use of trigger tools to determine paediatric ADE rates	Use of trigger tool to determine paediatric ADR rates	Comparisons of trigger tools with medical record review among paediatric patients	Assessments of reliability of trigger tools among paediatric patients	Comparison of trigger tools with other tools to detect harm among paediatric patients
Matlow et al (2011)	Ferranti et al (2008)	Haffner et al (2005)	Matlow et al (2011)	Kirkendall et al (2012)	Haffner et al (2005)
Matlow et al (2012)	Takata et al (2008a)	Neubert et al (2006)	Neubert et al (2006)	Lander et al (2010)	Neubert et al (2006)
Kirkendall et al (2012)	Yeesoonpan et al (2011b)		Lander et al (2010)	Matlow et al (2011)	Takata et al (2008b)
Lander et al (2010)	Takata et al (2008b)		Lemon and Stockwell (2012)		Takata et al (2008a)
Sharek et al (2006)	Agarwal et al (2010)		Takata et al (2008)		Sharek et al (2006)
Agarwal et al (2010)			Sharek et al (2006)		Lemon and Stockwell (2012)
Larsen et al (2007)					

Use of trigger tools in ICUs

Six studies have assessed the use of trigger tools to identify the rate of AEs in the ICU among adults (Resar et al 2006; Nilsson et al 2012; Pravinkumar et al 2009) and children (Agarwal et al 2010; Resar et al 2006; Larsen et al 2008). Other studies have focused on the recognition of ADEs among adults hospitalised in the ICU (Fayed et al 2009; Agarwal et al 2010, Resar et al 2006; Seynaeve et al 2011). A specially adapted version of the IHI GTT has been developed for ICU use (Resar et al 2003). Pravinkumar et al (2009) report that the IHI model can be readily adapted for use in the ICU setting.

Relatively few studies have explored the accuracy of the use of trigger tools among ICU patients in comparison with record review (one study) or other methods to ascertain harm (two studies).

The use of trigger tools suggest that AEs have frequently occurred among intensive care inpatients, many of whom (28%) suffered more than one AE during their stay (Resar et al 2006). Among both adults and paediatric patients, rates of AEs in the ICU identified by trigger tools are generally considerably higher than those located by other methods (Resar et al 2006; Sharek et al 2006; Stockwell 2010). However, most AEs were associated with only temporary harm and relatively few led to permanent harm or death (Resar et al 2006). A small number of triggers identified many of the AEs in the ICU – for example, haemoglobin drop was associated with 201 episodes of harm in one study (Nilsson et al 2012). The most common AEs in the PICU were catheter complications, uncontrolled pain, and endotracheal tube malposition (Agarwal et al 2010). Higher rates of AEs in the ICU were associated with surgical patients, those intubated and those who subsequently died. Adult inpatients with preventable events were more likely to be younger, have higher illness severity, longer stays and more likely to be surgical patients (Larsen et al 2008).

A small number of triggers (hypoglycaemia, hypokalaemia and prolonged partial thromboplastin time) also accounted for most (78%) of the ADEs (Seynaeve et al 2011). In common with AEs, most identified ADEs were not severe (96%) (Seynaeve et al 2011). Antimicrobials were also commonly associated with ADEs in the ICU (Fayed et al 2009). The days when an ADE occurred at the ICU were associated with higher nursing workloads and more severely unwell patients (Seynaeve et al 2011).

The various methods employed at Canadian ICUs to estimate the rate of AEs and ADEs has been surveyed by Louie et al (2007). Most (85%) Canadian ICUs operate a system to identify AEs and ADEs but only a minority (8%) employed a trigger tool. Most of the units instead provided a voluntary reporting system which was sometimes anonymous. Only half of the units reported that any changes to patient care had been made as a result of these measurements. The authors concluded that standardising methods to measure AEs and ADEs across the country was important for patient safety.

Table 12: Use of trigger tools with intensive care patients

Assessments of the rate of AEs at ICUs	Assessments of the rate of ADEs at ICUs	Accuracy of trigger tools when used among ICU patients	Comparisons with other methods to detect harm at ICUs
Sharek et al (2006)	Resar et al (2006)	Sharek et al (2006)	Sharek et al (2006)
Resar et al (2006)	Seynaeve et al (2011; 2010)		Nilsson et al (2012)
Nilsson et al (2012)	Fayed et al (2009)		
Agarwal et al (2010)	Agarwal et al (2010)		
Larsen et al (2007)			
Pravinkumar et al (2009)			

Use of the trigger tools among surgical patients

Twelve studies have applied trigger tools to identify AEs across a range of inpatients that have included surgical cases. These studies have included adult inpatients (7), paediatric inpatients (2) and intensive care patients (3). Some of these studies have reported that AEs may be more frequent among surgical cases (Matlow et al 2012; Asavaroengchai et al 2009) especially within 48 hours after surgery (Muething et al 2010). The findings from one study suggest that AEs among surgical cases may be more readily preventable than those occurring among medical inpatients (Larsen et al 2008). Two studies reported that an unplanned return to the operating theatre was a trigger associated with a high PPV for an AE (Kandpal et al 2012; Naessens et al 2011).

A specially modified version of the IHI GTT has been developed to assess AEs among surgical inpatients (Griffin and Klassen 2008). The surgical tool with 23 triggers considered most relevant to surgical care has been tested at 11 hospitals in the United States (Griffin and Klassen 2008). Almost 15% of surgical patients sustained an AE; 8.7% of these AEs were severe – requiring life preserving intervention or associated with either permanent harm or death. However, this tool has not been extensively evaluated. More experience has been accumulated with the IHI GTT applied to groups of patients that include surgical admissions recognising that the IHI GTT includes a surgical care module (Asavaroengchai et al 2009; Kenneley et al 2013; Kandpal et al 2012; Pravinkumar et al 2009). Other

researchers have adapted a modified version of the IHI GTT and then applied it to groups of inpatients that have included surgical admissions (Matlow et al 2011). A version of the trigger tool was developed specifically to evaluate the occurrence of AEs related to ENT surgical care (Lander et al 2010). Although the tool was useful to identify most AEs it did not reliably detect complex cases.

Trigger tools have also been used to detect ADEs among surgical inpatients (Franklin et al 2009; 2010). However, they were associated with a large number of false positives and it was suggested that their sensitivity needed to be improved before they were ready for more widespread use in that setting. A version of trigger tools has been developed for use with ambulatory surgery (Rosen et al 2010). The tool was applied to three large health care organisations in the United States and between 1% and 22% of cases were categorised as being associated with an AE (Rosen et al 2010).

Table 13: Use of trigger tools with surgical patients

Assessments of the rate of AEs among inpatients including surgical patients	**Assessments of the rate of AEs among primarily surgical patients**	**Assessments of the rate of ADEs among inpatients including surgical patients**	**Assessments of the rate of ADEs among primarily surgical patients**	**Assessments of the reliability of trigger tools including surgical patients**	**Assessments of the accuracy of trigger tools including surgical patients**
Adults					
Asavaroengchai (2009) (IHI GTT)	Griffin and Classen (2008)	Jha et al (1998)	Franklin et al (2009)	Kenneley et al (2013) (IHI GTT)	Kenneley et al (2013) (IHI GTT)
Kandpal et al (2012) (IHI GTT)	Lander et al (2010)	Meuthing et al (2010)	Franklin et al (2010)	Naessens et al (2011)	Naessens et al (2011)
Kenneley et al (2013) (IHI GTT)	Lipczak et al (2011b)			Lander et al (2010)	Lander et al (2010)
Rajeshet al 2012 (2011)	Marini et al (2012)			Marini et al (2012)	Marini et al (2012)
Naessens et al (2011)	**Outpatients**			**Paediatric**	**Paediatric**
Paediatric	Rosen et al (2010)			Matlow et al (2011)	Matlow et al (2011)
Matlow et al (2011) (IHI GTT)					
Matlow et al (2012) (IHI GTT)					
ICU patients					
Agarwal et al (2010)					
Larsen et al (2007)					
Pravinkumar et al (2009) (IHI GTT)					

Outpatient and primary care setting

The use of trigger tools in the outpatient or primary care setting has mainly been used in order to study ADEs (Singh et al 2009; Gurwitz et al 2003; Brenner et al 2012; Hope et al 2003). However, one Scottish study has examined the frequency of AEs by means of an adapted version of the IHI GTT (De Wet and Bowie 2009). The authors concluded that the trigger tool was able to successfully identify otherwise undetected AEs in primary care but raised concerns about the feasibility of the methodology due to its resource requirements.

Likewise, in relation to ADEs, Singh et al (2009) have also concluded that trigger tools have an important role in primary care in relation to quality improvement but suggested that a shorter version of the tool may be needed as it is less resource-intensive. By contrast, Brenner et al (2012) highlighted the shortcomings of an abbreviated trigger tool consisting of just six abnormal laboratory values and concluded that more complex tools were required to effectively identify ADEs in the outpatient setting. Finally, Rosen et al (2010) have suggested that triggers may serve a useful role in the identification of AEs specifically related to ambulatory surgical practice.

Table 14: Use of trigger tools in the outpatient setting

Outpatient assessments of AE rates	Outpatient assessments of ADE/ADR rates	Outpatient-based assessments of the accuracy of trigger tools to identify ADEs/ADRs	Outpatient-based comparisons of trigger tools with other methods to detect ADEs/ADRs
De Wet and Bowie (2009)	Singh et al (2009)	Rosen et al (2010)	Gurwitz et al (2003)
	Gurwitz et al (2003)	Brenner et al (2012)	Hope et al (2003)
	Brenner et al (2012)	Singh et al (2009)	

Assessments of the costs and cost-effectiveness of the use of trigger tools to identify harm

Although a number of authors have commented on the resource requirements associated with measuring harm either by means of trigger tools or with other methods, only two studies (Cohen et al 2005; Dormann et al 2000) have considered the costs associated with the introduction of trigger tools and mapped whether any savings occurred as a result of this intervention. The study by Cohen et al (2005) is an important example as it charted the costs associated with the introduction of a patient safety programme that included the provision of the IHI GTT at a community hospital in the United States. The researchers observed that both the frequency and severity of ADEs significantly declined after the programme was commenced and cost savings of over US$10 million were noted. The other study to measure costs associated with the provision of a computerised trigger tool on one ward at a German hospital to locate ADEs concluded that the potential for savings can be estimated at EUR 56,200/year. A study of 33 ambulatory practices in Indiana (Hope et al 2003) compared the cost per ADE identified for intensive pharmacist review with that of a tiered approach that included the IHI methodology. The tiered IHI approach was found to be more cost-effective than pharmacist review (US$68.7 per ADE identified vs US$42.4).

Table 15: Assessments of the costs and cost-effectiveness of the use of trigger tools

Assessments of costs before and after application of trigger tools and other interventions to improve patient safety	Comparisons of cost-effectiveness of trigger tools versus other methods to monitor harm
Cohen et al (2005)	Hope et al (2003)
Dormann et al (2000)	

Application of trigger tools in the New Zealand setting

Two published studies have assessed the use trigger tools in New Zealand (Seddon et al 2013; Schade et al 2006). Both studies focused on the use of trigger tools to identify ADEs. One of them used the IHI GTT and observed that a high rate of ADEs occurred at New Zealand hospitals (28.9 ADEs per 100 admissions) (Seddon et al 2013). Both noted that morphine and anticoagulants were commonly associated with ADEs. Both also compared the use of trigger tools with voluntary reporting to ascertain the frequency of ADEs among inpatients. Trigger tools in both studies consistently identified far more occurrences of ADEs than voluntary reporting. The study by Seddon et al (2013) documented 128 ADEs but noted that not even a single event had been voluntarily reported by any health professional.

