International Journal for Quality in Health Care, 2023, **35(3)**, 1–8
DOI: https://doi.org/10.1093/intqhc/mzad056
Advance Access Publication Date: 13 July 2023
Systematic Review

Comparing rates of adverse events detected in incident reporting and the Global Trigger Tool: a systematic review

Peter D. Hibbert[1,2,3,*], Charlotte J. Molloy[1,2,3], Timothy J. Schultz[4], Andrew Carson-Stevens[5], Jeffrey Braithwaite[1]

[1]Australian Institute of Health Innovation, Macquarie University, 75 Talavera Rd, Macquarie Park, New South Wales 2109, Australia
[2]IIMPACT in Health, Allied Health and Human Performance, University of South Australia, GPO Box 2471, Adelaide, South Australia 5001, Australia
[3]South Australian Health and Medical Research Institute, North Terrace, Adelaide, South Australia 5000, Australia
[4]Flinders Health and Medical Research Institute, Flinders University, Sturt Rd, Bedford Park 5042, South Australia, Australia
[5]PRIME Centre Wales & Division of Population Medicine, Cardiff University, Heath Park, Cardiff, Wales CF14 4XN, United Kingdom

[*]Corresponding author. Australian Institute of Health Innovation, Macquarie University, 75 Talavera Rd, Sydney Macquarie Park, New South Wales 2109, Australia. E-mail: peter.hibbert@mq.edu.au
Handling Editor: Dr Phillip Phan

Abstract

Many hospitals continue to use incident reporting systems (IRSs) as their primary patient safety data source. The information IRSs collect on the frequency of harm to patients [adverse events (AEs)] is generally of poor quality, and some incident types (e.g. diagnostic errors) are under-reported. Other methods of collecting patient safety information using medical record review, such as the Global Trigger Tool (GTT), have been developed. The aim of this study was to undertake a systematic review to empirically quantify the gap between the percentage of AEs detected using the GTT to those that are also detected via IRSs. The review was conducted in adherence to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. Studies published in English, which collected AE data using the GTT and IRSs, were included. In total, 14 studies met the inclusion criteria. All studies were undertaken in hospitals and were published between 2006 and 2022. The studies were conducted in six countries, mainly in the USA (nine studies). Studies reviewed 22 589 medical records using the GTT across 107 institutions finding 7166 AEs. The percentage of AEs detected using the GTT that were also detected in corresponding IRSs ranged from 0% to 37.4% with an average of 7.0% (SD 9.1; median 3.9 and IQR 5.2). Twelve of the fourteen studies found <10% of the AEs detected using the GTT were also found in corresponding IRSs. The >10-fold gap between the detection rates of the GTT and IRSs is strong evidence that the rate of AEs collected in IRSs in hospitals should not be used to measure or as a proxy for the level of safety of a hospital. IRSs should be recognized for their strengths which are to detect rare, serious, and new incident types and to enable analysis of contributing and contextual factors to develop preventive and corrective strategies. Health systems should use multiple patient safety data sources to prioritize interventions and promote a cycle of action and improvement based on data rather than merely just collecting and analysing information.

Keywords: patient safety; systematic review; quality of care; hospital incident reporting; adverse events; healthcare Global Trigger Tool

Introduction

Hospitals require real- or near real-time information to understand whether they are delivering safe care to patients and to inform interventions to reduce adverse events (AEs) (harm to patients) [1]. The lack of adequate detection and monitoring of AEs is a major factor in their persistence [1, 2]. Measurement of types and frequencies of AEs informs patient safety priorities for corrective strategies and tracking progress over time and against peers [1]. The challenges for patient safety measurement in healthcare systems have been outlined, with solutions for adoption by member states and their healthcare systems, in the WHO Patient Safety Acton Plan (2021–2030) which calls on governments to 'strengthen synergies and data-sharing channels between sources of patient safety information for timely action and intervention …' [3].

One reason why hospitals do not adequately detect AEs and monitor their prevalence may be their use of incident reporting systems (IRSs) as their primary patient safety information data source [4, 5]. IRSs tend to collect poor quality information on the frequency of harm, and certain incident types, such as diagnostic errors are consistently under-reported [4, 6]. Over-reliance on IRSs thereby can compromise a hospital's quantitative understanding of AEs [4].

One frequently used method of collecting patient safety prevalence information is the Global Trigger Tool (GTT) [7–9]. The GTT was designed to provide hospitals with 'an easy-to-use method for accurately identifying AEs and measuring the rate of AEs over time' [8]. The GTT involves the screening of medical records for the presence of triggers, followed by a more in-depth manual review for the presence of an AE. After AEs have been detected with the GTT, their rates may be calculated and displayed graphically over time [8]. Originally developed for adult inpatients in 2003, the GTT has since been modified for hospital specialties [10–13] and

Received 2 January 2023; Editorial Decision 3 July 2023; Revised 21 June 2023; Accepted 11 July 2023
© The Author(s) 2023. Published by Oxford University Press on behalf of International Society for Quality in Health Care.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/intqhc/article/35/3/mzad056/7223853 by guest on 07 August 2023

primary care [13–15] with a second edition ('the GTT Protocol') published in 2009 [8]. Medical record review, using structured tools like the GTT, is considered to be one of the patient safety data sources most amenable to measuring rates of AEs, whilst IRSs are not suitable for reliable measurement purposes largely owing to reporting biases [16]. Healthcare services may use the GTT as an adjunct to IRSs to detect and measure AEs [8]. Tchijevitch et al. [17] found that an IRS alone was insufficient as a single method for quantifying the occurrence of serious or fatal adverse drug events (ADEs) and that the GTT could be beneficial as one other data source. In a secondary finding of our 2016 systematic review of the GTT, IRSs detected only an average of 4% (range 2–8%) of AEs detected using the GTT across eight studies [18]. However, there are no syntheses of direct comparisons of AE data collected by IRS and the GTT. Given many hospitals' widespread use of and arguably over-dependence on IRSs [4, 5], the poor quality of information that IRSs provide on the prevalence of harm and that the GTT is designed as a more reliable tool to be used by hospitals to measure AEs, we sought to compare the two methods. The aim of this study was to undertake a systematic review to empirically quantify the gap between the percentage of AEs detected using the GTT to those that are also detected via IRSs.

