RESEARCH ARTICLE

REVISED Assessing the development and implementation of the Global Trigger Tool method across a large health system in Sicily [version 2; peer review: 1 approved, 1 approved with reservations]

Vincenzo Parrinello[1], Elena Grasso[1], Giuseppe Saglimbeni[1], Gabriella Patanè[1], Alma Scalia[1], Giuseppe Murolo[2], Peter Lachman [3]

[1]U.O. Qualità e Rischio Clinico, Azienda Ospedaliero-Universitaria "Policlinico-Vittorio Emanuele, Catania, 95129, Italy
[2]Servizio 8 "Qualità, Governo Clinico e Sicurezza del Paziente", Assessorato della Salute, Regione Siciliana, Palermo, 90145, Italy
[3]International Society for Quality in Healthcare, Dublin, D02NY63, Ireland

v2 **First published:** 07 Mar 2019, **8**:263 (https://doi.org/10.12688/f1000research.18025.1)
Second version: 11 Sep 2019, **8**:263 (https://doi.org/10.12688/f1000research.18025.2)
Latest published: 09 Oct 2019, **8**:263 (https://doi.org/10.12688/f1000research.18025.3)

Abstract

Background: The Institute for Healthcare Improvement (IHI) has proposed a new method, the Global Trigger Tool (IHI GTT), to detect and monitor adverse events (AEs) and provide information to implement improvement. In 2015, the Sicilian Health System adopted IHI GTT to assess the number, types and severity levels of AEs. The GTT was implemented in 44 of 73 Sicilian public hospitals and 18,008 clinical records (CRs) were examined. Here we present the standardized application of the GTT and the preliminary results of 14,706 reviews of CRs.
Methods: IHI GTT was adapted, developed and implemented to the local context. Reviews of CRs were conducted by 199 professionals divided into 71 review teams consisting of three individuals: two of whom had clinical knowledge and expertise, and a physician to authenticate the AE. The reviewers entered data into a dedicated IT-platform. All 44 of the public hospitals were included, with approximately 300,000 yearly inpatient admissions out of a population of approximately 5 million. In total, 14,706 randomized CRs of inpatients from medicine, surgery, obstetric and ICU wards, from June 2015 to June 2018 were reviewed.
Results: In 975 (6.6%) CRs at least one AE was found. Approximately 20,000 patients of the 300,000 discharged each year in Sicily have at least one AE. In 5,574 (37.9%) CRs at least one trigger was found. A total of 1,542 AEs were found. The analysis of ROC curve shows that the presence of two triggers in a CR indicates with high probability the presence of an AE. The most frequent type of AE was in-hospital related infection.
Conclusions: The GTT is an efficient method to identify AEs and to track improvement of care. The analysis and monitoring of some triggers is important to prevent AEs. However, it is a labor-intensive method, particularly if the CRs are paper-based.

Open Peer Review

Reviewer Status ✔ ?

	Invited Reviewers	
	1	**2**
REVISED **version 3** published 09 Oct 2019		
REVISED **version 2** published 11 Sep 2019	✔ report	
version 1 published 07 Mar 2019	? report	? report

1 **James M. Naessens**, Mayo Clinic, Rochester, USA

2 **Brent C. James**, Stanford University School of Medicine, Stanford, USA

Any reports and responses or comments on the article can be found at the end of the article.

Keywords
Global Trigger Tool, patient safety, adverse events detections, quality of care, medical errors, harm

Corresponding author: Vincenzo Parrinello (parrinello.qualita@gmail.com)

Author roles: Parrinello V: Conceptualization, Investigation, Methodology, Project Administration, Supervision, Writing – Original Draft Preparation; **Grasso E**: Data Curation, Formal Analysis, Methodology, Validation; **Saglimbeni G**: Conceptualization; **Patanè G**: Conceptualization; **Scalia A**: Conceptualization, Writing – Review & Editing; **Murolo G**: Funding Acquisition; **Lachman P**: Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work is supported by Italian Healthcare Ministry and Sicilian Healthcare Ministry (PSN1316.3).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Parrinello V, Grasso E, Saglimbeni G *et al.* **Assessing the development and implementation of the Global Trigger Tool method across a large health system in Sicily [version 2; peer review: 1 approved, 1 approved with reservations]** F1000Research 2019, **8**:263 (https://doi.org/10.12688/f1000research.18025.2)

First published: 07 Mar 2019, **8**:263 (https://doi.org/10.12688/f1000research.18025.1)

REVISED **Amendments from Version 1**

Our review consisted in a revision of contents and typos indicated by the reviewers. We tried to clear some aspects of the research that seemed to be unclear.

We have clarified the characteristics of the CRs sample examined and the CRS selection methods. We have modified Table 1 and Table 2 and removed Figure 2, which generated confusion. Figure 1 has also been modified for clarity.

We also tried to clarify the reason why the number of some triggers is less than the number of the identified AEs. In the CRs with one trigger like C11, C04, C14, C03, C08, M04, I01, M02, I02, C12 and M11 has been found more than one AEs. We have better explained the role of the reviewers and in particular the role of the physician, the third reviewer. We have tried to briefly clarify the meaning of triggers without AE, attributing them the near miss function.

We have corrected numerous translation errors.

Any further responses from the reviewers can be found at the end of the article

Introduction

Safety is one of the domains of quality in healthcare. Improving the safety of patients is a political priority worldwide, as studies on the safety of patients have drawn attention to the high rates of health care-related harm[1–3]. Improving patient safety requires effective and reliable methods to identify and monitor adverse events (AEs) so that learning can take place and improvements can be made.

Even though several methods to detect AEs are available, there is no universally recognized method that reliably provides a comprehensive overview of the extent of the problem. These methods include incident reporting, clinical records (CRs) review and automated extraction using hospital administrative data, for example Patient Safety Indicators (PSI) as developed by the Agency for Healthcare Research and Quality (AHRQ). Incident reporting is the most commonly used method to detect AEs in hospitals, but is based on voluntary reporting. Despite considerable efforts by local hospitals, reporting systems only detect a limited number of AEs[4]. The effectiveness of automated extraction using hospital administrative data for detecting AEs depends on the accuracy of data compilation. CRs review, as used in the Harvard Medical Practice Study, is very labor intensive, thereby limiting its use[4]. As a result, health services, governments and researchers have focused on developing harm detection tools.