Table 16: Application of trigger tools in the New Zealand setting

Use of trigger tools to describe rate of ADEs in New Zealand	Comparisons of trigger tools with other methods to detect harm in the New Zealand setting
Zolezzi et al (2007)	Zolezzi et al (2007)
Seddon et al (2013)	Seddon et al (2013)

Excluded studies

Table 17: Details of excluded studies

Author	Reason for exclusion
Fairclough et al (2009)	Not assessing AEs
Tinoco et al (2011)	Triggers vs ADEs with ADEs determined by text mining
Heenan (2009)	Not assessing trigger tools
Klopotowska et al (2011)	Study protocol only
Wolff and Bourke (2002)	General outcome based ‘triggers’ only (death, transfer, readmission)
Hogan et al (2008)	Short case note review but no clear use of triggers
Olsen et al (2007)	Short case note review but no clear use of triggers
Woloshynowych et al (2003)	Short case note review but no clear use of triggers
Grasela et al (1993)	Not assessing triggers
O’Neil et al (1993)	Assessing structured case note review and not clearly assessing triggers
Sari et al (2007)	Assessing structured case note review and not clearly assessing triggers
Alonzo (2010)	Protocol only
Anonymous (2009)	No description of methods etc

Author	Reason for exclusion
Dolores (et al (2010)	Spanish text
Meyer-Massetti and Cohen (2012)	German text
McKinney (2010)	Not assessing trigger tools
Mull and Nebeker (2008)	Unable to access full text of conference abstract
Moore and Childs (2011–12)	Unable to access full text of opinion article
Najjar et al (2012)	Unable to access full text of conference abstract
Paruthi et al (2011)	Unable to access full text
Robinson et al (2012)	Different type of trigger tool – to identify patients with end-stage heart failure
Vozikis et al (2012)	Greek text
Tomlin (et al (2012)	Natural language searching but no trigger evaluation
Anonymous (2008)	Danish text
Govindan et al (2010)	Limited to automatic detection only
Singh et al (2012)	Limited to automatic detection only
Trillo-Alvarez et al (2010)	Unable to access full text of conference abstract
Vangekrantz and Hvarfner (2009)	Unable to access full text of conference abstract
O'Leary et al (2013)	Text mining vs triggers
Berry et al (1988)	Published 1988

References

Adler L, Denham C, McKeever M et al. 2008. Global trigger tool: implementation basics. *J Patient Saf* 4(4): 245–9.

Agarwal S, Classen D, Larsen G et al. 2010. Prevalence of adverse events in pediatric intensive care units in the United States. *Pediatric Critical Care Medicine* 11(5): 568–78. PubMed PMID: 2010552639. English.

Alonzo J. 2010. Design and implementation of a methodology to benchmark adverse drug event occurrence rates. *Journal of the American Pharmacists Association* 50 (2): 305. PubMed PMID: 70480765. English.

Anonymous. 2008. *Systematisk journalgennemgang med IHI's GLOBAL TRIGGER TOOL.* Southern Denmark: Center for Quality. URL: http://wwwcenterforkvalitet.dk/wm293462.

Anonymous. 2009. Premier/IHI algorithms automate process for identifying patient harm. *Healthcare Benchmarks & Quality Improvement* 16(9): 97–9. PubMed PMID: 19697805. English.

Anonymous. 2010a. *Evidence scan: Global trigger tools.* London: Health Foundation.

Anonymous. 2010b. *How to use trigger tools: Cymru NHS Wales* [March 2013]. URL: www.1000livesplus.wales.nhs.uk.

Anonymous. 2011. *Interreg4A Strengthening patient safety between Denmark and Germany European Union: Syddanmark-Schlesweig-Kern* [March 2013]. URL: http://www.patientsafety-interreg.com/home.html.

Asavaroengchai S, Sriratanaban J, Hiransuthikul N et al. 2009. Identifying adverse events in hospitalized patients using Global Trigger Tool in Thailand. *Asian Biomedicine* 3(5): 545–50. PubMed PMID: 2011121614. English.

Baker GR, Norton PG, Flintoft V et al. 2004. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ* 170(11): 1678–86. PubMed PMID: 15159366. Pubmed Central PMCID: PMC408508. English.

Berry LL, Segal R, Sherrin TP et al. 1988. Sensitivity and specificity of three methods of detecting adverse drug reactions. *Am J Hosp Pharm* 45(7): 1534–9. PubMed PMID: 3046347. English.

Beyea SC. 2005. Using trigger tools to enhance patient safety. *AORN journal* 82(1): 115–6. PubMed PMID: 16114610. English.

Brennan T, Leape L, Laird N et al. 1991. Incidence of adverse events and negligence in hospitalized patients – results of the Harvard Medical Practice Study I. *New England Journal of Medicine* 324: 370–6.

Brenner S, Detz A, Lopez A. 2012. Signal and noise: applying a laboratory trigger tool to identify adverse drug events among primary care patients. *BMJ Quality & Safety* 21(8): 670–5. PubMed PMID: 22626736. Pubmed Central PMCID: PMC3402751. English.

Classen D, Lloyd R, Provost L et al. 2008. Development and evaluation of the Institute for healthcare Improvement Global Trigger Tool. *J Patient Saf* 4(3): 169–77.

Classen DC, Pestotnik SL, Evans RS et al. 1991. Computerized surveillance of adverse drug events in hospital patients.[Erratum appears in *JAMA* 1992 Apr 8; 267(14): 1922]. *JAMA* 266(20): 2847–51. PubMed PMID: 1942452. English.

Classen DC, Pestotnik SL, Evans RS. 2005. Computerized surveillance of adverse drug events in hospital patients. 1991. *Quality and Safety in Health Care* 14(3): 221–5; discussion 5–6. PubMed PMID: 15933322. Pubmed Central PMCID: PMC1744018. English.

Classen DC, Resar R, Griffin F et al. 2011. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. *Health Affairs* 30(4): 581–9. PubMed PMID: 2011247029. English.

Cohen MM, Kimmel NL, Benage MK et al. 2005. Medication safety program reduces adverse drug events in a community hospital. *Quality and Safety in Health Care* 14(3): 169–74. PubMed PMID: 2005271082. English.

Danish Safer Hospital Programme. 2012. *The Danish Safer Hospital Programme* [March 2013]. URL: www.sikkerpatient.dk.

Davis P, Lay-Yee R, Briant R et al. 2002. Adverse events in New Zealand public hospitals I: occurrence and impact. *NZ Med J* 115(1167): U271. PubMed PMID: 12552260. English.

de Boer M, Ramrattan MA, Kiewiet JJ et al. 2011. Cost-effectiveness of ward-based pharmacy care in surgical patients: protocol of the SUREPILL (Surgery & Pharmacy In Liaison) study. *BMC health services research* 11: 55. PubMed PMID: 21385352. English.

De Wet C, Bowie P. 2009. The preliminary development and testing of a global trigger tool to detect error and patient harm in primary-care records. *Postgraduate Medical Journal* 85(1002): 176–80. PubMed PMID: 2009236311. English.

De Wet C, Bowie P. 2011. Screening electronic patient records to detect preventable harm: a trigger tool for primary care. *Quality in Primary Care* 19(2): 115–25. PubMed PMID: 21575334. English.

Dolores Menendez M, Rancano I, Garcia V et al. 2010. Use of different patient safety reporting systems: Much ado about nothing?. [Spanish]Uso de diferentes sistemas de notificacin de eventos adversos: ?mucho ruido y pocas nueces? *Revista de Calidad Asistencial* 25(4): 232–6. PubMed PMID: 20494603. Spanish.

Dormann H, Criegee-Rieck M, Neubert A et al. 2004. Implementation of a computer-assisted monitoring system for the detection of adverse drug reactions in gastroenterology. *Aliment Pharmacol Ther* 19(3): 303–9. PubMed PMID: 14984377. English.

Dormann H, Muth-Selbach U, Krebs S et al. 2000. Incidence and costs of adverse drug reactions during hospitalisation: computerised monitoring versus stimulated spontaneous reporting. *Drug Saf* 22(2): 161–8. PubMed PMID: 10672897. English.

Doupi P. 2012. Using EHR data for monitoring and promoting patient safety: reviewing the evidence on trigger tools. *Stud Health Technol Inform* 180: 786–90. PubMed PMID: 22874299. English.

Egger T, Dormann H, Ahne G et al. 2003. Identification of adverse drug reactions in geriatric inpatients using a computerised drug database. *Drugs Aging* 20(10): 769–76. PubMed PMID: 12875612. English.

Fairclough E, Cairns E, Hamilton J et al. 2009. Evaluation of a modified early warning system for acute medical admissions and comparison with C-reactive protein/albumin ratio as a predictor of patient outcome. *Clinical Medicine, Journal of the Royal College of Physicians of London* 9(1): 30–3. PubMed PMID: 2009154171. English.

Fayed A, El-Hosseiny T, Shehata A et al. 2009. Using trigger tools to measure adverse drug events in an intensive care unit. *Critical Care Medicine* 37 (12 SUPPL.): A416. PubMed PMID: 70192224. English.

Ferranti J, Horvath MM, Cozart H et al. 2008. Reevaluating the safety profile of pediatrics: a comparison of computerized adverse drug event surveillance and voluntary reporting in the pediatric environment. *Pediatrics* 121(5): e1201–7. PubMed PMID: 18450863. English.

Franklin BD, Birch S, Savage I et al. 2009. Methodological variability in detecting prescribing errors and consequences for the evaluation of interventions. *Pharmacoepidemiology and Drug Safety* 18(11): 992–9. PubMed PMID: 2009611658. English.

Franklin BD, Birch S, Schachter M et al. 2010. Testing a trigger tool as a method of detecting harm from medication errors in a UK hospital: A pilot study. *International Journal of Pharmacy Practice* 18(5): 305–11. PubMed PMID: 2010488013. English.

Good VS, Saldana M, Gilder R et al. 2011. 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. *BMJ Quality and Safety* (1): 25–30. PubMed PMID: 2011316480. English.

Govindan M, Van Citters AD, Nelson EC et al. 2010. Automated detection of harm in healthcare with information technology: a systematic review. *Quality and Safety in Health Care* 19(5): e11. PubMed PMID: 20671081. English.

Grasela TH, Walawander CA, Kennedy DL et al. 1993. Capability of hospital computer systems in performing drug-use evaluations and adverse drug event monitoring. *Am J Hosp Pharm* 50(9): 1889–95. PubMed PMID: 8135235. English.

Griffin F, Resar R. 2009. *IHI Global Trigger Tool for measuring adverse events*. IHI Innovation Series [Internet]. 2nd edition.

Griffin FA, Classen DC. 2008. Detection of adverse events in surgical patients using the Trigger Tool approach. *Quality and Safety in Health Care* 17(4): 253–8. PubMed PMID: 18678721. English.

Gurwitz JH, Field TS, Harrold LR et al. 2003. Incidence and preventability of adverse drug events among older persons in the ambulatory setting. *JAMA* 289(9): 1107–16. PubMed PMID: 12622580. English.

Haffner S, von Laue N, Wirth S et al. 2005. Detecting adverse drug reactions on paediatric wards: intensified surveillance versus computerised screening of laboratory values. *Drug Saf* 28(5): 453–64. PubMed PMID: 15853446. English.

Handler SM, Hanlon JT. 2010. Detecting adverse drug events using a nursing home-specific trigger tool. *Annals of Long-Term Care* 18(5): 17–22. PubMed PMID: 2010381450. English.

Heenan M. 2009. Engaging physicians in and measuring medical quality: Development of a medical quality scorecard. *Quality and Safety in Health Care* 18 (4): e1. PubMed PMID: 70007761. English.

Hogan H, Olsen S, Scobie S et al. 2008. What can we learn about patient safety from information sources within an acute hospital: a step on the ladder of integrated risk management? *Quality and Safety in Health Care* 17(3): 209–15. PubMed PMID: 18519628. English.

Hope C, Overhage JM, Seger A et al. 2003. A tiered approach is more cost effective than traditional pharmacist-based review for classifying computer-detected signals as adverse drug events. *J Biomed Inform* 36(1–2): 92–8. PubMed PMID: 14552850. English.

Hougland P, Xu W, Pickard S et al. 2006. Performance of International Classification of Diseases, 9th Revision, Clinical Modification codes as an adverse drug event surveillance system. *Med Care* 44(7): 629–36. PubMed PMID: 16799357. English.