Methods

A systematic review and narrative synthesis was conducted in adherence to the PRISMA statement [19]. We searched MEDLINE, EMBASE, and CINAHL for articles for all time up to 27 March 2023 using the search term 'Global Trigger Tool'. We also hand-searched the key journals including 'BMJ Quality and Safety', the International 'Journal for Quality in Health Care', 'Health Services Research', 'BMJ Open', 'Pediatrics', 'Journal of Evaluation Clinical Practice', 'Joint Commission Journal', 'Journal of General Internal Medicine', 'Journal of Patient Safety Risk Management', 'Journal of Patient Safety', 'American Journal of Medical Quality', and included all eight studies previously identified [18]. Snowball searching of included articles was also undertaken. Variants of the GTT were included. Studies were limited to those published in the peer-reviewed literature; doctoral theses were excluded. Fig. 1 depicts the search process.

Study selection

Two authors (C.J.M. and P.H.) independently screened all titles, abstracts, and potentially relevant articles for full-text review. Any disagreements about the eligibility of studies were resolved through discussion until consensus was reached. Studies published in English that compared AE rates using a variant of the GTT with AEs detected by IRS were included. In total, 14 studies met the inclusion criteria.

Data extraction

Two authors (C.J.M. and P.H.) extracted and compiled data from each paper. The publication data and study demographics (authors, year of publication, and country), speciality (healthcare type and speciality), GTT methodology (number of institutions, sample size, AE definitions, number of reviewers, use of inter-rater reliability (IRR), and patient safety classifications), and results data (AE rate measured by GTT and IRS), were all extracted. The GTT methodology was abstracted due to the considerable heterogeneity and deviations from the GTT protocol that we previously found within studies [18].

Quality assessment

The included studies were critically appraised using the Quality Assessment Tool for Studies with Diverse Designs

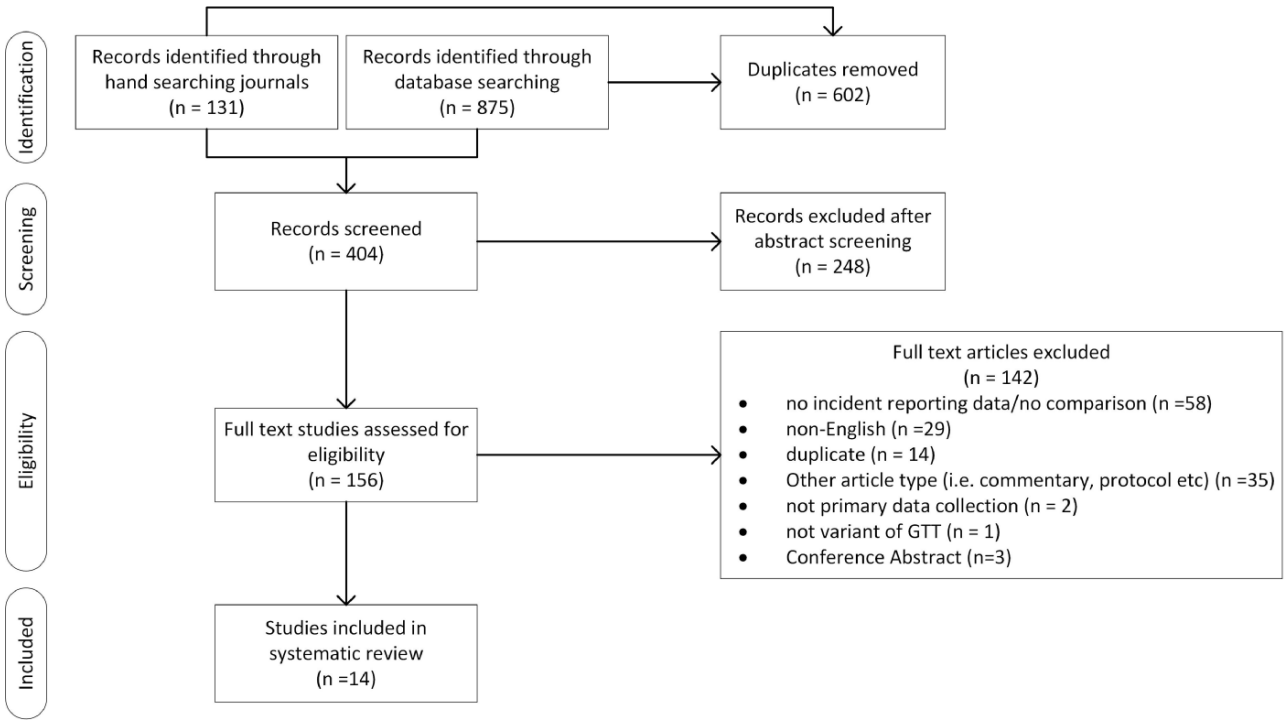

Figure 1 Systematic review flow diagram detailing the numbers of articles found, abstracts screened, and full texts reviewed.