This paper is one of the first to report the findings of the application of the Global Trigger Tool, developed by the Institute for Healthcare Improvement (IHI), across the whole health system in Sicily. In 2015, the Sicilian Health System adopted IHI GTT[5] to assess the number, types and severity levels of AEs.

Methods

IHI GTT

The measurement instrument used in our study is the Italian version of the IHI GTT[6]. The Italian version was adapted to be appropriate to the regional context[7]. The triggers are grouped into the same seven categories as the original version (care, medication, surgical, intensive care, obstetric, pediatrics and emergency care); however, changes to some triggers have been introduced. We did not consider triggers and AEs that were present on admission and we added three new triggers: change in procedure anesthesia, duration of surgery greater than 6 hours, and hospital stay greater than five days after delivery[7].

Sample selection

89 of medicine, surgery, obstetrics and intensive care wards have participated in the study. Some of these have participated from June 2015 to June 2018, while others did not continue to the end of the observation period. Others were added during the study. Some departments participated in the study from June 2015 to June 2018, while others did not continue until the end of the observation period. Others were added during the study.

We used a random method of records selection, as recommended by the IHI GTT. In each ward, we selected 10 inpatient CRs randomly selected during a month, one for every ten. (Table 1) From the ordered sequence of the numbering of the CRs of the period under evaluation, a CR was selected every 10 (ie 10th, 20th, 30th, 40th, 50th, etc). In the case in which the number of patients discharged during that month was less than 100, we proceeded to remove the previously selected CRs and to select another CR in the same way (one CR every 10). Eligibility criteria were an admission lasting more than 24 hours, and all the administrative data completed. In the Intensive Care Unit (ICU), all patient CRs, discharged during the reference period, were reviewed.

Review team

As per the IHI protocol, each review team was composed of three individuals: two with clinical knowledge and expertise on patient clinical documentation, and a physician whose role was to authenticate the findings and the severity rating of the AEs.

The total number of the reviewers was 199 divided into a 71 team. Where possible, the review team remained consistent over time.

Review process

We excluded triggers and/or events that took place outside the time of the patient admission to the hospital and we considered only triggers and AEs that occurred during hospitalization. The two clinical reviewers audited all the CRs on their own, independently. We used five worksheets: general care, medication, surgical, obstetric and intensive care, with some changes in accordance with the IHI GTT. The third reviewer was always a physicians. The physician who did not reviewed the records, but had to authenticate the consensus of the two primary record reviewers. The physician authenticated the findings of the adverse events and the rating of severity, and provided answers to questions of the record reviewers have about findings in a specific record.

We have used the IHI GTT definition of an adverse event: *unintended physical injury resulting from or contributed to by medical care that requires additional monitoring, treatment or hospitalization, or that results in death*.

Triggers and AEs present at the time of hospitalization were excluded in this study.

Table 1. Distribution of CRs per clinical area.

	Medicine	Surgery	Obstetric	ICU	Total
CRs examined, n (%)	4571 (31.1)	4826 (32.8)	3336 (22.7)	1973 (13.4)	14706 (100)
CRs with triggers per CRs examined, n (%)	1571 (34.3)	1709 (35.4)	676 (20.2)	1672 (84.7)	5574 (37.9)
CRs with isolated trigger CRs examined, n (%)	930 (20.3)	1085 (22.5)	491 (14.7)	272 (13.7)	2778 (18.9)
CRs with AEs, n (%)	191 (19.5)	128 (14.2)	57 (5.8)	599 (61.5)	975 (100)
AEs, n (%)	210 (13.5)	138 (9.0)	61 (3.9)	1133 (73.5)	1542 (100)
CRs with AEs/CRs examined, (%)	4.2	2.7	1.7	30.4	6.6
CRs with AEs/CRs with triggers, (%)	11.1	8.4	8.4	35.9	17.5

CRs, clinical records; ICU, intensive care unit; AES, adverse events

The CRs were examined following the order of the sections described in the IHI GTT. The revison time should have been no longer than 20 minutes. The "20-minute rule" was applied to all records regardless their size[5]. The reviewers entered data into a specially developed dedicated IT-platform, developed by our IT team (based on Jawascript HTML and PHP)[8]

Statistical analysis

For the statistical analysis we used the software SPSS ver. 20. We used it also to develop the Receiver Operating Characteristic (ROC) curve analysis.

Results

From June 2015 to June 2018, 18,008 CRs from 105 wards of 44 Sicilian public hospitals were examined. In this study, we analyzed just 14,706 CRs relating to patients discharged from 89 medicine, surgery, obstetrics and intensive care wards, without including the CRs of the emergency and pediatric wards. In 5,574 (37.9%) CRs at least one trigger was found. In 7 CRs, the reading of the discharge diagnosis aroused interest by reviewers and an AE was detected and the triggers were not looked for. AEs were determined in 1,542 CRs (Table 1). The identification of triggers allowed us to identify corresponding AEs (Table 2).

In this study, 37.9% (n=5,574) of all CRs examined had at least 1 positive trigger. Of those, 2,778 CRs had a single positive trigger (49.8% of all CRs with positive triggers) while 2,796 CRs had more than one positive trigger (51.2% of all CRs with positive triggers).

CRs with triggers (n=5,574) are significantly present in surgery wards (n=1,709; 35.4%), medicine wards (n=1,517; 34.3%) and ICU (n=1,672; 84.7%). while CRs with triggers in obstetrics wards are significantly less frequent (n=676; 20.3%). CRs with isolated triggers are more common in medical wards (n=1,085; 23.7%) and rarer in ICU (n=272; 13.7%) (Table 1).

We detected 975 CRs (i.e patients) with AEs in 1,542 CRs (i.e patients) with at least 1 positive trigger. In 652 patients (66.8%) a single AE was present. In the remaining 323 (33.2%) were more than one AEs.

This analysis allowed us to highlight how isolated triggers are not always a good indicator of AEs. A Receiver Operating Characteristic (ROC) curve analysis demonstrates that the presence of two triggers in a CR has a high probability that an AE having occurred (Figure 1). In CRs with a high frequency of triggers, a corresponding number of AEs was not always detected. As indicated in Table 3, on the contrary, some triggers were associated with a large number of AEs. For example, the trigger C11 Health care–associated Infection was detected in 504 CRs. However, often, a single CR with trigger C11, presented more than one AE.