Huddleston J, Gabriel S, Fowler J. 2011. Reliability of safe hospital care for congestive heart failure patients. *Clinical and Translational Science* 4(2): 110. PubMed PMID: 70587867. English.

Jha AK, Kuperman GJ, Teich JM et al. 1998. Identifying adverse drug events: development of a computer-based monitor and comparison with chart review and stimulated voluntary report. *J Am Med Inform Assoc* 5(3): 305–14. PubMed PMID: 9609500. Pubmed Central PMCID: PMC61304. English.

Jick H. 1974. Drugs – remarkably toxic. *N Engl Med J* 291: 824–8.

Kaafarani HM, Rosen AK, Nebeker JR et al. 2010. Development of trigger tools for surveillance of adverse events in ambulatory surgery. *Quality and Safety in Health Care* 19(5): 425–9. PubMed PMID: 20513790. English.

Kandpal S, Bulgar A, McHugh R. 2012. IHI global trigger tools and adverse events. *Journal of Hospital Medicine* 7: S21. PubMed PMID: 70698024. English.

Kenneley DA, Saldana M, Kudyakov R et al. 2013. Description and evaluation of adaptations to the Global Trigger Tool to enhance value to adverse event reduction efforts. *Journal of Patient Safety* 9(2): 00–.

Kilbridge PM, Campbell UC, Cozart HB et al. 2006. Automated surveillance for adverse drug events at a community hospital and an academic medical center. *J Am Med Inform Assoc* 13(4): 372–7. PubMed PMID: 16622159. Pubmed Central PMCID: PMC1513675. English.

Kirkendall ES, Kloppenborg E, Papp J et al. 2012. Measuring adverse events and levels of harm in pediatric inpatients with the global trigger tool. *Pediatrics* 130(5): e1206–e14. PubMed PMID: 2012656833. English.

Klopotowska JE, Wierenga PC, de Rooij SE et al. 2011. The effect of an active on-ward participation of hospital pharmacists in Internal Medicine teams on preventable Adverse Drug Events in elderly inpatients: protocol of the WINGS study (Ward-oriented pharmacy in newly admitted geriatric seniors). *BMC health services research* 11: 124. PubMed PMID: 21612624. English.

Lander L, Roberson DW, Plummer KM. 2010. A trigger tool fails to identify serious errors and adverse events in pediatric otolaryngology. *Otolaryngology – Head and Neck Surgery* 143(4): 480–6. PubMed PMID: 2010541994. English.

Landrigan CP, Parry GJ, Bones CB et al. 2010. Temporal trends in rates of patient harm resulting from medical care. *New England Journal of Medicine* 363(22): 2124–34. PubMed PMID: 2010649530. English.

Larsen GY, Donaldson AE, Parker HB et al. 2008. Preventable harm occurring to critically ill children. *Pediatric Critical Care Medicine* 8(4): 331–6. PubMed PMID: 2007340419. English.

Lau H, Litman KC. 2011. Saving lives by studying deaths: using standardized mortality reviews to improve inpatient safety. J*oint Commission Journal on Quality & Patient Safety* 37(9): 400–8. PubMed PMID: 21995256. English.

Leape L. 2007. Is hospital patient care becoming safer? A conversation with Lucian Leape. Interview by Peter I. Buerhaus.[Erratum appears in *Health Aff* (Millwood). 2007 Nov–Dec; 26(6):following w696]. *Health Affairs* 26(6): Nov–Dec. PubMed PMID: 17925308. English.

Lemon V, Stockwell DC. 2012. Automated Detection of Adverse Events in Children. *Pediatric Clinics of North America* 59(6): 1269–78. PubMed PMID: 2012640005. English.

Lessing C, Schmitz A, Albers B et al. 2010. Impact of sample size on variation of adverse events and preventable adverse events: systematic review on epidemiology and contributing factors. *Quality and Safety in Health Care* 19(6): e24. PubMed PMID: 20679137. Pubmed Central PMCID: PMC3002821. English.

Levinson D. 2010. *Adverse events in hospitals: methods for identifying events.* Washington: Department of Health and Human Services, Office of Inspector General.

Levy M, Azaz-Livshits T, Sadan B. 1999. Computerized surveillance of adverse drug reactions in hospital: implementation. *Eur J Clin Pharmacol* 54(11): 887–92. PubMed PMID: 10027665. English.

Lipczak H, Knudsen JL, Nissen A. 2011a. Safety hazards in cancer care: Findings using three different methods. *BMJ Quality and Safety* 20(12): 1052–6. PubMed PMID: 2011663047. English.

Lipczak H, Neckelmann K, Steding-Jessen M et al. 2011b. Uncertain added value of Global Trigger Tool for monitoring of patient safety in cancer care. *Danish Medical Bulletin* 58(11). PubMed PMID: 2011629462. English.

Louie K, Wilmer A, Wong H et al. 2010. Medication error reporting systems: A survey of Canadian intensive care units. *Canadian Journal of Hospital Pharmacy* 63(1): 20–4. PubMed PMID: 2011699385. English.

Mack EH, Brilli RJ. 2007. To err is human; to improve, divine. *Pediatric Critical Care Medicine* 8(4): 398–9. PubMed PMID: 2007340433. English.

Mangoni AA. 2012. Predicting and detecting adverse drug reactions in old age: Challenges and opportunities. Expert Opinion on Drug Metabolism and Toxicology 8(5): 527–30. PubMed PMID: 2012227219. English.

Marini H, Merle V, Derrey S et al. 2012. Surveillance of unplanned return to the operating theatre in neurosurgery combined with a mortality-morbidity conference: Results of a pilot survey. *BMJ Quality and Safety* 21(5): 432–8. PubMed PMID: 2012233335. English.

Matlow A, Flintoft V, Orrbine E et al. 2005. The development of the Canadian paediatric trigger tool for identifying potential adverse events. *Healthcare Quarterly* (Toronto, Ont) 8 Spec No: 90–3. PubMed PMID: 16334079. English.

Matlow AG, Baker GR, Flintoft V et al. 2012. Adverse events among children in Canadian hospitals: The Canadian Paediatric Adverse Events Study. *CMAJ* 184(13): E709–E18. PubMed PMID: 2012562074. English.

Matlow AG, Cronin CMG, Flintoft V et al. 2011. Description of the development and validation of the Canadian paediatric trigger tool. *BMJ Quality and Safety* 20(5): 416–23. PubMed PMID: 2011306173. English.

McKinney S. 2010. Venous thromboembolism (VTE) prophylaxis in hospital patients. Canadian *Journal of Hospital Pharmacy* 63(1): 79. PubMed PMID: 70657310. English.

MERP N. 2013. *National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) Index for Categorizing Errors 2001* [February 2013]. URL: http://www.nccmerp.org/medErrorCatIndex.html.

Meyer-Massetti C, Cheng CM, Schwappach DLB et al. 2011. Systematic review of medication safety assessment methods. *American Journal of Health-System Pharmacy* 68(3): 227–40. PubMed PMID: 2011269551. English.

Meyer-Massetti C, Conen D. 2012. Assessment, frequency, causes, and prevention of medication errors – acritical analysisErfassung, haufigkeit, ursachen und pravention von medikationsfehlern – eine kritische analyse. *Therapeutische Umschau* 69(6): 347–52. PubMed PMID: 2012317497. English, German.

Moore C, Childs L. 2011–12. A tool to identify falling care quality. *Nursing Times* 107(49–50): 14–6. PubMed PMID: 22239045. English.

Muething SE, Conway PH, Kloppenborg E et al. 2010. Identifying causes of adverse events detected by an automated trigger tool through in-depth analysis. *Quality and Safety in Health Care* 19(5): 435–9. PubMed PMID: 20798069. English.

Mull H, Shimada S, Nebeker J et al. 2008. *Review of the trigger literature: Adverse events targeted and gaps in detection. Triggers and targeted injury detection systems expert panel meeting: conference summary February 2009.* Rockville, MD: AHRQ.

Mull HJ, Nebeker JR, Shimada SL et al. 2011. Consensus building for development of outpatient adverse drug event triggers. *Journal of Patient Safety* 7(2): 66–71. PubMed PMID: 21587117. English.

Mull HJ, Nebeker JR. 2008. *Informatics tools for the development of action-oriented triggers for outpatient adverse drug events.* AMIA; Annual Symposium proceedings/AMIA Symposium. AMIA Symposium: 505–9. PubMed PMID: 18999297. English.

Naessens JM, Campbell CR, Huddleston JM et al. 2009. A comparison of hospital adverse events identified by three widely used detection methods. *International Journal for Quality in Health Care* 21(4): 301–7. PubMed PMID: 2009396775. English.

Naessens JM, O'Byrne TJ, Johnson MG. 2011. Measuring hospital adverse events: Assessing inter-rater reliability and trigger performance of the Global Trigger Tool. *International Journal for Quality in Health Care* 22(4): 266–74. PubMed PMID: 2010414201. English.

Najjar R, Tuffaha H, Abdel-Rahman F et al. 2012. Adverse drug reactions in bone marrow transplant patients. *Bone Marrow Transplantation* 47: S204–S5. PubMed PMID: 70722986. English.

Nebeker J, Stoddart G, Rosen AK. 2008. *Considering sensitivity and positive predictive value in comparing the perforamnce of triggers systems for iatrogenic adverse events. Triggers and targeted injury detection systems expert panel meeting: conference summary.* Rockville, MD: AHRQ.

Neubert A, Dormann H, Weiss J et al. 2006. Are computerised monitoring systems of value to improve pharmacovigilance in paediatric patients? *Eur J Clin Pharmacol* 62(11): 959–65. PubMed PMID: 17021890. English.

Nicol N. 2007. Case study: An interdisciplinary approach to medication error reduction. *American Journal of Health-System Pharmacy* 64(14 SUPPL.): S17–S20. PubMed PMID: 2007350204. English.

Nilsson L, Pihl A, Tagsjo M et al. 2012. Adverse events are common on the intensive care unit: Results from a structured record review. *Acta Anaesthesiologica Scandinavica* 56(8): 959–65. PubMed PMID: 2012485190. English.

O'Leary KJ, Devisetty VK, Patel AR et al. 2013. Comparison of traditional trigger tool to data warehouse based screening for identifying hospital adverse events. *BMJ Quality and Safety* 22(2): 130–8. PubMed PMID: 2013146628. English.

Olsen S, Neale G, Schwab K et al. 2007. Hospital staff should use more than one method to detect adverse events and potential adverse events: incident reporting, pharmacist surveillance and local real-time record review may all have a place. *Quality and Safety in Health Care* 16(1): 40–4. PubMed PMID: 17301203. Pubmed Central PMCID: PMC2464933. English.

O'Neil AC, Petersen LA, Cook EF. 1993. Physician reporting compared with medical-record review to identify adverse medical events. *Ann Intern Med* 119(5): 370–6. PubMed PMID: 8338290. English.

Parry G, Cline A, Goldmann D. 2012. Deciphering harm measurement. *JAMA* 307(20): 2155–6. PubMed PMID: 22618920. English.

Paruthi P, Pansare G, Khairnar A. 2011. Use of triggers to detect adverse drug reaction induced by cardiovascular drugs in outpatient department in Nasik City. *Research Journal of Pharmacy and Technology* 4(12): 1819–21. PubMed PMID: 2012245774. English.

Pravinkumar SE, Warren ML, Bruno JJ et al. 2009. Implementation of the Institute for Healthcare Improvement Global Trigger Tool in an oncological ICU: Pilot data. Chest Conference: American College of Chest Physicians Annual Meeting, CHEST 136(4). PubMed PMID: 70203203. English.

Rajesh V, Surulivel Rajan M, Vijayanarayana K et al. 2012. Development of trigger tool for identifying adverse events in surgery department of an Indian tertiary care teaching hospital. *International Journal of Pharmacy Practice* 20: 65–6. PubMed PMID: 70879066. English.

Resar R. 2009. *Reflections on the Institute for Healthcare Improvement Global Trigger Tool. Triggers and targeted injury detection systems expert panel meeting: conference summary*. Rockville, MD: AHRQ.