Downloaded from https://academic.oup.com/intqhc/article/35/3/mzad056/7223853 by guest on 07 August 2023

Table 1. Reasons for undertaking the study.

Reason for undertaking the study	*n*	Reference number
Measuring AE rates	11	[10, 17, 21–29]
Characterising AEs	7	[10, 21, 23, 25–27, 29]
Comparing the GTT with other AE data sources	7	[17, 22, 24, 26, 30–32]
Developing GTT specialty versions	4	[21, 27–29]
Assessing GTT utility	3	[22, 24, 25]

(QATSDD) [20]. This 16-item tool allows for methodological evaluation of studies, with papers scored from 0 to 3 on each item. Potential overall scores range from 0 to 48, with a higher score indicating greater methodological rigour. Two reviewers completed the quality assessment and consensus was reached through discussion.

Results

The literature identified 404 potentially relevant, non-duplicate articles. After reviewing the titles and abstracts, we excluded 248 articles and 156 were read in full-text form (Fig. 1). Fourteen articles met our inclusion criteria. Of these, 11 (79%) studies cited measuring the AE rate as a reason for undertaking the research (Table 1). Characterising AEs (for example, using incident types, preventability, or severity) was the second most frequently cited reason (7/14, 50%), followed by comparing the GTT with other AE data sources (7/14, 50%).

Demographics and methodology—sampling

All 14 included studies were undertaken in hospitals of six countries with nine (64%) in the USA (Table 2, Table A.1). The studies were published between 2006 and 2022. ADEs only were collected in two studies [17, 29]. Over one-third of the studies were undertaken in a single institution (6/14, 43%) (Table 2), and a total of 22 589 medical records were reviewed across 107 institutions. A total of 7166 AEs were found.

Methodology—data collection and analysis

Definition of AE

While the GTT method was stated in all studies, only four of 14 (29%) explicitly used the GTT protocol's [8] AE definition ('unintended physical injury resulting from or contributed to by medical care that requires additional monitoring, treatment or hospitalization, or that results in death'). Five (36%) studies used the following definition (or a modification of it) 'an injury, large or small, caused by the use (including non-use) of a drug, test or medical treatment', four studies (29%) reported no definition and one (7%) study used an Institute of Medicine [33] definition ('an event leading to patient harm and caused by medical management rather than the underlying condition of the patient').

Number of reviewers

The GTT protocol recommends assignment of two primary reviewers and one authenticating physician [8]. Just under half of the studies (6/14, 43%) (Table A.2) used this method. The most frequently used other method was one primary and one secondary reviewer (3/14, 21%).

Table 2. GTT studies by country, speciality, and number of hospitals.

	n	Reference number
Country		
United States	9	[21, 23, 24, 27–32]
Sweden	2	[10, 26]
Australia	1	[22]
Canada	1	[27]
Denmark	1	[17]
Palestine	1	[25]
Total	**15**[a]	
Speciality		
General inpatients	5 (11,125)	[23–26, 30]
Paediatric	2 (1560)	[28, 29]
Paediatric intensive care unit (PICU)	2 (794)	[21, 22]
Inpatient oncology	1 (88)	[32]
Inpatient psychiatric	1 (8005)	[31]
Intensive care unit (ICU)[b]	1 (128)	[10]
Neonatal intensive care unit (NICU)	1 (749)	[27]
Unplanned transfer to ICU or death	1 (141)	[17]
Total	16 (22,589)	
Number of hospitals		
1	6	[10, 17, 22, 24, 26, 32]
2–5	2	[25, 30]
6–10	2	[23, 28]
11–15	3	[21, 27, 29]
>15	1	[31]
Publication year		
2006–2010	4	[21, 27, 29, 30]
2011–2015	7	[10, 22–26, 28]
2016–2020	1	[31]
2021–2023	2	[17, 32]

[a]One study [27] took part in two countries so total is greater than number of papers.
[b] Nilsson applied the GTT to patients who had died in ICU.

Inter-rater reliability

Only two studies measured IRR and only one of those reported the results ($k = 0.58$ between two primary reviewers and $k = 0.89$ between primary and secondary) [25].

Use of severity of harm scale

The National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) [34] was used as the scale of harm to report GTT AEs in 93% (13/14) of the studies (Table A.3).

Methods of reporting AE rates

The percentage of admissions with an AE was the most frequent method used (11/14, 79%) followed by AEs per 1000 (or 100) patient-days (8/14, 57%) then AEs per 100 admissions (5/14, 36%) (Table A.4). All three methods were used in four studies (29%).

Quality of included studies

The mean QATSDD score was 26.71 (SD = 2.6, range 22–31) with a score of 48 being the maximum possible. Quality scores for individual studies are presented in Table 3 and in more detail in Table A.5. Studies largely performed best on criteria related to fit between research question and method, statement of aims/objectives, and clear description of research

Downloaded from https://academic.oup.com/intqhc/article/35/3/mzad056/7223853 by guest on 07 August 2023

Table 3. Included studies, demographics, and results.