Triggers and AEs were analyzed when isolated triggers were identified (Table 4). For example, the isolated trigger C01 (Blood products use) was present in 483 cases, but identified only two AEs, and the trigger M05 (Rising BUN or serum creatinine >2 times the baseline) did not identify any AEs. AEs were classified using the 2009 edition of the WHO International Classification for Patient Safety (ICPS)[9], and a clinical classification developed by our group (Table 5).

The most frequent type of AEs observed: in-hospital related infections; surgical complications; pressure ulcers; acute kidney injury; and procedure complications.

Discussion

The evaluation of the quality and safety of health systems is difficult, but has become a priority of healthcare funders and organizations. Outcome, management and patient satisfaction indicators are available to measure the different dimensions of health care quality, but reliable measurements of safety have been elusive. Many methodologies and indicators, such as the PSI developed by AHRQ, the review of health documentation incident reporting and prospective clinical surveillance methodology are currently used.

The documentation and study of AEs, i.e. where they occur, and the type and degree of harm, is essential to promote specific opportunities for interventions improvement and to evaluate effectiveness of any intervention over time.

Table 2. Distribution of triggers and AEs.

Trigger	Description	Number of times detected, n (%)	Number of times associated, with AEs, n (%)	Trigger	Description	Number of times detected, n (%)	Number of times associated, with AEs, n (%)
C01	Blood products use	2002 (26.7)	958 (15.7)	**S01**	Return to surgery	64 (20)	47 (31.8)
C02	Emergency and rescue	719 (9.6)	617 (10.1)	**S02-A**	Change in procedure: surgery	57 (18)	8 (5.4)
C03	Acute dialysis	279 (3.7)	438 (7.2)	**S02-B**	Change in procedure: anesthesia	6 (1.8)	0 (0)
C04	Positive blood culture	291 (3.9)	485 (7.9)	**S03**	Admission to ICU	31 (9.6)	13 (8.8)
C05	X-ray or Doppler studies for emboli or DVT	435 (5.8)	112 (1.8)	**S04**	Intubation/reintubation/BiPap in PACU	10 (3.1)	7 (4.7)
C06	Decrease of Hb or Ht >25%	1433 (19.1)	896 (14.7)	**S05**	X-ray intraoperative or in PACU	4 (1.2)	0 (0)
C07	Patient fall	29 (0.4)	20 (0.3)	**S06**	Intraoperative or postoperative death	9 (2.8)	2 (1.4)
C08	Pressure ulcers	254 (3.4)	449 (7.4)	**S07**	Mechanical ventilation >24 hours post-op	18 (5.6)	9 (6.1)
C09	Readmission within 30 days	294 (3.9)	115 (1.9)	**S08**	Intraoperative epinephrine, norepinephrine, naloxone, or flumazenil	7 (2.2)	2 (1.4)
C10	Restraint use	260 (3.5)	38 (0.6)	**S09**	Postoperative troponin level >1.5 ng/mL	19 (5.9)	3 (2.0)
C11	Health care–associated infection	504 (6.7)	894 (14.6)	**S10**	Injury, repair, or removal of organ	13 (4.0)	9 (6.1)
C12	In-hospital stroke	35 (0.5)	78 (1.3)	**S11**	Any operative complication	60 (19)	41 (27.7)
C13	Transfer to higher level of care	678 (9.0)	595 (9.7)	**S12**	Duration of surgery > 6h	25 (7.7)	7 (4.7)
C14	Any procedure complication	284 (3.8)	408 (6.7)		**TOTAL**	**323 (100)**	**148 (100)**
	TOTAL	**7.497 (100)**	**6.103 (100)**	**P01**	3rd- or 4th-degree lacerations	13 (2.6)	7(10.9)
M01	Clostridium difficile–positive stool	28 (1.0)	21 (1.4)	**P02**	Platelet count less than 50,000	2 (0.4)	0 (0)
M02	PTT >100 seconds	99 (3.4)	122 (8.0)	**P03**	Estimated blood loss >500 mL (vaginal) or >1000 mL (C-section)	27 (5.5)	13 (20.3)
M03	INR > 6	48 (1.7)	33 (2.2)	**P04**	Specialty consult	96 (19.5)	11 (17.2)
M04	Glucose < 50 mg/dl	225 (7.8)	238 (15.6)	**P05**	Administrate prostaglandins postpartum	97 (19.7)	8 (12.5)
M05	Rising BUN or serum creatinine >2 times baseline	1194 (41.5)	718 (47.1)	**P06**	Instrumented delivery	101 (20.5)	7 (10.9)
M06	Vitamin K administration	231 (8.0)	155 (10.2)	**P07**	General anesthesia	75 (15.2)	10 (15.6)
M07	Anti-allergic use	145 (5.0)	49 (3.2)	**P08**	Hospital stay> more than 5 days	81 (16.5)	8(12.5)
M08	Flumazenil use	37 (1.3)	24 (1.6)		**TOTAL**	**492 (100)**	**64 (100)**
M09	Naloxone use	7 (0.2)	6 (0.4)	**I01**	Pneumonia onset	193 (10.6)	380 (19.8)
M10	Anti-emetic use	843 (29.3)	133 (8.7)	**I02**	Readmission ICU	59 (3.2)	125 (6.5)
M11	Over-sedation	21 (0.7)	26 (1.7)	**I03**	In-unit procedure	761 (41.9)	679 (35.3)
	TOTAL	**2,878 (100)**	**1,525 (100)**	**I04**	Intubation/reintubation	804 (44.2)	740 (38.5)
					TOTAL	**1,817 (100)**	**1,924 (100)**

AEs, adverse events.