Resar RK, Rozich JD, Classen D. 2003. Methodology and rationale for the measurement of harm with trigger tools. *Quality and Safety in Health Care* 12(SUPPL. 2): ii39–ii45. PubMed PMID: 2003514270. English.

Resar RK, Rozich JD, Simmonds T et al. 2006. A trigger tool to identify adverse events in the intensive care unit. *Joint Commission journal on quality and patient safety/Joint Commission Resources*. 32(10): 585–90. PubMed PMID: 17066996. English.

Robinson KS, Stephens C, Harris DT et al. 2012. Data collection process to reduce heart failure readmissions. Heart and Lung: *Journal of Acute and Critical Care* 41: 431–2. PubMed PMID: 70865012. English.

Rosen AK, Kaafarani H, Mull H et al. 2010. Use of trigger tools to detect adverse events in ambulatory surgery. *Journal of General Internal Medicine* 25: S423–S4. PubMed PMID: 70315257. English.

Rozich JD, Haraden CR, Resar RK. 2003. Adverse drug event trigger tool: A practical methodology for measuring medication related harm. *Quality and Safety in Health Care* 12(3): 194–200. PubMed PMID: 2003242860. English.

Sari AB, Sheldon TA, Cracknell A et al. 2007. Sensitivity of routine system for reporting patient safety incidents in an NHS hospital: retrospective patient case note review. *BMJ* 334(7584): 79. PubMed PMID: 17175566. Pubmed Central PMCID: PMC1767248. English.

Schade CP, Hannah K, Ruddick P et al. 2006. Improving self-reporting of adverse drug events in a West Virginia hospital. *Am J Med Qual* 21(5): 335–41. PubMed PMID: 16973950. English.

Schildmeijer K, Nilsson L, Arestedt K et al. 2012. Assessment of adverse events in medical care: Lack of consistency between experienced teams using the global trigger tool. *BMJ Quality and Safety* 21(4): 307–14. PubMed PMID: 2012170320. English.

Seddon M, Jackson A, Cameron C et al. 2013. The Adverse Drug Event Collaborative: a joint venture to measure medication-related patient harm. *New Zealand Medical Journal* 126(1368).

Seynaeve S, Reyntiens D, Vandenplas D et al. 2010. Adverse drug events in the critical care unit. *Critical Care* 14: S150. PubMed PMID: 70255691. English.

Seynaeve S, Verbrugghe W, Claes B. 2011. Adverse drug events in intensive care units: A cross-sectional study of prevalence and risk factors. *American Journal of Critical Care* 20(6): e131–e40. PubMed PMID: 22045149. English.

Sharek P. 2009. *The North Carolina Harm Study: Validating the IHI Global Trigger Tool (GTT) as a potential national harm measure.* Stanford University. URL: http://spice.stanford.edu/events/the_north_carolina_harm_study_validating_the_ihi_global_trigger_tool_gtt_as_a_potential_national_harm_measure/.

Sharek PJ, Classen D. 2006. The Incidence of Adverse Events and Medical Error in Pediatrics. *Pediatric Clinics of North America* 53(6): 1067–77. PubMed PMID: 2006570502. English.

Sharek PJ, Horbar JD, Mason W et al. 2006. 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* 118(4): 1332–40. PubMed PMID: 17015521. English.

Sharek PJ, Parry G, Goldmann D et al. 2011. Performance characteristics of a methodology to quantify adverse events over time in hospitalized patients. *Health Services Research* 46(2): 654–78. PubMed PMID: 2011136474. English.

Shimada S, Rivard PE, Mull H et al. 2008. *Triggers and targeted injury detection systems: Aiming for the right target with the appropriate tool. Triggers and targeted injury detection systems expert panel meeting: conference summary*. Rockville, MD: AHRQ.

Singh R, Anderson D, McLean-Plunkett E et al. 2012. IT-enabled systems engineering approach to monitoring and reducing ADEs. *American Journal of Managed Care* 18(3): 169–75. PubMed PMID: 2012263019. English.

Singh R, McLean-Plunckett EA, Kee R et al. 2009. Experience with a trigger tool for identifying adverse drug events among older adults in ambulatory primary care. *Quality and Safety in Health Care* 18(3): 199–204. PubMed PMID: 2009269571. English.

Stockwell DC. 2010. Pulling the triggers on adverse events in the pediatric intensive are unit. *Pediatric Critical Care Medicine* 11(5): 632–3. PubMed PMID: 2010552651. English.

Suresh GK. 2012. Measuring patient safety in neonatology. *American Journal of Perinatology* 29(1): 19–26. PubMed PMID: 2012057590. English.

Szekendi MK, Sullivan C, Bobb A et al. 2006. Active surveillance using electronic triggers to detect adverse events in hospitalized patients. *Quality and Safety in Health Care* 15(3): 184–90. PubMed PMID: 2006291938. English.

Takata GS, Mason W, Taketomo C. 2008a. Development, testing, and findings of a pediatric-focused trigger tool to identify medication-related harm in US children's hospitals. *Pediatrics* 121(4): e927–35. PubMed PMID: 18381521. English.

Takata GS, Taketomo CK, Waite S. 2008b. Characteristics of medication errors and adverse drug events in hospitals participating in the California Pediatric Patient Safety Initiative. *American Journal of Health-System Pharmacy* 65(21): 2036–44. PubMed PMID: 2008532727. English.

Tegeder I, Levy M, Muth-Selbach U et al. 1999. Retrospective analysis of the frequency and recognition of adverse drug reactions by means of automatically recorded laboratory signals. *Br J Clin Pharmacol* 47(5): 557–64. PubMed PMID: 10336581. Pubmed Central PMCID: PMC2014178. English.

Thomas E, Petersen L. 2003. Measuring errors and adverse events in health care. *J Gen Intern Med* 18: 61–7.

Thomas EJ, Studdert DM, Burstin HR et al. 2000. Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care* 38(3): 261–71. PubMed PMID: 10718351. English.

Thuermann PA, Windecker R, Steffen J et al. 2002. Detection of adverse drug reactions in a neurological department: comparison between intensified surveillance and a computer-assisted approach. *Drug Saf* 25(10): 713–24. PubMed PMID: 12167067. English.

Tinoco A, Evans RS, Staes CJ et al. 2011. Comparison of computerized surveillance and manual chart review for adverse events. *J Am Med Inform Assoc* 18(4): 491–7. PubMed PMID: 21672911. Pubmed Central PMCID: PMC3128408. English.

Tomlin A, Reith D, Dovey S et al. 2012. Methods for retrospective detection of drug safety signals and adverse events in electronic general practice records. *Drug Saf* 35(9): 733–43. PubMed PMID: 22861670. English.

Trillo-Alvarez C, Hill S, Leach L et al. 2010. Effect of a nurse lead rapid response team on out of ICU CODE blue survival: A county hospital experience with two novel interventions. American Journal of Respiratory and Critical Care Medicine Conference: American Thoracic Society International Conference, ATS. 181(1 MeetingAbstracts). PubMed PMID: 70844442. English.

Vangekrantz S, Hvarfner A. 2009. Adverse events, suboptimal assessments of vital signs and descisions to forgo treatment prior to ICU admission from the general wards among patients who dies within 30 days of ICU admission. *Intensive Care Medicine* 35: S101. PubMed PMID: 70190601. English.

Vincent C, Neale G, Woloshynowych M. 2001. Adverse events in British hospitals: preliminary retrospective record review.[Erratum appears in *BMJ* 2001 Jun 9;322(7299): 1395]. *BMJ* 322(7285): 517–9. PubMed PMID: 11230064. Pubmed Central PMCID: PMC26554. English.

Vincent C. 2010. *Patient safety*. London: BMJ Publishing Group Limited.

Von Plessen C, Kodal AM, Anhoj J. 2012. Experiences with global trigger tool reviews in five Danish hospitals: An implementation study. *BMJ Open* 2(5). PubMed PMID: 2012683515. English.

Vozikis A, Pollalis Y, Riga M et al. 2012. A system for the detection, recording and analysis of adverse events: Implementation in intensive care units (ICU–MERIS). *Archives of Hellenic Medicine* 29(3): 345–53. PubMed PMID: 2012355776. English.

Wilson RM, Runciman WB, Gibberd RW et al. 1996. Quality in Australian Health Care Study. *Med J Aust* 164(12): 754. PubMed PMID: 8668084. English.

Wolff AM, Bourke J. 2002. Detecting and reducing adverse events in an Australian rural base hospital emergency department using medical record screening and review. *Emerg Med J* 19(1): 35–40. PubMed PMID: 11777869. Pubmed Central PMCID: PMC1725773. English.

Woloshynowych M, Neale G, Vincent C. 2003. Case record review of adverse events: a new approach. *Quality and Safety in Health Care* 12(6): 411–5. PubMed PMID: 14645755. Pubmed Central PMCID: PMC1758034. English.

Yeesoonpan N, Tragulpiankit P, Ningsananda T. 2011a. The pilot study of evaluation of IHI ADE trigger tool in thai hospitalized patients: Multicenter study. *Pharmacoepidemiology and Drug Safety* 20: S349. PubMed PMID: 70665638. English.

Yeesoonpan N, Wisetsuwonaphum K, Tragulpiankit P. 2011b. The sensitivity and specificity of ADE detection by trigger tool in Thai pediatric hospitalized patients. *Pharmacoepidemiology and Drug Safety* 20: S349–S50. PubMed PMID: 70665639. English.

Zimmerman R, Pierson S, McLean R et al. 2010. Aiming for zero preventable deaths: using death review to improve care and reduce harm. *Healthcare Quarterly* (Toronto, Ont). 13 Spec No: 81–7. PubMed PMID: 20959735. English.

Zolezzi M, Forbes A, Parsotam N et al. 2007. Investigation of trigger tools for detecting adverse drug reactions. *Journal of Pharmacy Practice and Research* 37(3): 225–7. PubMed PMID: 2007507713. English.