Ref.no.	First author	Speciality	Quality score	Sample size of the GTT—number of medical records reviewed	Number of admissions	AE *N* and rates detected by the GTT				% of AEs detected using the GTT that were also detected in the corresponding IRSs
						AEs (*N*)	Per 1000 patient days	AEs per 100 admissions	% of admissions with an AE	
[21]	Agarwal	Paediatric intensive care unit (PICU)	29	734		1488	286[a]	203[b]	62	4
[22]	Hooper	Paediatric intensive care unit (PICU)	23	59		98	599[a]	166[b]	56	2
[23]	Kennerly	General inpatients	28		9017	3430	61	38	32	3.5
[24]	Mull	General inpatients	28	273	288	109	52	38	21	0.9
[30]	Naessens	General inpatients	23		235	65	–	28[b]	28	14
[25]	Najjar	General inpatients	28	640		91	–	14[b]	14	4.4
[10]	Nilsson	Intensive care unit (ICU)	25	128		41	–	32	20	2.4
[31]	Reilly	Inpatient psychiatric	31	8005		583	–	7[b]	–	37.4
[26]	Rutberg	General inpatients	26	960		271	33	28[b]	21	6.3
[32]	Samal	Inpatient oncology	22	88		79	–	90[b]	–	2.5
[27]	Sharek	Neonatal intensive care unit (NICU)	28	749		554	32	74[b]	–	8
[28]	Stockwell	Paediatric	26	600		240	55	40	24	9.2
[29]	Takata	Paediatric	30	960		107	16	11	7	3.7
[17]	Tchijevitch	Unplanned transfer to ICU or death	27	141		10[c]	–	7[b]	7[c]	0

[a] Reported as 100 patient days in original papers (calculated to 1000 patient days).
[b] Calculated by authors of this paper.
[c] Fatal and life threatening ADEs only.

Downloaded from https://academic.oup.com/intqhc/article/35/3/mzad056/7223853 by guest on 07 August 2023

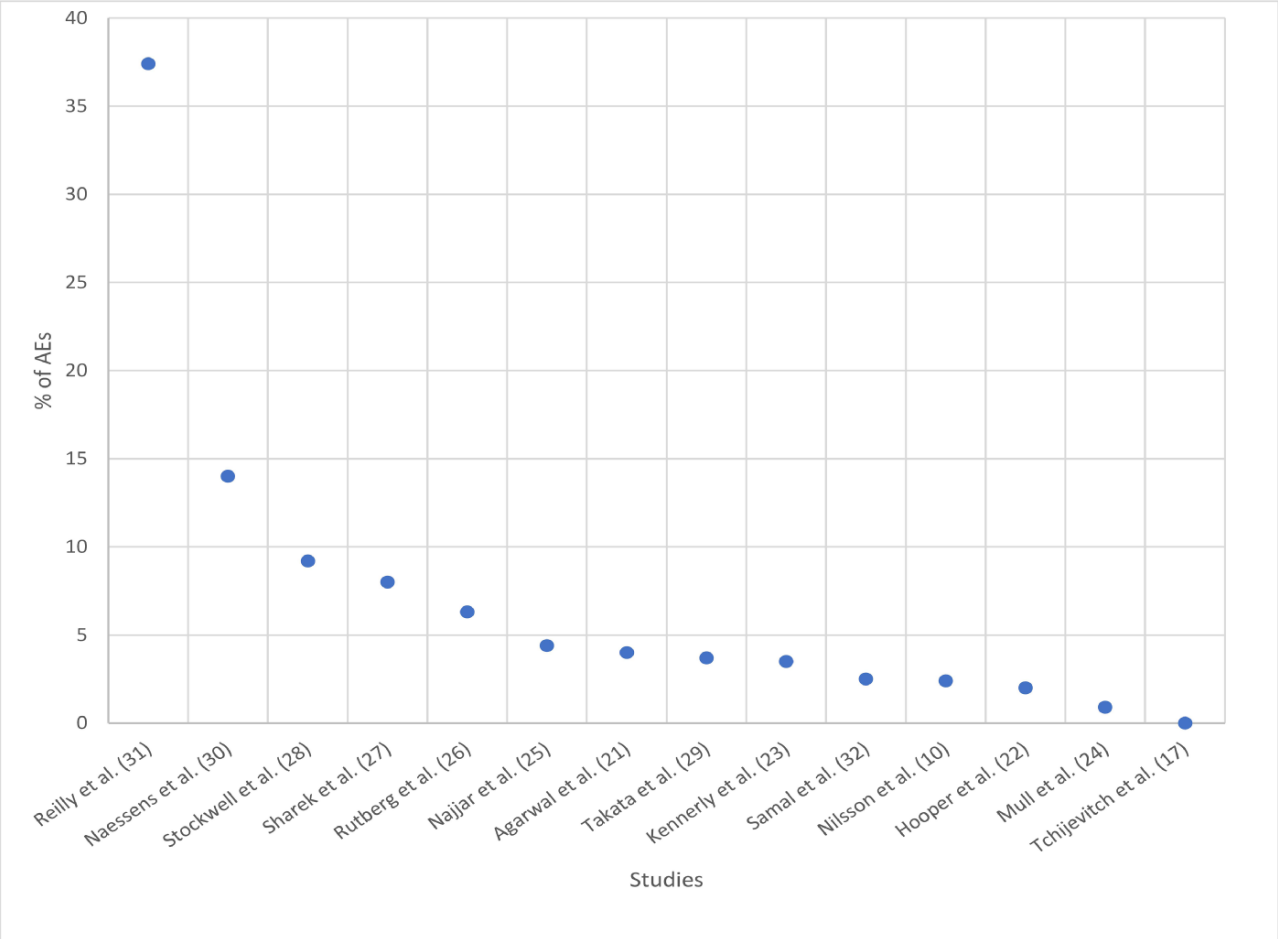

Figure 2 Percentage of adverse events detected using the Global Trigger Tool that were also found in corresponding incident reporting system by research study.

setting. However, studies typically performed poorly on items related to sample size, theoretical frameworks, and user involvement in design.