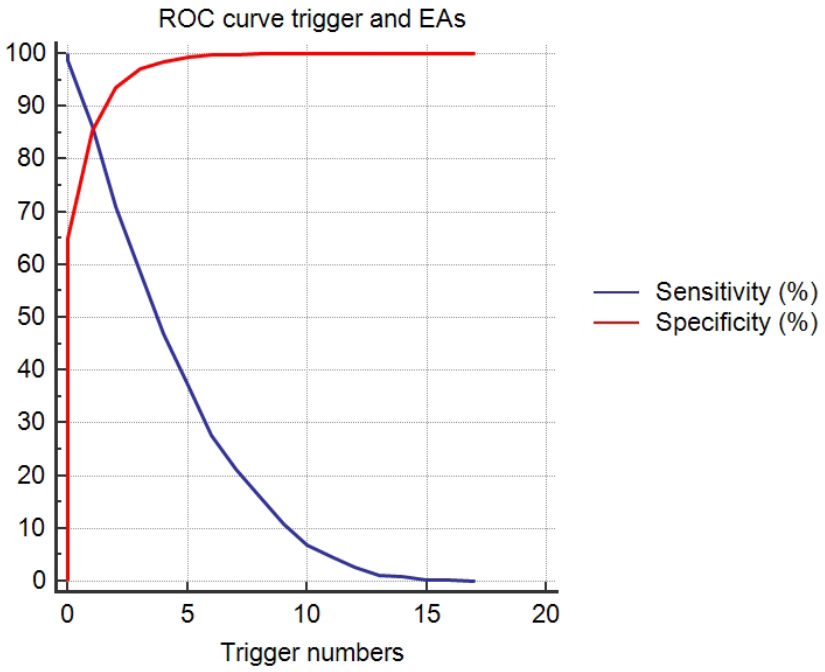

Figure 1. ROC analysis of two random triggers. ROC curve shows that the presence of two triggers in clinical records indicates an adverse event with a high probability.

Table 3. Distribution of triggers and AEs.

Trigger		Triggers with AEs (n)	AEs with trigger (n)
C11	Health care–associated Infection	504	894
C04	Positive blood culture	291	485
C14	Any procedure complication	284	408
C03	Acute dialysis	279	438
C08	Pressure ulcers	254	449
M04	Glucose < 50 mg/dl	225	238
I01	Pneumonia onset	193	380
M02	PTT >100 seconds	99	122
I02	Readmission ICU	59	125
C12	In-hospital stroke	35	78
M11	Over-sedation	21	26

AEs, adverse events.

The IHI GTT is one methodology proposed to detect and monitor AEs and provide information to implement improvement. At present, compared to other methods, it may be the best methodology to use[4]. A systematic review reported the use of GTT methodology in 15 countries in 44 hospitals, with 79,004 clinical records examined[10]. The data are an underestimation, as the report did not include some comprehensive Swedish and Norwegian studies[11]. Recently, papers have been published in Italy, Austria, China and Russia[12–16]. A critical appraisal of the studies and their results is difficult, as the methodologies use are heterogeneous, protocols are often locally adapted to the local context, the populations studied are different, and the skills of the reviewers vary. We adapted the IHI GTT to the local context in Sicily for this study and did not consider triggers and AEs identified at the

Table 4. Distribution of isolated triggers and AEs.

Trigger	Description	Number of times detected isolated, n (%)	Number of times associated with AEs, n (%)	Trigger	Description	Number of times detected isolated, n (%)	Number of times associated with AEs, n (%)
C01	Blood products use	483 (33.7)	2 (2.5)	**S01**	Return to surgery	12 (15.0)	5 (41.7)
C02	Emergency and rescue	91 (6.4)	4 (5.0)	**S02-A**	Change in procedure: surgery	32 (40.0)	1 (8.3)
C03	Acute dialysis	6 (0.4)	4 (5.0)	**S02-B**	Change in procedure: anesthesia	3 (3.8)	0 (0)
C04	Positive blood culture	26 (1.8)	11 (13.8)	**S03**	Admission to ICU	5 (6.3)	0 (0)
C05	X-ray or Doppler studies for emboli or DVT	186 (13.0)	0 (0)	**S04**	Intubation/reintubation/BiPap in PACU	1 (1.3)	0 (0)
C06	Decrease of Hb or Ht >25%	190 (13.3)	0 (0)	**S05**	X-ray intraoperative or in PACU	2 (2.5)	0 (0)
C07	Patient fall	9 (0.6)	4 (5.0)	**S06**	Intraoperative or postoperative death	1 (1.3)	0 (0)
C08	Pressure ulcers	23 (1.6)	15 (18.8)	**S07**	Mechanical ventilation >24 hours post-op	1 (1.3)	0 (0)
C09	Readmission within 30 days	114 (8.0)	9 (11.3)	**S08**	Intraoperative epinephrine, norepinephrine, naloxone, or flumazenil	0 (0)	0 (0)
C10	Restraint use	115 (8.0)	2 (2.5)	**S09**	Postoperative troponin level >1.5 ng/mL	2 (2.5)	0 (0)
C11	Health care–associated infection	32 (2.2)	17 (21.3)	**S10**	Injury, repair, or removal of organ	4 (5.0)	1 (0)
C12	In-hospital stroke	3 (0.2)	0 (0)	**S11**	Any operative complication	16 (20.0)	5 (41.7)
C13	Transfer to higher level of care	97 (6.8)	1 (1.3)	**S12**	Duration of surgery > 6h	1 (1.3)	0 (0)
C14	Any procedure complication	58 (4)	11 (13.8)		**TOTAL**	**80 (100)**	**12 (0)**
	TOTAL	**1433 (100)**	**80 (100)**	**P01**	3rd- or 4th-degree lacerations	9 (3.3)	4 (36.4)
M01	Clostridium difficile–positive stool	7 (0.7)	1 (5.6)	**P02**	Platelet count less than 50,000	1 (0.4)	0 (0)
M02	PTT >100 seconds	13 (1.4)	0 (0)	**P03**	Estimated blood loss >500 mL (vaginal) or >1000 mL (C-section)	5 (1.9)	3 (27.3)
M03	INR >6	5 (0.5)	0 (0)	**P04**	Specialty consult	52 (19.3)	0 (0)
M04	Glucose < 50 mg/dl	53 (5.6)	4 (22.2)	**P05**	Administrate prostaglandins postpartum	48 (17.8)	0 (0)
M05	Rising BUN or serum creatinine >2 times baseline	263 (27.7)	0 (0)	**P06**	Instrumented delivery	77 (28.6)	2 (18.2)
M06	Vitamin K administration	31 (3.3)	2 (11.1)	**P07**	General anesthesia	35 (13.0)	2 (18.2)
M07	Anti-allergic use	60 (6.3)	7 (38.9)	**P08**	Hospital stay > 5 days after delivery	42 (15.6)	0 (0)
M08	Flumazenil use	10 (1.1)	0 (0)		**TOTAL**	**269 (100)**	**11 (100)**
M09	Naloxone use	1 (0.1)	0 (0)	**I01**	Pneumonia onset	8 (17.4)	0 (0)
M10	Anti-emetic use	504 (53.1)	4 (22.2)	**I02**	Readmission ICU	1 (2.2)	0 (0)
M11	Over-sedation	3 (0.3)	0 (0)	**I03**	In-unit procedure	15 (32.6)	0 (0)
	TOTAL	**950 (100)**	**18 (100)**	**I04**	Intubation/reintubation	22 (47.8)	1 (100)
					TOTAL	**46(100)**	**1(100)**

AEs, adverse events.