Appendix One: Descriptions of included studies

Table 18: Descriptions of included studies

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Agarwal et al (2010) PICU AE	15 US paediatric ICUs	22 trigger tools developed by 8 physicians based on 32 common AEs. Training process with standard charts and webcasts and instruction manual and data collection sheets. Randomised review of 734 patient records staying > 2 days in PICU in 2005.	62% of PICU patients had a least 1 AE. 1488 AEs were identified including 256 ADEs, 28.6 AEs and 4.9 ADEs per 100 patient days. The most common AEs were catheter complications, uncontrolled pain, and endotracheal tube malposition. 10% of AEs were life threatening or permanent, 45% were preventable. Higher rates of AEs were associated with surgical patients, those intubated or those or died. The cumulative risk of an AE per PICU stay was 5.3%.	AEs and ADEs occur frequently in the PICU.
Asavaroengchai et al (2009) AE reliability	576 randomly sampled records were reviewed with 4460 patient days for patients at King Chulalongkorn Hospital, Bangkok in 2008	The GTT was compared with retrospective record review by trained nurses and physician.	Among the records 776 triggers were recorded (1.35 per patient). Inter-rater reliability for the triggers was high (kappa = 0.86). 138 records were identified with AEs (24%, 20.5%–27.5%). 236 AEs were identified. 41 AEs per 100 patients (32.3–49.6) or 50.4 events per 1000 patient days (40.7–60). 9 were judged severe (level G, H or I). 57.6% were preventable. 75 AEs were related to patient care processes, 48 were in surgery and 42 were related to medication.	The GTT detects more AEs than previously noted but most events are low severity. No gold standard was used to determine AEs.
Brenner et al (2012) ADE Selected triggers Accuracy Outpatient	Outpatient clinic at San Francisco, November 2008 to November 2009	6 abnormal laboratory values were used as triggers to search a clinical/administrative database. Trigger positive charts were reviewed by 2 physicians.	1342 triggers occurred and 622 ADEs among 516 patients. The trigger tool identified 91 ADEs (15% of all present). 49 (54%) of these ADEs occurred during medication monitoring and 41 (45%) during patient self-administration. 96% of INR abnormal values were ADEs but PPVs were 12% or less for the other triggers.	Other tools or more complex screening rules are needed to effectively screen for ADEs in sick adults in primary care.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Classen et al (1991; 2005) Computer screening ADE Voluntary reports	LDS Hospital, Salt Lake City, May 1989 to October 1990	Electronic drug monitoring included in an integrated hospital record system detected potential ADEs with algorithms (such as medication discontinuations or dose changes, antidotes, lab test abnormalities), which were checked by a pharmacist and an ADE was assigned if relevant using Naranjo criteria.	731 ADEs identified in 648 patients. 9 ADEs were voluntarily reported and 91 of the alerts. 100 of the ADEs were severe. Antidote use and therapeutic drugs for ADEs were most reliable signals.	Computer screening offers a potential method for improving the detection and characterisation of ADEs in hospitals.
Classen et al (2011) AE Comparisons Accuracy	3 large unnamed US hospitals with well-developed patient safety programmes. 1 academic and 2 community hospitals. Random selection of 795 patients in October 2004.	GTT and AHRQ indicators and incident reporting compared at 3 hospitals with full record review. 1 review team undertook IHI 2-stage and full record review processes at all hospitals.	393 AEs were detected. The GTT identified 354 (90%) of AEs, incident reporting identified 4 (1%) and the AHRQ indicators identified 35 (9%). AEs occurred in 33% of admissions or 91 events per 1000 patient days. Patients with an AE were older, had higher mortality and longer length of stay. GTT was associated with 95% sensitivity and 100% specificity. The indicators had sensitivity of 9% and specificity 99%. 26/354 AEs detected by the GTT were severe (life threatening, fatal or permanent injury).	Reliance on voluntary reporting or indicators may give misleading conclusions about safety in US hospitals and misdirect efforts to improve safety.
Cohen et al (2005) ADE GTT Intervention	Audit of ADEs at Missouri Baptist Medical Center from January 2001 to December 2003.	10–20 records reviewed each month using IHI protocol. Audit undertaken at baseline and after range of initiative to improve safety culture including provision of various medication protocols, new staff and safety council and new reporting opportunities.	Median ADEs per 1000 doses of medication declined from 2.04–0.65 ($p < 0.001$). Median ADEs per 100 inpatient days also reduced from 5.07 to 1.3 ($p < 0.001$). The percentage of inpatients with an ADE decreased from 31% to 10% ($p < 0.001$) The severity of ADEs declined. Cost savings of over US$10 million were noted.	A series of low-cost interventions focused on high-risk medications led to a significant decrease in harm.
De Wet and Bowie (2009) Outpatient	5 urban general practices in Scotland	IHI outpatient trigger tool developed for use with general practice by group of 20 general practitioners using Delphi technique. A 10-item trigger was tested with 100 randomly selected clinical records on electronic clinical database. Reviewers trained with IHI process.	730 triggers were records from 2251 consultations. Further review of triggers identified 47 episodes of patient harm (9.4%) and another 17 near miss episodes. Error/AEs occurred 1 per 35 consultations and harm 1 per 45 consultations. 2 events were associated with permanent harm but the events occurred in secondary care. Most AEs related to prescribing.	Trigger tool successful in identifying undetected patient harm primary care but feasibility remains unclear as it is time and labour intensive.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Dormann et al (2000) ADR Comparisons Automated trigger Accuracy	Single medical ward at German University Hospital in 1997	Computer-based monitoring of laboratory values outside of a defined range compared with stimulated spontaneous reporting where medical staff were asked 3 times a week about AEs. ADRs were classified by Navanjo algorithm.	501 computer alerts were generated and 34 ADRs whereas 17 ADRs were identified by spontaneous reporting. Only 5 ADRs were identified by both methods. Computer monitoring had 74% sensitivity and 75% specificity whereas spontaneous reporting had 37% sensitivity and 98% specificity. All 3 serious ADRs were noted by computer monitoring but 2 were reported. The PPV of the alerts was 13%. ADRs were associated with 3.5 days excess length of stay and savings from introducing monitoring were estimated to be EUR 56,200 per year.	Computer monitoring is an effective method for detecting ADRs. Large excess length of stay and costs from ADRs may be reduced by monitoring.
Dormann et al (2004) ADR Accuracy	Single gastroenterological ward at University Hospital, Erlangen-Nuremberg, Germany, Sept 2000 to March 2001	All charts were assessed daily by a physician and a pharmacist. A computerised monitoring system generated daily alerts for laboratory related data.	The computer monitoring system generated 2328 alerts of which 914 (39%) were related to 109 ADRs. Most alerts related to hepatotoxicity and coagulation disorders. Central nervous system agents were the most common drug class related to ADRs. The sensitivity of the ADRs was 91% and specificity improved from 23% to 76% by including trend monitoring with the computer program.	Computer monitoring is a useful tool for the detection of ADRs.
Egger et al (2003) ADR Comparison Geriatric	Geriatric rehabilitation ward at St Marien Hospital, Erlangen, Germany, October 2001 to February 2002	Daily review of charts by pharmacist and physician and computerised drug database review providing range of ADR alert types. ADRs categorised by Naranjo.	60.7% of 163 patients experienced at least 1 ADR. The database detected 309 potential ADRs but only 21 were of high frequency (> 1%). In 48% of ADR positive patients the database detected at least 1 ADR. In 14 of 24 drug-drug interaction cases the database provided an alert (sensitivity = 58%).	ADRs are common among geriatric patients. Computerised drug databases are useful for detecting ADRs but the software also provides a large number of false signals so needs refinement.
Fayed et al (2009) Abstract ICU AE	ICU at single hospital in Egypt	20 admissions per month reviewed by electronic screening using 16 triggers with review by a pharmacist.	Among the 240 records 139 triggers were noted in 66 records. 24 ADEs occurred among 21 patients (8.75% ADEs per 100 ICU admissions. 5% were serious severity and antimicrobials were the most commonly associated medication.	Trigger tools were effective in identifying medication related AEs during ICU stays.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Ferranti et al (2008) Electronic ADE Comparison Paediatrics	Duke University Hospital 2004–2006. Comparison of computerised trigger system and voluntary reporting	Computerised ADE surveillance using Duke University system involving 57 warnings about medication and laboratory triggers. Chart review then undertaken by pharmacist who also assign causality and severity scores.	849 voluntary reports gave 93 AEs. ADE rate was 1.8 (1.5–2.2) per 100 inpatient days. 1537 triggers were made and 78 ADEs were noted 1.6 (1.2–2.1) per 1000 inpatient days. There was little overlap between the events identified by different methods. Most reporting occurred in the ICU while triggers were spread across wards.	Multiple systems are needed to assess the epidemiology of ADEs. Voluntary reporting is good at identifying administration errors while surveillance was good at identifying problems with high-risk medications. Paediatric surveillance did worse than adult systems suggesting some tailoring was needed.
Franklin et al (2009) ADE Comparisons	93 patients at a 28-bed general surgery ward in a London teaching hospital	Prescribing errors were identified by a ward pharmacist, health record review, trigger tool, spontaneous reporting over 4 week-long periods before and after the introduction of computerised physician order entry.	Overall 135 prescribing errors were detected (10.7% of medication orders) pre computerised physician order entry (CPOE) and 127 post CPOE (7.9%)(relative risk reduction 26%). There was little overlap in the AEs identified by each method. Pharmacist detected 48 (36% of all PEs) pre and 30 (24%) post CPOE, record review identified 923 (69%) pre and 105 (83%) post CPOE, trigger tool 0 pre and 2% post (2%) and reporting 1 (1%) and 1 (1%) post.	Trigger tools were less useful for detecting events in this pilot study and authors concluded that a combination of methods was needed to assess the effectiveness of the intervention.
Franklin et al (2010) Comparison ADE UK	Single surgical ward at hospital in London in 2004 with 207 patients	US trigger tool adapted for UK use by changing units and some drugs. Full record review undertaken by research pharmacist on 207 records (69 patient records unavailable). Trigger tool then applied to paper records and positive triggers further assessed for ADEs by same pharmacist.	168 positive triggers identified in 127 patients. 7 ADEs were recognised (5 non-preventable). ADE rate = 3.4% of patients or 0.7 per 100 patient days. Preventable ADEs were 1% of patients or 0.2 per 100 patient days. Overall PPV = 0.04 and 0.01 for preventable ADEs. PPVs for individual triggers varied widely from 0–100%. 5 preventable ADEs were found by record review. Sensitivity of locating preventable ADEs was 0.4 compared with record review. Record review required on average 44 minutes and triggers 4 minutes.	Some ADEs were identified by trigger tool but more work is needed to reduce false positives and increase sensitivity. Retrospective health record review is still needed.
Good et al (2011) GTT example Enhanced AE	Application of GTT to 12 hospitals in Baylor Health Care System, Texas, US, June 2006 to July 2007.	GTT applied by professional nurse reviewer with additional information about the AEs in order to help characterise learning opportunities.	Among 2369 admissions reviewed there were 68.1 AEs per 1000 patient days, 50.8 AEs per 100 encounters and 39.8% of admissions had at least 1 AE. Most AEs were acquired as inpatients – 41.6 per 1000 patient days or 25% of admissions were inpatient related. Some 13.4% of AEs were permanent, required immediate life-saving help or were fatal.	The GTT can be refined to support learning opportunities and quality improvement activities.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Griffin and Classen (2008) Surgical AE	Initial pilot testing in 5 hospitals, then subsequent use of surgical GTT in 11 US hospitals, October 2003 to October 2004	Development of 23 surgical triggers using literature and expert group. Standard harm severity rating. Pilot in 5 hospitals with subsequent deletion of 1 trigger. Teams at hospitals included surgeons, nurses, anaesthetists, and quality improvement staff. Training was provided and standardisation given. Review of triggers was by a doctor. Data sent to IHI where it was checked. 11 hospitals reviewed 20 records per month.	In 854 patients 138 surgical AEs detected in 125 patients. 16 surgical AEs per 100 (14.6%) patients. 61 (44%) of the surgical AEs increased length of stay and 12 (8.7%) required life-saving treatment or led to permanent harm or death.	The surgical trigger tool may offer a practical easy-to-use approach to detecting safety problems in surgical patients. It can estimate the frequency of AEs and the impact of any interventions to prevent them.
Gurwitz et al (2003) Primary care ADE Comparison	Medicare enrolees aged over 64 years at a single group practice in New England, 1 July 1999 to 30 June 2000	Pharmacist employed multiple methods to detect ADEs using incident reports, review of discharge summaries, review of ED notes, computer-generated alerts (elevated drug levels, abnormal laboratory values, antidotes and ICD diagnoses of ADEs), administrative incident reports and automated free text review of notes. All events were confirmed by a physician.	Among the 1523 ADEs identified from 30,397 enrolees 28.7% were identified by computer alerts, 11% by incident reports, 11% by discharge summaries, 12% by ED notes review, 37% by free text searching, and 1% by administrative incident reports. Overall rate of ADEs was 50.1 per 1000 person years and 13.8 preventable ADEs per 1000 person years.	Comparison of methods to identify events was not main focus of study.
Haffner et al (2005) ADE Paediatrics Computerised	Comparison of ADRs between intensive surveillance by a physician and computer-assisted screening at 3 wards at HELIOS hospital, Germany, 2001	Intensified surveillance used a physician to undertake ward rounds and chart review while the computer-assisted tool used triggers that screened pathology results for values outside of a normal range. The records of these patients were then reviewed.	Intensified surveillance identified 101 ADRs in 11.8% of patients. Computer-assisted surveillance identified 45 ADRs in 5.7% of patients. The sensitivities of the surveillance system and the computer-assisted scheme were 67.2% and 44.8% and the specificity of the computer screening was 72.8%. The mean PPV of the triggers was 18.6%. ADRs detected by the intensified method were more severe, affected younger children, and had closer causal attributability than trigger-detected ADRs.	Triggers and intensive surveillance have different specificities. A higher number and more severe ADRs can be detected by intensified surveillance than by computerised surveillance but require more personnel resources.
Hope et al (2003) ADE Comparison Outpatient	33 ambulatory care clinics from Wishart Health Services, Indiana, US, during 4 months of 2001	Comparison of tiered approach vs nurse reviewer. Tiered approach began with trained data analysts applying queries to electronic health records for antidotes, toxicity and lab results, followed by nurse reviewers then pharmacist/physician check.	The PPV of the signal for ADEs was 10.2% and 9.6% for the 2 approaches (p = 0.36) but the cost per ADE was US$68.7 for pharmacist review and US$42.4 for the tiered approach.	Tiered review of ADEs is more cost-efficient than pharmacist review.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Huddleston et al (2011) Abstract only AE	All patients hospitalised at the Mayo Clinic, US, with congestive heart failure from 1 January 2005 to 31 December 2007 were included	Clinical records and administrative data assessed with GTT. Multivariate analyses used to assess. Multivariate regression analyses determined patient characteristics related to occurrence and timing of an AE. Time-dependent analyses were performed to determine cumulative density, hazard and probability density functions.	Among 1711 patients hospitalised with CHF, 38% had at least 1 AE. Hazard rate in the time to first AE was 0.019 events per hour. None of the patient-specific characteristics statistically influenced the probability of an AE occurring. However, age and Charlson Index were related to time to first AE. 70% of AEs occurred within 72 hours of admission.	Majority of work to date focused on the patient state. Analysis methods for assessing AE must begin to include aspects of care delivery system. These offer the highest potential to mitigating AE.