Comparison of adverse event reporting rates using the GTT and IRSs

The percentage of AEs detected using the GTT that were also detected in corresponding IRSs ranged from 0% to 37.4% (Fig. 2, Table 3). There was an average of 7.0% of AEs detected with the GTT also detected by IRSs (SD 9.1; median 3.9 and IQR 5.2). Twelve of the fourteen studies found <10% of the AEs detected using the GTT were also found in corresponding IRSs.

Two studies included in this review compared rates of serious AEs collected by the GTT and IRS. Tchijevitch et al. [17] detected 10 serious or fatal AEs from 141 medical records using the GTT with none of these (0%) being detected by IRS. In another study related to inpatient psychiatry, the IRS detected just over half (53%) of moderate to severe harm AEs that were detected by GTT [31].

Three studies also analysed AEs detected by IRS that were not detected by the GTT [22, 30, 32]. These studies found 50% ($n = 2/4$) [22], 18% ($n = 2/11$) [30], and 90% (19/21) [32] of AEs detected by the IRS were not detected by the GTT.

Discussion

Statement of principal findings

We found that most AEs occurring in hospitals and identified by the GTT are unlikely to be detected by IRSs. In 12 of the 14 studies reviewed, the rate of AE detection by IRSs was <10% of those detected by the GTT. The average of AEs detected with the GTT that were also detected by IRSs was 7%.

Strengths and limitations

This review is the first to undertake a standalone systematic review to answer a recurring question in the quality and safety field: to compare rates of AEs using the GTT with AEs detected via IRS. Strengths include a systematic search strategy involving multiple data sources and reporting according to PRISMA guidelines, and included studies were critically appraised for quality. The quality of included studies as measured by the mean QATSDD score was higher (26.7 versus 19.7 [35] and 21.8 [36]) compared with two other systematic reviews assessing patient safety data sources.

The rate of AEs detected by the GTT is the comparator in this study. The GTT AE rates varied widely across studies: 7–203 AEs per 100 admissions across the 14 studies and 7–28 AEs per 100 admissions for the five general medical studies [23–26, 30] (Table 3). This wide variation is likely to have three explanations. Firstly, differences in setting: the

Downloaded from https://academic.oup.com/intqhc/article/35/3/mzad056/7223853 by guest on 07 August 2023

four highest AE rates (74–203 per 100 admissions [21, 22, 27, 32]) (Table 3) occurred in paediatric or neonatal intensive care units or inpatient oncology, where higher rates are more likely. Secondly, heterogeneity of GTT methods outlined in the results and thirdly, reviewers' judgements, for example, AEs involving minor harm are generally less easily identified and the GTT reviewers need to apply considerable discretion which may result in variation of perceptions [18]. The heterogeneity in methods, notably definitions of AE, and reviewer judgement, may reduce the utility of the GTT as a comparator to IRSs.

As to limitations, one is the small number of included studies. It is possible that some relevant studies were not captured by the search strategy. There remains a possibility of bias because non-English publications were omitted. It is also possible that publication bias affected the results of this study. The information within the included studies is reliant on the information captured in the medical records and, as our previous systematic review on the GTT's use found, there is heterogeneity in how this is captured and recorded in medical records between and within health services and studies [18]. The GTT was designed to be used in general inpatients; however, 9/14 papers included in this review applied the GTT in other specialties which tend to have different triggers for AEs which may impact on the rate of AEs detected. Research using the GTT methodology that yielded low levels of AEs may be less likely to be published than studies with higher AE levels.

Interpretation within the context of the wider literature

This study compared the rates of incidents from two data sources, IRSs and the GTT. However, there are other data sources available to health services to allow them to measure and characterize their safety profile. These include patient complaints, medico-legal claims, executive walk-arounds, investigations, observation of patient care, and administrative data analysis [2, 37–39]. Each have strengths and weaknesses; for example, medical record reviews, such as the GTT may allow health services to compare rates of incidents over time, but they are time-consuming and resource intensive often requiring experienced clinicians to identify and judge AEs [2].

Not only do different data sources have particular methodological strengths and weaknesses, they also tend to collect different incident types. Levtzion-Korach et al. in 2010 compared five data sources and found there was little overlap in incident types between sources [38]. For example, IRSs tend to collect incidents related to patient identification, falls, and medication; medicolegal claims collect incidents related to clinical judgement related to diagnosis and treatment, communication, and problems with medical records; whilst patient complaints tend to collect incidents on communication and administrative issues, such as admission and discharge [16]. The implication of patient safety data sources' methodological strengths and weaknesses together with the heterogeneous incident types collected means that best practice is for health services to collect and strategically analyse multiple types of data sources [16, 39].

The utility of patient safety data collection and analysis methods are evolving. One is the use of trigger tools, such as the GTT 'prospectively' or in 'real-time' [40]. This involves an integrated clinician working within a medical department reviewing medical records within 48 h of patient admission (and who still may be an in-patient); this is followed by a multi-disciplinary discussion to elicit staff perspectives of what happened and why, to determine contributing factors and change ideas to inform possible interventions to improve the safety of future care. This method was designed to overcome a weakness of retrospective medical reviews that can be poor at understanding contributing factors from a review of medical notes alone. The method also fits with Macrae's notion that learning from incidents is socially participative, not merely a formal data collection and analysis exercise [4].