F1000Research 2019, 8:263 Last updated: 09 OCT 2019

Table 5. Categorization of adverse events.

International Classification for Patient Safety (ICPS) - WHO ed. 2009		Clinical classification	
INCIDENT TYPE	**AEs, N (%)**	**INCIDENT TYPE**	**AEs, N (%)**
Healthcare Associated Infection	742 (48.1)	Healthcare Associated Infection	742 (48.1)
Clinical Process/Procedure	697 (45.2)	Surgical complications	175 (11.3)
Medication/IV Fluids	89 (5.7)	Pressure ulcers	172 (11.2)
Patient Accidents	12 (0.1)	Acute kidney injury	133 (8.6)
Blood/Blood Products	2 (0.1)	Procedure complications*	109 (7.1)
TOTAL	1542 (100)	Hypoglycemia	62 (4.0)
		Delivery complications	47 (3.0)
		In-hospital Stroke	18 (1.2)
		Anesthetic complications	17 (1.1)
		Hemorrhage	5 (0.3)
		Various	62 (4.0)
		TOTAL	1542 (100)

AEs, adverse events.

* Endoscopic procedures, central catheterization, urinary catheterization, orotracheal intubation

admission of the patient as well as modifying some triggers. In this study, the triggers were analyzed both when associated with other triggers and when isolated. In both cases the correlation with AEs was analyzed.

Rates of triggers

In this study, 37.9% (n=5,574) of all CRs examined had at least 1 positive trigger. Of those, 2,778 CRs had a single positive trigger (49.8% of all CRs with positive triggers) while 2,796 CRs had more than one positive trigger (51.2% of all CRs with positive triggers).

The connection between the number of CRs with triggers and the number of CRs examined is not always reported in the literature and when reported it is not always clear. Xu *et al.*[17] report that during the review of 240 clinical records, 51.0% triggers (26/51) were identified 206 times. Mortaro *et al.*[12] report that during review of 1,320 clinical records, a total of 130 triggers were detected.

In our study, few AEs were identified by isolated triggers and many isolated triggers were not associated with AEs. Even though these are not very useful for identifying AEs, some triggers may be direct measures of "near misses". Today's trigger could be tomorrow's adverse event.

In general, it would appear that little attention is paid to triggers if they are not related to an AE. Instead, many triggers of the IHI GTT protocol could be considered to be a measure of near misses and potential AEs. These include, decrease of Hb or Ht >25%, readmission within 30 days, transfer to higher level of care, clostridium difficile–positive stool, PTT >100 seconds, INR > 6, glucose < 50 mg/dl, rising BUN or serum creatinine >2 times baseline, blood loss >500 mL (after vaginal delivery) or >1000 mL (Cesarean section), and readmission to ICU.

Rates of AEs

In a systematic review, de Vriess *et al.*[18] reported that in 8 studies that included 74,485 CRs, the median overall incidence of in-hospital AEs was 9.2%. Another systematic review reported 44 hospitals with 79,004 CRs, had an incidence between 7 and 51%[10]. In the Sicilian public hospitals, 1,542 AEs were detected in 975 clinical records, corresponding to an incidence of 6.6% of CRs examined, and to 17.5% of CRs with triggers.

The results of this study are not comparable with other studies due to the diversity of detection protocols. This is also demonstrated by the wide frequency variability of AEs reported in the literature (7–51%)[10].

In this study the most significant difference, compared to the other studies, is the exclusion of triggers and AEs present at the time of the patient admission in the hospital.

ICUs have the highest incidence of AEs, in comparison to CRs examined (30.4%) and those with triggers (35.9%) (Table 1). This could be due to patients being transferred to ICU and the cause of the AE was in another clinical setting. We have analyzed the AEs comparing them to the triggers to allow for their identification:

- Most AEs are associated with general care triggers (from C1 to C14) (n=6,103) (Table 2). If triggers are isolated, AEs are more frequently associated with care triggers (n=80) (Table 4).
- The triggers related to general care have been identified 7,497 times, with an AE in 6,103 (81.4%).
- Medications-related triggers (From M1 to M11) have been identified 2,878 times, with an AE in 1,525 (53%).
- Surgery-related triggers (form S1 to S12) have been identified 323 times, with an AE in 148 (45.8%).
- Obstetrics-related triggers (from P1 to P8) have been identified 492 times, with an AE in 61 (12.4%).
- Intensive-care-related triggers (from I1 to I4) have been identified 1,817 times with an AE in 1,924 (i.e. it is very common for triggers to identify more AEs in the same patient) (Table 1 and Table 2).

However, if isolated triggers are considered, the intensive-care-related triggers were detected 46 times and they were correlated with only one AE (Table 4). This observation suggests that isolated triggers rarely allow to identify an AE and that the strength of the IHI's GTT methodology is linked to the association of triggers. It is evident that the detection of many triggers in a CR is associated with a high probability of AEs. The analysis of the ROC curve (Figure 1) shows that it is sufficient to detect two triggers in a CR because it can be almost certain that in that CR there may be an AE.

We have classified the AEs using the ICPS 2009 classification and a clinical classification, developed by our group. In both classifications, hospital acquired infections are the most frequent AEs present (Table 5), observed in the ICU in 625 clinical records (84.3%). Surgical complications (n=175) were observed in 55.9% (n=99) in ICU, i.e. they were AEs in patients undergoing surgery and then transferred to ICU due to the onset of a complication. In 44.6% (n=80), the AEs are represented by hemorrhagic complications (intra- and post-operative hemorrhages or hematomas). Pressure ulcer lesions were detected in 172 cases, usually in the ICUs (n=112 - 65.1%). Also, the complications from procedures (n=109) were observed mainly in the ICU (n=78; 71.5%). In total, 33 complications from procedures are related to orotracheal intubation, 22 at central venous catheter, and 10 at childbirth analgesia. The complications of child birth (n= 47) were represented more frequently in 59.5% (n=28) by bleeding and in 29.8% (n=14) by lacerations.