Jha et al (1998) Electronic screening ADE Comparison Reliability Accuracy	9 medical and surgical wards, Brigham Hospital, US, October 1994 to May 1995	Computerised detection rules based on out-of-threshold laboratory values, new medications, medications related to laboratory values. Based on Classen et al (1991). Rules modified during study, at the end there were 52 rules. Each rule was investigated by a trained reviewer. ADEs were defined by an additional review by a physician. Comparison with daily chart review by trained reviewers and stimulated voluntary reporting. All ADEs evaluated for severity and preventability in the same manner.	Reliability reviewers identifying ADEs, kappa = 0.53 and judgements by physicians 0.81–0.98, for preventability 0.92 and severity 0.32–0.37. 2620 alerts and 275 ADEs (9.6 per 1000 patient days) were identified. Chart review identified 398 ADEs (13.3 per 1000 patient days). Voluntary reporting identified 23 ADEs (0.7 per 1000 patient days). 76 of the 617 ADEs detected by all methods were detected by chart review and computer monitor, 3 were detected by computer monitoring and voluntary reporting. 139 (409) of the severe ADEs were identified by computer monitoring. Monitoring identified relatively more severe ADEs than chart review (p = 0.04) but not preventable ADEs. The PPV of the rules was 17%. The PPV of the individual rules varied from 9–28%. Monitoring required 11 person hours per week, voluntary reporting 5 and chart review 55.	Computer monitor identified fewer events compared with chart review but more than voluntary reporting. Small overlap of events from the methods so different methods may identify different types of events. Computer monitoring is an efficient approach to detect ADEs.
Kandpal et al (2012) Abstract only GTT AE	Unnamed venue February 2010 and February 2011	Application of IHI GTT at a tertiary facility. Every 2 weeks, 10 charts were randomly selected. A 20-min limit was set for review of each patient record. The review team consisted of 3 reviewers: a pharmacist and a nurse from Nursing Quality and a physician. Agreement by team on determination of AEs.	260 randomly selected patients' records were reviewed; 1067 triggers and 192 AEs were identified (74% of admissions). Top triggers associated with AEs include any operative complication, decrease in haemoglobin > 25%, any procedure complication, readmission within 30 days, partial thromboplastin time >100, investigations for DVT/PE. Top AEs include DVT, intra-op blood loss, pressure ulcers, healthcare associate infections, atrial fibrillation, bleeding from incisional site, hypoglycaemia and return to surgery. There were 108 AEs per 1000 patient days.	The IHI GTT is a springboard to identify areas to focus resources. IHI GTT identifies AEs that are missed using the voluntary reporting system.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Kenneley et al (2013) AE Accuracy Reliability	8 acute general hospitals at Baylor Medical System, Texas, US	Application of IHI GTT to hospitals and used as an ongoing monitoring tool with additional information about presence on admission, preventability relation to care provided or not and narrative descriptions about contributing factors. Patients with length of stay of 3 days or more were only included between January 2008 to June 2010. Patients admitted for addictive care, psychiatric illness or rehabilitation were excluded. Between 10 and 35 patient records were included each month depending on the hospital size. Records were reviewed by 1 of 4 nurse reviewers from an external company dedicated to the task. Periodic assessments of inter-rater reliability were made with small number of charts (approximately 94). Training sessions were conducted and information was provided to medical consultants for consideration.	16,172 records were reviewed and there were 14,184 positive triggers and 2772 AEs. There were 23.2 AEs per 100 discharges for patients with length of stay > 2 days and 5.5 per 100 discharges for length of stay less than 3 days. Trigger yield varied between 0 (4 triggers) – 100% (4 triggers). Overall trigger yield was 17.1% and surgical and medication modules provided most positive yields. Approximately 40% of the AEs were present on admission. 72% of AEs were deemed preventable. The inter-rater reliability between nurse reviewers in relation to whether or not there was an AE was 0.62. Some triggers had lower PPV than other reports suggesting some organisational refinement of the triggers is indicated (eg, mechanical ventilation had PPV = 7% in this study but 82% in the study by Naessens, 2010).	The GTT can be adapted to health care organisations' goals and resource limitations.
Kilbridge et al (2006) ADE Comparison	Automated surveillance system employed at a university hospital and a community hospital in Durham, North Carolina, March to October 2005	Duke University ADE surveillance system (antidotes, toxic drug levels, lab values) alerts are reviewed by a pharmacist who applies Naranjo algorithm. Physicians then review the ADEs.	1116 ADEs (900 patients) at the university hospital (4.4 ADEs per 100 admissions) and 501 ADEs (399 patients; 6.2 ADEs per 100 admissions) at the community hospital. Rates of antibiotic associated colitis, drug-induced hypoglycaemia, and anticoagulation-related events were higher at the community hospital. Computerised surveillance was 3.6 or 12.3 times higher than voluntary reporting at the university and community hospitals.	Automated surveillance detects higher rates than voluntary reporting. Community hospitals may experience higher rates of ADEs than academic centres.
Kirkendall et al (2012) Paediatric AE Reliability	Cincinnati Children's Hospital, US, 2009	Application of all 53 triggers of adult GTT to paediatric population. Trained nurse reviewers assessed triggers applied to 20 random records per month using IHI protocol with physician assessment of AEs.	404 triggers were detected and 88 AEs identified. 36.7 (27.8–45.6) AEs per 100 admissions and 76.3 (59.0–93.5) AEs per 1000 patient days. 25.8% (20.5–31.2%) of patients had a least 1 AE. 2 AEs required intervention to preserve life. 2 modules (cares and medication) identified 95% of the AEs. Inter-rater reliability between the 2 nurses for AEs was 0.63.	Utility of GTT shown in paediatric setting. Harm found to be 2–3 times higher than previously noted using other measures. The tool could be further modified for the paediatric setting.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Lander et al (2010) Paediatric surgical Reliability Comparison AE ADE	ENT Service, Children's Hospital, Boston	Development of an ENT-specific trigger tool based on Rozich et al (2003) and ENT clinicians. Training was undertaken. Final tool included 43 triggers and 6 domains (administrative, operative, discharge, nursing notes, clerical and medication). 50 inpatient charts randomly selected. 2 clinicians reviewed 20 charts to test reliability. Medical record review was conducted on all 553 charts by staff blind to trigger tool results.	236 triggers were identified, 92 of which were associated with errors. Admission triggers were found in 78% of records, medical record errors in 32%, operative triggers in 30%, discharge triggers in 30%, clerical triggers in 46%, medication triggers in 68%. Inter-rater reliability ranged between 0.35–0.90 for the trigger categories. Record review found errors in all admission (553 total) and 34 AEs. The trigger tool had 17% (14–20%) sensitivity, 82% (79–84%) specificity), 39% (33–46%) PPV and 59% (56–62%) negative predictive value. Triggers identified only 92 errors	Trigger tool was successful at identifying clerical and administrative errors and AEs but failed to identify complex AEs. A hybrid approach may be cost-effective for ENT.
Landrigan et al (2010) Reliability AE	Stratified random sample of 10 North Carolina hospitals, January 2002 to December 2007	100 admissions per quarter reviewed in random order by nurse reviewers from the hospital and external reviewers using GTT after training and standardisation. 2-stage review process with 52 triggers. Random effects Poisson regression model undertaken adjusting for hospital clustering, demographic variables, hospital service and risk conditions.	Among 2341 admissions 588 harms were identified for 423 admissions (18.1%), or 56.5 (52–61.2) per 1000 days or 25.1 (23.1–27.2) per 100 admissions. 2.9% of harms were permanent, 8.5% life threatening and 2.4% contributed to death. 17.9% were present on admission. There was no significant change over time. The reduction factor was 0.99 (0.94–1.04) for internal reviewers and 0.98 (0.93–1.04) for external reviewers. Inter-rater reliability kappa was 0.64–0.93 for internal review teams and 0.40–0.72 for external teams	Harms remain common – further efforts are needed to translate safety interventions into routine practice and to monitor health care over time.
Larsen et al (2008) ICU Paediatrics AE	Primary Children's Medical Centre, Salt Lake City, March 2002 – March 2003	Classen et al (1991) triggers were modified for paediatric ICU use. 2-stage process with chart review for triggers then detailed review if trigger positive.	507 AEs were identified from 259 admissions. 0.53 (0.47–0.57) AEs per patient day. 3% of AEs were serious. 183 AEs among 88 patients were preventable. 0.19 (0.16–0.22) per patient day. Patients with preventable events were younger, had higher illness and longer stays and were more likely to be surgical patients.	Preventable AEs are frequent but serious AEs are rare. Improved patient monitoring under increased risk conditions and improving early detection of harm will be more effective than strategies aimed at general error prevention.
Lemon and Stockwell (2012) Automated Comparison ADE	Children's National Medical Center Washington DC, US	2-stage review with first an automated assessment, second-stage physician review and determination of severity by NCC MERP system.	9143 triggers over 4 years. 2441 (34%) identified AEs. Only 75 (3%) of the AEs were identified by voluntary reporting. 552 (19%) of the AEs were considered preventable. The individual triggers ranged in PPV between 15% and 92.5%.	Automated AE identification by triggers has greatly improved quality of care.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Levinson (2010) Comparison AE	Acute hospitals in 2 counties, August 2008	Random sample of 278 Medicare beneficiary hospitalisations during 1 week. Comparison of 5 methods to screen for AEs: IHI nurse review of records, interview of patients or family members, hospital incident reports, use of POA coding to identify hospital-acquired conditions and AHRQ patient safety indicators. All positive flags then reviewed by physicians.	5 methods generated 662 flags. Physician review identified 256 events but as more than 1 flag identified many events there were 114 AEs. Plus another 6 from medical record review. IHI nurse review (93/120) and POA analysis (61/120) identified the most AEs. IHI nurse review also identified 35 events not flagged by any other method.	Nurse review is an effective way to identify AEs.
Levy et al (1999) ADRs Automated	34-bed medical ward, Hadassah University Hospital, Israel, 2 months, 1997	199 admissions subjected to screening by computerised alerts (lab values outside range, followed by chart review by clinical pharmacologists using Naranjo score.	295 alerts detected 43 ADRs among 40 patients. 10 ADRs were serious. 19% of the alert positive ADRs were not recognised by clinical staff.	The implementation of the monitoring doubled the number of ADRs recognised in the ward. The system is simple and valid.
Lipczak et al (2011a) Cancer care AE Comparisons	Application of trigger tool to 4 cancer surgery wards and 1 oncology ward at 5 different hospitals in Denmark during 2008	Comparisons made with incidents related to cancer-specific care reported to mandatory database and complaints provided to a patient database maintained by the Danish Cancer Society.	Some 260 events were noted among 570 records. Most (120) were related to clinical processes particularly healthcare associated infections (64) or medications (56).	The types of identified AEs varied in relation to the methods used, but each one generated different information
Louie et al (2010) Adult ICU ADE	Survey of medication errors and AE measurement methods at Canadian ICUs	Questionnaire of 146 pharmacist members of Canadian critical care pharmacy group at 79 ICUs in Canada in 2007.	34 responses from 31 (39%) of the ICUs. 26 (84%) of responders had a system for tracking medication errors and AEs: non-anonymous voluntary reporting 19 (73%), direct observation 14 (54%), anonymous voluntary reporting 12 (46%), chart review 6 (23%), computerised system 3 (12%), trigger tools 2 (8%), pharmacist intervention 2 (8%), weekly meeting 1 (4%). 14 (54%) of the ICUs with measurement methods had implemented changes to reduce AEs.	Most Canadian ICUs were measuring medication errors and AEs but a wide variety of methods were used. Only half had made any changes as a result of the measurements. Standardisation of measurement of medication error and AEs could be improved.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Marini et al (2012) AE Accuracy Reliability	Rouen University Hospital, France	For consecutive patients who underwent a neurosurgical procedure between 1 November 2008 and 30 April 2009, return to the operating theatre (ROT) within 30 days was identified based on the hospital information system associated to prospective payment. ROT was classified as planned or unplanned (UROT). UROT was further classified as related to the natural history of the disease or related to an AE (AEUROT). Meetings were organised to discuss results.	Among the 1009 procedures and 879 patients the information system identified 73 UROT cases (8.4%, 6.7–10.5%).The PPV was 61% (95% confidence interval (CI) 53–69%). Infectious AEs (*n* = 24, 2.4% (1.5–3.5%)), haemorrhagic AEs (*n* = 23, 2.3% (1.5–3.4%)), other cause AEs (*n* = 26, 2.8% (1.9–4.0%)), and infectious and other cause AEs (*n* = 2, 0.2% (0.0–0.7%)) were the most common reasons. Agreement between reviewers was high kappa = 0.88. Identification of required 4 hours/month time frame. 8 UROTs related to AE cases were discussed during mortality and morbidity meetings, leading to the identification of non-conforming care processes and practical improvement actions.	Unplanned return to theatre related to AE surveillance in neurosurgical patients was feasible and was a practical and useful tool to stimulate improvement.
Matlow et al (2011) Paediatric AE Accuracy	Various Canadian paediatric hospitals	5 existing trigger tools were consolidated using a delphi process to derive 47 triggers. The tool was validated on 591 randomly selected charts across 4 age groups with half medical and half surgical diagnoses at 6 academic paediatric hospitals. The triggers were applied with 2-stage process first by nurses and then physicians assessed for AEs.	Nurses rated the tool easy to use and identified triggers in 61.1% (95% CI 57.2 to 65.0) of patient charts; physicians identified AEs in 15.1% (89/591, 95% CI 0.23 to 0.43). Over a third of patients with AEs were neonates. The sensitivity and specificity were 0.88 and 0.44, respectively. Nurse and physician AE assessments correlated poorly. The PPV for each trigger ranged from 0–88.3%. Triggers with a false/true-positive ratio of > 0.7 were eliminated, resulting in the final 35-trigger.	This Canadian Tool is the first validated, comprehensive trigger tool available to detect AEs in children hospitalised in acute care facilities. This 35-trigger tool is reliable and robust, and can be used in quality-improvement initiatives and for more rigorous research agendas.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Matlow et al (2012) Paediatric AE	8 academic paediatric hospitals and 14 community hospitals in Canada	Random samples from 4 age groups. Records reviewed by nurses for triggers after training using standard form. Triggers assessed by physicians for AEs. 2-stage review with nurses.	1692 (46%) charts reviewed at academic hospitals and 1977 (54%) from community hospitals. The overall rate of AEs was 9.2% (95% CI 5.1–13.3%) Children in academic paediatric centres had significantly more AEs than those in community hospitals (11.2% (95% CI 6.4–15.9%) vs 3.3% (95% CI 1.2–5.3%)). The incidence of preventable AEs was not significantly different between types of hospital, but non-preventable AEs were more common in academic paediatric centres (adjusted odds ratio 4.39, 95% CI 2.08–9.27). Surgical events predominated overall and occurred more frequently in academic paediatric centres than in community hospitals (37.2% vs 21.5%, relative risk (RR) 1.7, 95% CI 1.0–3.1), whereas events associated with diagnostic errors were significantly less frequent (11.1% vs 23.1%, RR 0.5, 95% CI 0.2–0.9).	More children have AEs in academic paediatric centres than in community hospitals; however, AEs in the former are less likely to be preventable. There are many opportunities to reduce harm affecting children in hospital in Canada, particularly related to surgery, intensive care and diagnostic error.
Meuthing et al (2010) ADE	Cincinnati Children's Hospital	Triggers for AEs were developed using the hospital's computerised medical record (naloxone for opiate-related over-sedation and administration of a glucose bolus while on insulin for insulin-related hypoglycaemia). Triggers were identified daily. Based on information from the medical record and interviews, a subject expert determined if an ADE had occurred and then conducted a real-time analysis to identify event characteristics. Expert groups, consisting of frontline staff and specialist physicians, examined event characteristics and determined the apparent cause.	30 insulin-related hypoglycaemia events and 34 opiate-related over-sedation events were identified by the triggers over 21 months. The PPV of the triggers was 0.58 or 0.6. Only 5 of the 64 AEs (7.8%) were voluntarily reported. Patients receiving continuous-infusion insulin and those receiving dextrose only via parenteral nutrition were at increased risk for insulin-related hypoglycaemia. Lack of standardisation in insulin-dosing decisions and variation regarding when and how much to adjust insulin doses in response to changing glucose levels were identified as common causes of the AEs. Opiate-related over-sedation events often occurred within 48 h of surgery. Variation in pain management in the operating room and post-anaesthesia care unit was identified by the experts as potential causes.	Identification of ADEs through an automated trigger system, supplemented by in-depth analysis, can help identify targets for intervention and improvement.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Naessens et al (2009) AE Comparison	Inpatients discharged from Mayo Clinic hospitals, Rochester, Minnesota, 2005 (*n* = 60,599)	AEs were identified by: (1) AHRQ patient safety indicators excluding present on admission data (2) voluntary reported events (3) GTT (including physician confirmation).	2401 discharges (4%) had an AE identified by at least 1 method. Patient safety indicators were reported on 1576 discharges (2.6%). Mostly accidental puncture/lacerations (761/1576). 825 discharges had a reported event, most were skin integrity problems (43%) medication events (23%) or falls (21%). 235 discharges were reviewed by the trigger tools and 65 AEs (27.7%) were detected. AEs detected by 1 method were seldom identified by another. Only 97 (6.2%) of PSI events had a reported event and only 10.5% of reported events had a PSI.	Different detection methods identify different AEs. Combined approach may be best to measure patient safety in organisations.