Another innovation relates to using the IRS in a different way by focussing reporting on incidents that a health service may require more information about [41]. Marshall et al. focussed reporting on a clinical topic that is not well covered by IRS, paediatric diagnostic incidents, with the added aim of increasing reporting from doctors, who generally did not use the IRS at their institution [41]. Using small-scale iterative interventions, 44 paediatric diagnostic incidents were reported in 6 months from a baseline of 0. This was sufficient to characterize the main contributing factors to these incidents to allow interventions to be designed [41].

The manual and time-consuming nature of case note review and on-going digitalization of medical records has sparked continued research interest in collecting triggers electronically, thereby introducing considerable efficiencies [42]. Within this realm, querying of large electronic data repositories can be undertaken to detect incident types which are infrequent and difficult to collect, such as diagnostic incidents [43]. Artificial intelligence approaches, such as Natural Language Processes can also be incorporated into these data repository querying approaches to refine and make searches more specific [44]. These methods may only be applicable in health services with large volumes of digitized medical record information and the requisite data analyst capabilities [39, 43].

Only two studies included in this review collected information on serious AEs—with 0% and 53% of serious AEs detected using the GTT that were also detected in corresponding IRSs. Both of these studies collected data in specialty areas (medication [17] and inpatient psychiatry [31]) with the latter an outlier in the general results in this review (37%). The relatively high proportion of moderate to severe harm AEs detected by IRS in the inpatient psychiatry study is unlikely to be generalizable [31]. Another study compared GTT and IRS AE rates (but was not included in this systematic review due to GTT and IRR AEs being reported independently) and found in a sample of 795 medical records in three US hospitals, 26 serious AEs detected by the GTT with only two or 8% detected by an IRS [7]. This limited evidence indicates that serious AEs may not be detected reliably by IRS, further emphasising the need for health services to routinely use multiple patient safety data sources.

Implications for policy, practice, and research

Our findings of a >10-fold gap between the AE detection rates of the GTT and IRSs is strong evidence that the rate of AEs collected in IRSs in hospitals should not be used to measure or serve as a proxy for the level of safety of a hospital. However, recent prominent editorials and reviews of international patient safety expert opinions note that such a practice remains common in healthcare and that IRSs are the most widely employed patient safety practice [5, 45, 46]. The primary implication is for health services to incorporate

Downloaded from https://academic.oup.com/intqhc/article/35/3/mzad056/7223853 by guest on 07 August 2023

multiple methods to collect patient safety information in a way that is most efficient for them depending, for example, on whether their records are digitized and to use frameworks of decision making and prioritization setting for action. This should explicitly delineate the purpose of all patient safety data sources in policy, practice, education, and measuring results from interventions designed to reduce harm. For example, IRSs, as emphasized by our results, are poor at assessing incident rates. However, from the perspective of patient safety improvement, they can detect rare and new incident types and contribute to analysis of contributing and contextual factors to develop preventive and corrective strategies. If a health service is lacking information to understand particular incident types or clinical specialties, they may design bespoke IRS data collections to fill this gap and to design interventions, as Marshall et al. achieved in relation to paediatric diagnostic incidents [41].

The results of our study may provide a temptation for health services to run campaigns with clinicians for more incidents to be reported in the IRS. However, we would caution against this. Even though IRS detect a small percentage of AEs, large numbers of incidents are being collected, for example, over 2.4 million in one year (July 2021–June 2022) in the former National Reporting and Learning System in England and Wales [47]. The qualitative data contained in the IRS incident narratives are a highly valuable source of information to understand contributing and contextual factors, to inform improvement and research [48], notwithstanding their known limitations [4]. There are already too many resources devoted to the collection of IRS data, and not enough dedicated to the strategic prioritization, interpretation, and analysis of all patient safety data sources and the implementation of corrective strategies [4].

Conclusions

This systematic review found that in 14 studies, across a range of specialties, IRSs detected <10% of AEs than were detected by the GTT. This study provides clear evidence that hospitals should not use IRSs to estimate prevalence of harm to patients. Health systems should incorporate multiple patient safety data sources to prioritize interventions and promote a cycle of action and improvement based on data rather than merely just collecting and analysing information.

Author contributions

Peter D. Hibbert (initiated and led the project, extracted and analysed the data, undertook the first drafting of the manuscript, and reviewed and signed-off on the manuscript revisions), Charlotte J. Molloy (extracted and analysed the data and undertook the first drafting of the manuscript), and Timothy J. Schultz, Andrew Carson-Stevens, and Jeffrey Braithwaite (interpreted results, revised the manuscript critically for important intellectual content, and informed the discussion).

Supplementary data

Supplementary data is available at *INTQHC* online.

Funding

This work was not supported by any specific funding sources.

Data availability

As this is a systematic review, all data are incorporated into the article and its online supplementary material.

Ethics and other permissions

Not applicable as study was a systematic review with peer-reviewed publications as the source.