Limitations

Our study has some several limitations. The first concerns the inter-rater reliability assessment of review teams that is not available. The second limitation is the underlying quality of CRs, which may have affected the results. However, all reviewers received the same training and each team followed the same protocols to ensure reliability.

Conclusion

The Global Trigger Tool is an effective method to identify AEs and track improvement of care. It provides to clinical teams an understanding of the patient safety issues that are present in their clinical area, as well as opportunities to improve. With active involvement of clinical teams, it places patient safety in the centre of clinical activity and fosters a culture of safety. It also provides an effective way to assess the quality of the clinical records. The drawback is that the process is labor intensive, particularly if the clinical records are paper based. The introduction of electronic medical records would allow a quicker process with the automation of the identification of triggers and the possibility to link triggers together in the identification of adverse events and near misses, especially where there has been more than one trigger[19,20]. Finally, we conclude that the analysis and monitoring of some triggers, as potential indicators of near misses, is important to prevent adverse events. Today's trigger could be tomorrow's adverse event.

Ethical considerations

Since the data used in this study was gathered during routine practice and is used for analysis of hospital procedures, no ethical approval was obtained. Every patient gave written informed consent on admission to hospital for the use of their data for scientific research. This consent is the "Information on the processing of personal data" with reference to the Italian law n. 196/2003 and Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (UE).

Data availability

Underlying data

Harvard Dataverse: Replication data for Developing and implementing the Global Trigger Tool methodology across a large health system in Sicily, https://doi.org/10.7910/DVN/YQNKCC[21]

Acknowledgments

It was possible to realize this project thanks to the involvement and passion of the 199 reviewers, doctors, nurses and obstetrics, the cooperation of quality and patients safety managers of the Sicilian public hospitals, Antonino Drago and Rosario Raineri for their support on data analysis.

References

1. Leape LL, Brennan TA, Laird N, *et al.*: **The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II.** *N Engl J Med.* 1991; **324**(6): 377–384.
PubMed Abstract | Publisher Full Text

2. Localio AR, Lawthers AG, Brennan TA, *et al.*: **Relation between malpractice claims and adverse events due to negligence. Results of the Harvard Medical Practice Study III.** *N Engl J Med.* 1991; **325**(4): 245–251.
PubMed Abstract | Publisher Full Text

3. Brennan TA, Leape LL, Laird NM, *et al.*: **Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I.** *N Engl J Med.* 1991; **324**(6): 370–376.
PubMed Abstract | Publisher Full Text

4. 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**(4): 581–9.
PubMed Abstract | Publisher Full Text

5. Griffin FA, Resar RK: **IHI Global Trigger Tool for Measuring Adverse Events.** 2nd ed. (IHI Innovation Series white paper). Cambridge, MA: Institute for Healthcare Improvement. 2009.
Reference Source

6. Moretti F, Mortaro A, Pascu D, *et al.*: **Uno strumento per la quantificazione degli eventi avversi: il Global Trigger Tool dell'Institute for Healthcare Improvement (IHI).** 2013; (Accessed 7th November 2018).
Reference Source

7. **Azienda Ospedaliero-Universitaria "Policlinico-Vittorio Emanuele, Catania. U.O. per la Qualità e Rischio Clinico. Il protocollo di rilevazione.**
Reference Source

8. **Assessorato regionale della salute.** Regione Siciliana. IT platform.
Reference Source

9. **Conceptual Framework for the International. Classification for Patient Safety.** Version 1.1. Final Technical Report. 2009; (Accessed 7th November 2018).
Reference Source

10. 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**(6): 640–649.
PubMed Abstract | Publisher Full Text

11. Nilsson L, Borgstedt-Risberg M, Soop M, *et al.*: **Incidence of adverse events in Sweden during 2013-2016: a cohort study describing the implementation of a national trigger tool.** *BMJ Open.* 2018; **8**(3): e020833.
PubMed Abstract | Publisher Full Text | Free Full Text

12. Mortaro A, Moretti F, Pascu D, *et al.*: **Adverse Events Detection Through Global Trigger Tool Methodology: Results From a 5-Year Study in an Italian Hospital and Opportunities to Improve Interrater Reliability.** *J Patient Saf.* 2017.
PubMed Abstract | Publisher Full Text

13. Deilkås E, Bukholm G, Lindstrøm JC, *et al.*: **Monitoring adverse events in Norwegian hospitals from 2010 to 2013.** *BMJ Open.* 2015; **5**(12): e008576.
PubMed Abstract | Publisher Full Text | Free Full Text

14. Hoffmann-Völkl G, Kästenbauer T, Mück U, *et al.*: [**Detection of adverse events using IHI Global Trigger Tool during the adoption of a risk management system: A retrospective study over three years at a department for cardiovascular surgery in Vienna].** *Z Evid Fortbild Qual Gesundhwes.* 2018; **131–132**: 38–45.
PubMed Abstract | Publisher Full Text

15. Ji HH, Song L, Xiao JW, *et al.*: **Adverse drug events in Chinese pediatric inpatients and associated risk factors: a retrospective review using the Global Trigger Tool.** *Sci Rep.* 2018; **8**(1): 2573.
PubMed Abstract | Publisher Full Text | Free Full Text

16. Фастовец МН, Белорус АИ, Лысак ВП, *et al.*: [**Incidence of adverse medical events in the neonatal intensive care unit with the help of a global trigger tool].** *Wiad Lek.* 2017; **70**(3 pt 1): 483–488.
PubMed Abstract

17. Xu XD, Yuan YJ, Zhao LM, *et al.*: **Adverse Events at Baseline in a Chinese General Hospital: A Pilot Study of the Global Trigger Tool.** *J Patient Saf.* 2016.
PubMed Abstract | Publisher Full Text

18. de Vries EN, Ramrattan MA, Smorenburg SM, *et al.*: **The incidence and nature of in-hospital adverse events: a systematic review.** *Qual Saf Health Care.* 2008; **17**(3): 216–223.
PubMed Abstract | Publisher Full Text | Free Full Text

19. Mevik K, Hansen TE, Deilkås EC, *et al.*: **Is a modified Global Trigger Tool method using automatic trigger identification valid when measuring adverse events?: A comparison of review methods using automatic and manual trigger identification.** *Int J Qual Health Care.* 2018.
PubMed Abstract | Publisher Full Text

20. Bates DW, Evans RS, Murff H, *et al.*: **Policy and the future of adverse event detection using information technology.** *J Am Med Inform Assoc.* 2003; **10**(2): 226–228.
PubMed Abstract | Publisher Full Text | Free Full Text

21. Parrinello V: **"Replication data for to Developing and implementing the Global Trigger Tool methodology across a large health system in Sicily"**. Harvard Dataverse, V1. 2019.
http://www.doi.org/10.7910/DVN/YQNKCC

Open Peer Review

Current Peer Review Status: ✔ ?