Naessens et al (2011) AE Reliability Accuracy	Mayo Clinic campuses in Florida, Minnesota and Arizona	Electronic records (*n* = 1138) for 10 admissions randomly selected at each hospital every 2 weeks. 2 nurses independently review the records for 55 IHI triggers between 2004 and 2008. More detailed review after identification of a trigger established whether an AE had occurred. Second-stage physician review was included. 4 US hospitals.	PPVs for triggers varied between 80% (return to surgery) and 26% (intra-operative X-ray). Cases with AEs had more triggers than those without (average 4.7 vs 1.8, $p < 0.001$). Agreement between the nurses was good with mean kappa ranging from 0.53–0.73 for triggers and 0.4–0.6 for AEs. The agreement between nurses and physicians for AEs was higher (0.65–0.77). Agreement between nurses on individual triggers varied with lower levels with more subjective measures such as over-sedation kappa = 0.11 (-0.02–0.22) compared with more objective triggers such as INR > 6 kappa = 0.9 (0.76–1.0). Agreement about harm severity was low between nurses (kappa = 0.26–0.42) but higher between nurses and physicians (kappa = 0.48–0.76).	The trigger methodology appears to be a promising approach to the measurement of patient safety. However, the process was resource intensive and automated processes could make the process more efficient in identifying AEs.
Neubert et al (2006) Paediatric ADR Comparison	Paediatric ward at Children's University Hospital, Erlangen-Nuremberg, Germany	Intensive chart review by pharmacist and physician. Computer monitoring of hospital and laboratory records. Alerts were generated for abnormal values and important changes. ADRs classified by World Health Organization Adverse Reactions Terminology (WHO-ART) and Naranjo systems. In addition, comparison was made with reporting rates by treating physicians.	73 ADRs occurred for 439 admissions (396 patients). Computer alerts were generated for 31 ADRs (42%) at 27 admissions. 23 ADRs were identified by the treating physicians and not the computer. 8 ADRs were found by both the computer and physicians. The computer system had sensitivity = 90% and specificity = 20%.	Sensitivity is sufficient but specificity is too low for daily practice.
Nicol (2007) ADE Narrative case report	McLeod Regional Medical Centre, US	Institution report of introduction of series of process and automation improvements such as bar coding, medical management and medicine reconciliation. IHI GTT used to evaluate improvement.	Reduced harmful events from 35 per day to 1 or less between 2001 and 2006.	Minimal detail provided about use of tool.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Nilsson et al (2012) Adult ICU AE rate Comparison	6-bed ICU at 300-bed Swedish hospital	128 adult admissions who died in ICU or within 96 hours after discharge. 2-stage review, no time limit.	25 admissions (19.5%) suffered an AE. 41 AEs were noted or 32 AEs per 100 ICU admissions. 22 of the AEs (54%) were preventable. 12 were associated with death and 2 required intervention to avoid death. Health care associated infections, hypoglycaemia and pressure sores were the most common harmful events. 3 AEs were voluntarily reported.	About 1/5th of patients who died in the ICU were subject to harmful events. The trigger tool identified more AEs than traditional reporting systems. Limited by patient record. Subjectivity in assessments. Allowed longer time for record assessments and may have located more AEs.
Pravinkumar et al (2009) Abstract only ICU	ns	10 charts reviewed over 1 month by ICU team (5 medical and 5 surgical admissions).	41 triggers and 3 AEs were identified from a median of 30 (10–40) minutes chart review	The IHI model is effective at identifying triggers and AEs.
Rajesh et al (2012) Conference abstract only	Development and pilot testing of surgical triggers at academic hospital in India	List of triggers developed based on IHI and subjected to Delphi process selection with 5 clinicians. A list of 16 critical care, 19 surgical and 51 medication triggers were assessed against 247 case records.	60 triggers were identified in 140 cases (57%). Repeated request for lab investigations (43), use of laxatives (41), and Pyrexia (34) were common triggers.	Validating and implementation of this tool will enhance the identification of AEs.
Resar et al (2006) Adult ICU AE	62 ICUs in 54 hospitals in the US	Random sampling of admissions and stage review process employed between 2001 and 2004. 23 triggers were used for records of adults with stay > 1 day. Charts were assessed for up to 20 minutes. 20 charts per month.	12,074 records were reviewed and 11.3 AEs100 ICU days were noted (28% of the records had more than 1 AE. 60 AEs contributed to patient death and 165 required intervention to save life. Permanent harm was associated with 30 events and 353 (24.3%) prolonged stay. A small number of triggers were associated with most AEs – haemoglobin drop was associated with 201 episodes of harm. Medication-related AEs accounted for 18% (261) of AEs.	The trigger tool methodology is a practical approach to enhance AE detection which can direct improvement work.
Rosen et al (2010) Abstract only Ambulatory surgery	3 large healthcare systems in the US	Developed 6 ambulatory surgical AE trigger algorithms, 4 global and 2 specific. Applied triggers to a database of de-identified electronic data for patients with an ambulatory surgery between 01/01/05 and 12/31/05. 2 trained nurses reviewed a sample of 51 trigger-flagged cases per trigger from each healthcare system.	The ambulatory surgical AE triggers flagged between 1–22% of ambulatory surgery cases. There was a wide range in PPVs (6–62%).	Triggers have the potential to flag ambulatory surgeries with a possible surgical AE.
Schade et al (2006) AE	Pilot study at Bluefield Regional Medical Center, West Virginia, March 2005 to August 2005	Use of antidote (rescue) drugs was tracked across an electronic pharmacy system	1011 uses of a rescue drug were identified among 3572 discharges. For 109 cases an ADE was determined to have occurred and 29 cases were preventable. Most ADEs were related to anticoagulants or hypoglycaemic agents. 14% were severe but no deaths were identified.	Surveillance is feasible but labour-intensive. ADEs are under-reported.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Schildmeijer et al (2012) Reliability	5 hospitals Sweden with a team consisting of 2 nurses and 1 physician	50 cases from 1 hospital, 2009–2010, randomly selected for independent review by nurses in team looking for 53 triggers. The records were then reviewed by a physician who judged preventability.	The teams identified between 27.2 and 99.7 AEs per 1000 hospital days. Weighted kappa values for agreement for the detection of the number of triggers team by team was 0.32–0.6 with a combined unweighted kappa of 0.2 (0.12–0.3) and the weighted kappa for AE detection was 0.26–0.77 with combined unweighted kappa of 0.45 (0.26–0.63) which corresponded to moderate agreement.	The authors concluded the GTT should not be used for making comparisons between hospitals without substantially more training being given to reviewer teams. The study did not have a gold standard and included small number of cases and teams.
Seddon et al (2013) ADE NZ	3 district health boards, New Zealand	Random sample of charts March 2010 to February 2011 were reviewed for 19 triggers with positive charts further evaluated for ADEs by team with clinical pharmacist.	353 ADEs were identified among 1210 charts. The average ADE rate was 28.9 per 100 admissions or 38/1000 bed days. Most ADEs were minor but 5 were associated with fatalities, 4 permanent harm and 9 required intervention to preserve life. The most sensitive triggers were cessation of medication and anti-emetics. Morphine, warfarin and tramadol were most frequently associated with an ADE. None of the ADEs were reported at 1 district health board.	Higher rates of ADEs are identified by the trigger tool compared with voluntary reporting. The tool provides a standardised measure of harm that can be used to determine trends and the impact of safety programmes.
Seynaeve et al (2010) ADE ICU	Single Belgium ICU	1009 inpatient days for 79 patients assessed for prevalence of ADEs.	230 ADEs observed, the most frequent were hypoglycaemia and hypokalaemia. 4% were severe.	ADEs are common in the ICU.
Sharek et al (2006) ICU neonatal Comparions Development AE	15 NICUs in the US	6 neonatologists developed a list of 38 triggers thought likely to identify 24 AEs. The tool was piloted at 4 sites with 42 charts. 21 triggers were removed and the final tool of 17 triggers was applied to 749 randomly selected charts with 17,106 bed days. The version was tested and applied to each hospital with central coordination, training and standardisation. Retrospective chart review comparison was undertaken of the triggers.	2218 triggers were detected (2.96 per patient) and 554 AEs were identified (0.74 per patient). The mean PPV for the triggers was 0.38. The mean chart review time was 20 minutes. The mean AE rate per 1000 patient days was 32.4. 56% of all AEs were preventable, 16% could have been identified earlier and 6% could have been mitigated more effectively. Only 85 AEs were identified by voluntary reporting.	AEs rates in NICU setting are higher than previously described. Many result in permanent harm and many are preventable. The NICU tool is efficient and effective at identifying AEs. No gold standard for AE detection so assume that all AEs identified are the total of all AEs. Some subjectivity l in assessments was noted.
Sharek (2009) Comparison AE	Assessment of suitability of GTT as a measure of harm at individual hospitals and role in a national harm measurement system	Retrospective chart review of 10 charts per quarter from 10 randomly selected hospitals in North Carolina between 2002 and 2007. Charts were reviewed by internal hospital team, external team and an IHI group. Each team separately applied the GTT methodology.	Internal hospital teams found average AE rates of 22.9 per 100 patients (21.2–24.9), external teams identified rates of 17.2 (15.6–19) and IHI team found 36.6 per 100 patients (28.8–46).	The researchers concluded that there was relatively good agreement between the teams and the GTT could be used as a measure of harm for individual hospitals and nationally.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Sharek et al (2011) Reliability AE	10 North Carolina hospitals	Retrospective chart review of 10 charts per quarter from 10 randomly selected hospitals in North Carolina between 2002 and 2007. Charts were reviewed by internal hospital team, external team and an IHI group. Each team separately applied the GTT methodology.	Moderate (kappa 0.64) to almost perfect (kappa 0.93) agreement between internal reviewers and external reviewer team. The internal team had higher sensitivity (49% vs 34%) and specificity (94% vs 93%) compared with the external team.	GTT has favourable inter-rater and intra-rater reliability.
Singh et al (2009) Outpatient ADE Accuracy	6 ambulatory practices in New York state	Developed own trigger tools based on Gurwitz et al (2003) without administrative data related triggers leaving 39 triggers. Evaluation by trained reviewer then pharmacist/physician. 12-month retrospective chart review of patients aged 65 and older with cardiovascular diagnoses.	1289 charts were reviewed and 645 (50%) charts had at least 1 trigger (1733 in total). A random sample of 383 charts was further reviewed – 232 ADEs were identified of which 92 were preventable. 30% of the ADEs were severe (hospitalisation, permanent injury or death). The top 9 triggers identified 94% of the ADEs. The PPV of the triggers varied from 6.7–100%.	Trigger tools have a potential role in quality improvement. A briefer tool may be useful.
Szekendi et al (2006) Automated triggers ADE (not just ADE) Active surveillance	Northwestern Hospital, Chicago, date unspecified	21 electronic triggers based on laboratory values, high-risk and antidote medication used to identify records, subjected to medical record review by nurse and pharmacist and AEs then determined by a physician.	At least 1 AE identified in 243 (74%) of 327 records. 163 preventable AEs. 4% of AEs gave permanent harm, 10% required intervention to preserve life and 1% contributed to death. High INR and positive blood cultures were the most sensitive triggers.	The study provides a useful algorithm for defining an AE based on Harvard Medical Practice Study. The active surveillance methodology allows for identification of AEs among hospitalised patients that provides a unique opportunity to intervene to mitigate harm.
Takata et al (2008b) Paediatric Comparisons ADEs	5 California Pediatric Safety Initiative hospitals between Nov 2003 and April 2004 (25,763–41,831 inpatient days).	Comparison between pharmacy intervention medication errors (actions taken by pharmacist when they receive an order that contains an error) and IHI GTT (7 medication use and 3 laboratory tools) using a sample of 40 discharges per month and finally voluntary reporting.	Pharmacy intervention errors were 2.67 per 1000 inpatient days, trigger tools generated 22.3 AEs per 1000 inpatient days and voluntary reporting 1.7 per 1000 inpatient days. The methods identified different types of events. Trigger tools identified more ADEs by a factor of 11 and triggers had a PPV of 16.8%. ADEs identified by any method mostly occurred among patients aged 1 year or older during days 0 and 1 of admission and mainly concerned anti-infectives, analgesics and electrolyte and water balance replacements.	The authors concluded that the study provided useful baseline rates of AEs in paediatric hospitals and that trigger tools were the most effective at identifying AEs.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Takata et al (2008a) Paediatric Development comparison other ADE	80 patients randomly selected at each of 12 children's hospitals in the US	The IHI GTT was applied to 900 charts and a paediatric population. The paediatric version was tested and applied to each hospital with central coordination and standardisation Retrospective chart review comparison of triggers.	2388 triggers were identified with 960 patient charts. 107 ADEs were located. The PPV of the triggers was 3.7%. Trigger ADE rates were 9.27 per 100 patients, 13.1 per 1000 patient days. 22% of all ADEs were deemed preventable and 3% severe. The most frequent ADEs were pruritus and nausea and the most commonly associated drugs were opioid analgesics and antibiotics. Only 3.7% of the ADEs were identified by voluntary reporting. Trigger tool identified 89/107 ADEs and incident reporting 4/107.	The trigger tool is effective at identifying ADEs in inpatient paediatric populations.
Tegeder et al (1999) ADR Automated	Single ward at University of Erlangen Hospital January 1996 to May 1997.	19 laboratory values exceeding defined boundaries were used as triggers to prompt an evaluation of medical record for an ADR using Naranjo probability scale by a physician.	229 signals were generated for 98 patients. 18 cases of ADRs were noted. In 12 of the 18 cases the clinical team had not identified the reaction during the inpatient stay. 3 of the ADRs were serious.	Increased awareness of ADRs through automated laboratory signals will increase the recognition rate of ADRs and may help prevent them.
Thuermann et al (2002) ADR Accuracy	Neurology (86 beds) department at teaching hospital in Germany (Wuppertal GmbH) over 3 months in 1999	Computerised triggers using laboratory data outside of set boundaries. Alerts were checked by pharmacist and then a neurologist. Definitions of ADR were according to WHO criteria. Comparison with intensified surveillance using daily ward rounds by clinical pharmacologist with subsequent review by neurologist.	From 600 admissions there were 501 triggers among 231 patients. 121 of the triggers were judged related to an ADR in 111 patients (18.5%). 16 of the ADRs were severe (2 deaths). PPV of the triggers ranged from 0–100%. The highest were for high INR or increased serum concentrations. Only half of the ADRs could be detected by the triggers so sensitivity = 45.1% and specificity 78.9%.	High number of ADRs on neurology wards. The majority of ADRs could not be detected by the triggers.
Von Plessen et al (2012) GTT AE	5 hospitals in Denmark, January 2010 to June 2011	Application of translated GTT for use in Danish hospitals. Interviews with team members at each location. GTT results presented as run charts.	Background information about hospitals, GTT teams, training and procedures is presented. There were local differences between teams with their training and procedures. Reported incidents varied between 3–12 per 1000 patient days at the hospitals and the average GTT harm rates were 60 per 1000 patient days and the range 34–84 per 1000. The percentage of patient harmed was 25% (range 18–33%). Most harm was temporary (96%).	Variation in harm rates – differences in training, procedures and documentation probably contributed to this variation. Training reviewers as teams specifying roles and the use of training records and a database for results may improve the application of the tool.
Yeesoonpan et al (2011a, 2011b) GTT for ADE Abstract only	Pilot study (date unspecified)	IHI GTT applied to 20 charts from 7 hospitals across Thailand. Limited description of methods.	188 triggers were recorded from 136 charts. 17 ADEs were identified using 8 triggers.	Thai modification is feasible.