References

1. Forster AJ, Dervin G, Martin C *et al.* Improving patient safety through the systematic evaluation of patient outcomes. *Can J Surg* 2012;**55**:418–25. https://doi.org/10.1503/cjs.007811.
2. Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. *J Gen Intern Med* 2003;**18**:61–7. https://doi.org/10.1046/j.1525-1497.2003.20147.x.
3. World Health Organization. *Global Patient Safety Action Plan 2021–2030: Towards Eliminating Avoidable Harm in Health Care*. Geneva: WHO, 2021.
4. Macrae C. The problem with incident reporting. *BMJ Qual Saf* 2016;**25**:71. https://doi.org/10.1136/bmjqs-2015-004732.
5. Shojania KG. Incident reporting systems: what will it take to make them less frustrating and achieve anything useful? *Jt Comm J Qual Patient Saf* 2021;**47**:755–8. https://doi.org/10.1016/j.jcjq.2021.10.001.
6. Graber ML. The incidence of diagnostic error in medicine. *BMJ Qual Saf* 2013;**22**:ii21–7. https://doi.org/10.1136/bmjqs-2012-001615.
7. Classen DC, Resar R, Griffin F *et al.* 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff (Millwood)* 2011;**30**:581–9. https://doi.org/10.1377/hlthaff.2011.0190.
8. Griffin FA, Resar RK. *IHI Global Trigger Tool for Measuring Adverse Events*. 2nd edn. Cambridge, MA: Institute for Healthcare Improvement, 2009.
9. Resar RK, Rozich JD, Classen D. Methodology and rationale for the measurement of harm with trigger tools. *Qual Saf Health Care* 2003;**12**:ii39–45. https://doi.org/10.1136/qhc.12.suppl_2.ii39.
10. Nilsson L, Pihl A, Tågsjõ M *et al.* Adverse events are common on the intensive care unit: Results from a structured record review. *Acta Anaesthesiol Scand* 2012;**56**:959–65. https://doi.org/10.1111/j.1399-6576.2012.02711.x.
11. Gerber A, Da Silva Lopes A, Szüts N *et al.* Describing adverse events in Swiss hospitalized oncology patients using the global trigger tool. *Health Sci Rep* 2020;**3**:e160. https://doi.org/10.1002/hsr2.160.
12. Brösterhaus M, Hammer A, Kalina S *et al.* Applying the global trigger tool in German hospitals: a pilot in surgery and neurosurgery. *J Patient Saf* 2020;**16**:e340–e51. https://doi.org/10.1097/PTS.0000000000000576.
13. Hibbert PD, Runciman WB, Carson-Stevens A *et al.* Characterising the types of paediatric adverse events detected by the global trigger tool – CareTrack Kids. *J Patient Saf Risk Manag* 2020;**25**:239–49. https://doi.org/10.1177/2516043520969329.
14. Hibbert P, Williams H. The use of a glober trigger tool to inform quality and safety in Australian general practice: a pilot study. *Aust Fam Physician* 2014;**43**:723–6.
15. de Wet C, Bowie P. The preliminary development and testing of a global trigger tool to detect error and patient harm in primary-care records. *Postgrad Med J* 1002;**85**: 176–80. https://doi.org/10.1136/pgmj.2008.075788.
16. Shojania KG. The elephant of patient safety: what you see depends on how you look. *Jt Comm J Qual Patient Saf* 2010;**36**:399–401. https://doi.org/10.1016/s1553-7250(10)36058-2.

Downloaded from https://academic.oup.com/intqhc/article/35/3/mzad056/7223853 by guest on 07 August 2023