Version 2

Reviewer Report 19 September 2019

https://doi.org/10.5256/f1000research.22544.r53740

✔ **James M. Naessens**
Division of Health Care Policy and Research, Mayo Clinic, Rochester, MN, USA

I appreciate the revisions and responses that the authors have made. I believe that the manuscript is now more understandable. I have several editorial suggestions:

- The heading for the fourth column of Table 2 is still confusing. Can you change it to match the heading for the 4th column of Table 3? "AEs with trigger"?
- Delete the duplicate sentence under Sample Selection.
- Change 1st paragraph of Review Process as follows: "The third reviewer was always a physician. The physician did not review the CRs, but had to authenticate the consensus of the two primary record reviewers. The physician authenticated the findings of the adverse events, the rating of severity and provided answers to questions of the record reviewers about findings in a specific record."
- I'm not sure all readers would understand that an "isolated trigger" in Table 1 is a "single positive trigger" in the 2nd paragraph of results. I'd change the sentence to "single, positive trigger".
- Some of the numbers in the 3rd paragraph of Results do not match Table 1.
- First sentence of 4th paragraph of Results should read: "We detected 1542 AEs in 975 CRs (i.e. patients) with AEs."
- Finally, under Limitations, first sentence, delete "some".

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Health Services Research, Patient Safety and Quality of Care Measurement, Biostatistics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 01 May 2019

https://doi.org/10.5256/f1000research.19713.r46773

?

Brent C. James
Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA

General comments:

The Harvard Medical Practice Study (HMPS – cited by the authors) used post-discharge chart review to detect care-associated injuries (adverse events – AEs) that occurred during hospitalization. Under HMPS, trained nurses reviewed a random selection of charts. If those nurses discovered what they judged to be care-associated injuries, they flagged those events in the chart as a potential AE. Charts that contained one or more AEs were forwarded to 2 independent physician reviewers. If both physician reviewers judged, for each AE, that event had occurred, then it was reported as a confirmed AE. Those physicians also judged whether each AE was avoidable, and whether each could be considered negligent (substandard) care.

The IHI Global Trigger Tool (IHI GTT) built on the HMPS methodology. It attempted to improve the ability of initial nurse reviewers to detect potential AEs, by providing a set of 51 “review triggers,” falling into 7 major subcategories – explicit initial events that, if detected in the chart, chained to examination of other specific events. Initial assessments of the IHI GTT showed much higher detection rates for AEs than did the original HMPS methodology, which itself far exceeded typical voluntary reporting mechanisms.

This study reviews the use of the IHI GTT in 44 Sicilian public hospitals across a 3 year time period – June 2015, through June 2018. The study’s authors adapted the IHI GTT to their specific environment. They also added 3 additional review triggers, beyond those included in the original IHI GTT. They focused their analysis on a subset of all charts assessed with their modified IHI GTT: They analyzed only charts for patients hospitalized on medicine, surgery, obstetrics, and intensive care wards.

As this study notes, many other groups are using the IHI GTT to detect AEs. This work is useful because it supplies empiric observation of one such use, across a large system of hospitals and an extended period of time.

This report has at least 2 major differences from other work in the field. First, only 6.6% of records reviewed had at least one care-associated adverse event. Other studies have shown much higher AE rates. The authors note this in their text, and suggest it may reflect differences in local environments and

chart review methods. Second, most other IHI GTT reviews show adverse drug events (medication-related events - overdoses, drug-drug interactions, and allergic or idiosyncratic reactions) as the dominant category of AEs detected. In this study, they are a distant number 3. Why?

Specific suggestions:

1. It appears that the entire IHI GTT program assessed a total of 18,008 records from among 105 hospital wards from June 2015 through June 2018. However, this analysis examines only medicine, surgery, obstetrics, and intensive care wards. Those represented a total 89 wards, with a total of 14,706 records reviewed.

   This is unclear in both your abstract and your text. It would be very helpful to more clearly explain how you derived the CRs included in your analysis.

2. You mention that you had a total of 199 individuals who participated in 71 3-person review teams. Obviously, some people participated in more than 1 team. It is not clear how the 2 initial reviewers shared their work, before their findings were submitted to a physician for final validation. Did they both separately review all records? Did they divide their assigned records between them? Please clarify.

3. In your introduction, 2nd paragraph, you review alternative methods for detecting AEs, including incident reporting, CR review, and automated administrative data review (e.g., PSIs). Consider adding “prospective (concurrent) clinical trigger systems,” that track possible real-time clinical responses to AEs then track back to see if an AE actually occurred. Dr. R. Scott Evans at LDS Hospital in Salt Lake City, Utah; and Dr. David Bates at Brigham & Women’s Hospital in Boston, Massachusetts, developed and demonstrated such systems. Such approaches find AEs that never make their way into a traditional clinical record, and may detect far more events.

4. Under Methods, please change the heading “Sampling selection” to say something like “Sample selection.” More importantly, the frame within which “10 inpatient CRs” were selected monthly is not clear. Working the total number of charts reviewed backward, you were probably sampling “10 inpatient CRs” each month for each of the 44 Sicilian public hospitals. Please clarify. How was record selection balanced across the types of wards (medicine, surgery, obstetrics, and ICUs) in each hospital?

5. Under Results, please add clarity regarding the number of patients that had at least 1 AE during their index hospitalization (975); then break out the number of patients who experienced a single AE versus those who had more than 1 AE during their index hospitalization.

6. Table 1: Please clarify – I can’t understand what you mean by the phrases “CRs with triggers per inpatient wards” versus “CRs with trigger isolated per wards”; or “CRs with trigger” versus “CRs with trigger isolated”.

7. Consider moving the first 2 paragraphs of the section entitled “Rates of triggers” from the Discussion section into the Results section.