Author, date, reference, keywords	Setting	Methods	Results	Authors' conclusions
Yeesoonpan et al (2011a) ADE	Paediatric inpatients (date unspecified)	Tool applied to 20 charts at paediatric Thailand hospital	76 triggers found among 20 charts. 5 ADEs observed.	Suggestive results to facilitate trigger modifications.
Zimmerman et al (2010) GTT AE	Application of IHI GTT to mortality review process at a Canadian hospital between 2008-9	2-step process of AE identification based on IHI methodology.	Among the 1817 deaths reviewed 14% were associated with an AE.	The process resulted in a number of systems improvements.
Zolezzi et al (2007) ADE New Zealand Comparison	Assessment of a modified trigger tool at a single hospital in New Zealand	The trigger tool was modified from Classen (1991)[7] and focused on high-risk medications (warfarin heparin, morphine, benzodiazepines) looking for the use of any reversal agents or laboratory parameters used as triggers. 528 patients were assessed from July–August 2005.	Among the 286 patients who received at least 1 of the study medicines, 45 patients (8.5%) were identified as having an ADE. Agreement between the researchers for the identification of the ADEs was 88%. Morphine was associated with the highest number of ADEs (30). The trigger was able to identify considerably more ADEs than generated by the spontaneous reporting system (0.07%).	Modified trigger tools is a sensitive method to detect ADEs and yields more events than voluntary reporting. The seriousness of the ADEs was not assessed and the study considered a limited number of medications.