17. Tchijevitch OA, Nielsen LP, Lisby M. Life-threatening and fatal adverse drug events in a Danish university hospital. *J Patient Saf* 2021;**17**:e562–7. https://doi.org/10.1097/PTS.0000000000000411.
18. Hibbert PD, Molloy CJ, Hooper TD *et al.* The application of the Global Trigger Tool: a systematic review. *Int J Qual Health Care* 2016; **28**:640–9. https://doi.org/10.1093/intqhc/mzw115.
19. Moher D, Liberati A, Tetzlaff J *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;**339**:b2535. https://doi.org/10.1136/bmj.b2535.
20. Sirriyeh R, Lawton R, Gardner P *et al.* Reviewing studies with diverse designs: the development and evaluation of a new tool. *J Eval Clin Pract* 2012;**18**:746–52. https://doi.org/10.1111/j.1365-2753.2011.01662.x.
21. Agarwal S, Classen D, Larsen G *et al.* Prevalence of adverse events in pediatric intensive care units in the United States. *Pediatr Crit Care Medand Crit Care Societies* 2010;**11**:568–78. https://doi.org/10.1097/PCC.0b013e3181d8e405.
22. Hooper AJ, Tibballs J. Comparison of a trigger tool and voluntary reporting to identify adverse events in a paediatric intensive care unit. *Anaesth Intensive Care* 2014;**42**:199–206. https://doi.org/10.1177/0310057X1404200206.
23. Kennerly DA, Kudyakov R, da Graca B *et al.* Characterization of adverse events detected in a large health care delivery system using an enhanced Global Trigger Tool over a five-year interval. *Health Serv Res* 2013;**49**:1407–25. https://doi.org/10.1111/1475-6773.12163.
24. Mull HJ, Brennan CW, Folkes T *et al.* Identifying previously undetected harm: piloting the institute for healthcare improvement's global trigger tool in the veterans health administration. *Qual Manag Health Care* 2015;**24**:140–6. https://doi.org/10.1097/QMH.0000000000000060.
25. Najjar S, Hamdan M, Euwema MC *et al.* The Global Trigger Tool shows that one out of seven patients suffers harm in Palestinian hospitals: challenges for launching a strategic safety plan. *Int J Qual Health Care* 2013;**25**:640–7. https://doi.org/10.1093/intqhc/mzt066.
26. Rutberg H, Risberg MB, Sjödahl R *et al.* Characterisations of adverse events detected in a university hospital: a 4 year study using the Global Trigger Tool method. *BMJ Open* 2014;**4**:e004879.https://doi.org/10.1136/bmjopen-2014-004879.
27. Sharek PJ, Horbar JD, Mason W *et al.* Adverse events in the neonatal intensive care unit: development, testing, and findings of an NICU-focused trigger tool to identify harm in North American NICUs. *Pediatrics* 2006;**118**:1332–40. https://doi.org/10.1542/peds.2006-0565.
28. Stockwell DC, Bisarya H, Classen DC *et al.* A trigger tool to detect harm in pediatric inpatient settings. *Pediatrics* 2015;**135**:1036–42. https://doi.org/10.1542/peds.2014-2152.
29. Takata GS, Mason W, Taketomo C *et al.* Development, testing, and findings of a pediatric-focused trigger tool to identify medication-related harm in US children's hospitals. *Pediatrics* 2008;**121**:e927–35. https://doi.org/10.1542/peds.2007-1779.
30. Naessens JM, Campbell CR, Huddleston JM *et al.* A comparison of hospital adverse events identified by three widely used detection methods. *Int J Qual Health Care* 2009;**21**:301–7. https://doi.org/10.1093/intqhc/mzp027.
31. Reilly CA, Cullen SW, Watts BV *et al.* How well do incident reporting systems work on inpatient psychiatric units? *Jt Comm J Qual Patient Saf* 2019;**45**:63–9. https://doi.org/10.1016/j.jcjq.2018.05.002.
32. Samal L, Khasnabish S, Foskett C *et al.* Comparison of a voluntary safety reporting system to a global trigger tool for identifying adverse events in an oncology population. *J Patient Saf* 2022;**18**:611–6. https://doi.org/10.1097/PTS.0000000000001050.
33. Kohn LT, Corrigan JM, Donaldson MS. *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academies Press (US), 2000.
34. National coordination council for medication error reporting and prevention. NCC MERP index for categorizing medication errors, 2001.
35. Madden C, Lydon S, Curran C *et al.* Potential value of patient record review to assess and improve patient safety in general practice: a systematic review. *Eur J Gen Pract* 2018;**24**:192–201.https://doi.org/10.1080/13814788.2018.1491963.
36. Lydon S, Cupples ME, Murphy AW *et al.* A systematic review of measurement tools for the proactive assessment of patient safety in general practice. *J Patient Saf* 2021;**17**:e406–e12. https://doi.org/10.1097/PTS.0000000000000350.
37. Westbrook JI, Li L, Lehnbom EC *et al.* What are incident reports telling us? A comparative study at two Australian hospitals of medication errors identified at audit, detected by staff and reported to an incident system. *Int J Qual Health Care* 2015;**27**:1–9. https://doi.org/10.1093/intqhc/mzu098.
38. Levtzion-Korach O, Frankel A, Alcalai H *et al.* Integrating incident data from five reporting systems to assess patient safety: making sense of the elephant. *Jt Comm J Qual Patient Saf* 2010; **36**:402–10. https://doi.org/10.1016/s1553-7250(10)36059-4.
39. Rosen AK. *Are We Getting Better at Measuring Patient Safety?*. Rockville, MD: AHRQ, 2010.
40. Wong BM, Dyal S, Etchells EE *et al.* Application of a trigger tool in near real time to inform quality improvement activities: a prospective study in a general medicine ward. *BMJ Qual Saf* 2015;**24**:272–81. https://doi.org/10.1136/bmjqs-2014-003432.
41. Marshall TL, Ipsaro AJ, Le M *et al.* Increasing physician reporting of diagnostic learning opportunities. *Pediatrics* 2021;**147**:e20192400. https://doi.org/10.1542/peds.2019-2400.
42. Musy SN, Ausserhofer D, Schwendimann R *et al.* Trigger tool-based automated adverse event detection in electronic health records: systematic review. *J Med Internet Res* 2018;**20**:e198. https://doi.org/10.2196/jmir.9901.
43. Bhise V, Sittig DF, Vaghani V *et al.* An electronic trigger based on care escalation to identify preventable adverse events in hospitalised patients. *BMJ Qual Saf* 2018;**27**:241–6. https://doi.org/10.1136/bmjqs-2017-006975.
44. Murphy DR, Meyer AN, Sittig DF *et al.* Application of electronic trigger tools to identify targets for improving diagnostic safety. *BMJ Qual Saf* 2019;**28**:151–9. https://doi.org/10.1136/bmjqs-2018-008086.
45. Mitchell I, Schuster A, Smith K *et al.* Patient safety incident reporting: a qualitative study of thoughts and perceptions of experts 15 years after 'to err is human'. *BMJ Qual Saf* 2016;**25**:92–9. https://doi.org/10.1136/bmjqs-2015-004405.
46. Howell AM, Burns EM, Hull L *et al.* International recommendations for national patient safety incident reporting systems: an expert Delphi consensus-building process. *BMJ Qual Saf* 2017;**26**:150–63. https://doi.org/10.1136/bmjqs-2015-004456.
47. NHS England. *National patient safety incident report up to June 2022*. England: NHS England. 2022. https://www.england.nhs.uk/publication/national-patient-safety-incident-reports-up-to-june-2022/ (12 December 2022, date last accessed).
48. Carson-Stevens A, Hibbert P, Williams H *et al.* Characterising the nature of primary care patient safety incident reports in the England and Wales National Reporting and Learning System: a mixed-methods agenda-setting study for general practice. *Health Serv Res* 2016;**4**:1–76. https://doi.org/10.3310/hsdr04270.

Downloaded from https://academic.oup.com/intqhc/article/35/3/mzad056/7223853 by guest on 07 August 2023