8. Clarify your text in the first paragraph of your “Rates of triggers” section. For example, it might read:

   “In this study, 37.9% (n=5,574) of all CRs examined had at least 1 positive trigger. Of those, 2,778 CRs had a single positive trigger (49.8% of all CRs with positive triggers) while 2,796 CRs had more than one positive trigger (51.2% of all CRs with positive triggers).

9. Consider modifying Table 2:

   a) Label each of the subsections in Table 2. For example, in your text you cite the Table relative to “general care triggers”. What are “general care triggers”? Presumably you mean the subset of triggers with “C” labels. It would be quite helpful if you put headings on each of the subsections (for C, M, S, P, and I) that labelled the contents of each subsection.

   b) You are (quite appropriately) approaching AE detection as a screening test with 2 main steps. Step 1 involves evaluation of the initial triggers. Step 2 involves following up on positive initial triggers, to see if a validated AE emerges. This is done in the context of a general review by a trained clinical expert, who can (and sometimes does) detect AEs that fall outside of the trigger system.

   It would be very useful if you separated the total time spent on CR review into (a) initial assessment of the triggers; versus follow-up analysis of the positive triggers.

   It would similarly be very useful if, in the 3rd column of Table 2 (labelled “Times associated with AEs, n (%)” you showed the proportion of positive triggers that yielded a confirmed AE related to/deriving from the initial positive trigger. The Table, as currently constructed, is quite confusing – in 10 instances (C03, C04, C08, C11, C12, C14, M02, M04, I01, I03) the count in 3rd column is larger than the count in 2nd column. Presumably, this is because you are not connecting the trigger to derivative AEs (but I can’t really tell, with any certainty).

10. The section entitled “Rates of AEs”, first paragraph, contains the sentence:

    “It would appear that the percentage of clinical records with triggers increases with the increase of percentage of CRs examined, compared to patients discharged, while the percentage of CRs with AEs remains more stable.”

    I can’t translate quite what that means. Please simplify and clarify. It may relate to the 2 upper lines in Figure 2, which both appear to be increasing over time. However, you supply no statistical analysis to show that trends over time in the 2 lines are statistically associated, and you offer no reasoning as to why such a relationship may have meaning.

11. Table 3 appears to identify high frequency triggers, then count the number of AEs of any sort that occur in CRs that had a high frequency trigger. It would be much more useful if you tracked each

trigger to related or derivative AEs, rather than treating AEs generically.

12. In the section labelled “Rates of AEs” you have a bulleted list. The last item in that list says:

    “Intensive-care-related triggers have been identified 1,817 times with an AE in 1,924 (i.e., it is very common for triggers to identify more AEs in the same patient)”.

    I think I get the gist of what you mean, but (a) this sentence needs more clarity; and (b) the main purpose for identifying AEs is to move toward intervention and prevention (better, safer care). In that circumstance, treating AEs as generic items is not very useful. It is necessary to list them by specific causes. In that framing, associating triggers (which are cause specific) with a total generic AE count is not very informative or useful. You might want to rethink your framing.

Spelling corrections:

1. In the Abstract, Methods subsection, it should say “300,000 inpatient yearly admission**s**”.
2. In the Methods section, it should say “Sample selection,” not “Sampling selection”.
3. In the Discussion section, 2nd paragraph, it should say “A critical appraisal of the studies and their results is difficult, as the methodolog**ies** use **are** heterogeneous,”.
4. In Figure 1, the label for the blue line should say “**Sensitivity (%)**,” not “Sensibility (%)”.
5. In the section labelled “Rates of AEs” there is a bulleted list. The 4th bullet should say “Surgery-**r** elated” (not “Surgery-elated”).

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
No source data required

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Clinical quality improvement; patient safety; high-reliability organizations

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 28 May 2019

vincenzo parrinello, U.O. Quality management and patients safety - Azienda Ospedaliero-Universitaria, Catania, Italy

We appreciate the review by Professor Btrent of our paper, the accuracy of his observations and the relevance of his suggestions. We are honored that he considers our paper "*useful because it supplies empiric observation of one such use, across a large system of hospitals and an extended period of time.*"

We are aware that some of the points require further clarity, which may have been lost as we limited the word count.

Comment
The Harvard Medical Practice Study (HMPS – cited by the authors) used post-discharge chart review to detect care-associated injuries (adverse events – AEs) that occurred during hospitalization. Under HMPS, trained nurses reviewed a random selection of charts. If those nurses discovered what they judged to be care-associated injuries, they flagged those events in the chart as a potential AE. Charts that contained one or more AEs were forwarded to 2 independent physician reviewers. If both physician reviewers judged, for each AE, that event had occurred, then it was reported as a confirmed AE. Those physicians also judged whether each AE was avoidable, and whether each could be considered negligent (substandard) care.
The IHI Global Trigger Tool (IHI GTT) built on the HMPS methodology. It attempted to improve the ability of initial nurse reviewers to detect potential AEs, by providing a set of 51 "review triggers," falling into 7 major subcategories – explicit initial events that, if detected in the chart, chained to examination of other specific events. Initial assessments of the IHI GTT showed much higher detection rates for AEs than did the original HMPS methodology, which itself far exceeded typical voluntary reporting mechanisms.
Response
Harvard's medical practice study is a key pillar in patient safety and care research. We agree that the IHI Global trigger tool is an attempt to improve the Harvard Medical Practice Study methodology.
In our opinion, the IHI Global Trigger Tool has two significant differences from the Harvard medical practice study.
The first is that the IHI Global Trigger Tool is conducted on a random sample of clinical records that are analyzed only if triggers are present.
The second is that the IHI Global Trigger Tool "focuses on only those adverse events related to the active delivery of care (commission) and excludes, as much as possible, issues related to substandard care" (omission). (IHI Global Trigger Tool for Measuring Adverse Events. Second Edition 2009.
http://www.ihi.org/resources/Pages/Tools/IHIGlobalTriggerToolforMeasuringAEs.aspx)

Comment
This report has at least 2 major differences from other work in the field. First, only 6.6% of records

reviewed had at least one care-associated adverse event. Other studies have shown much higher AE rates. The authors note this in their text, and suggest it may reflect differences in local environments and chart review methods.

Response

The percentages of adverse events found in other studies are very different. In a recent systematic review of the literature, the percentage of adverse events is between 7% and 51%. (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 (6): 640-649.) We think that the most important reasons, which can explain these wide differences, could be twofold.

The first could be due to the different care settings (medicine, surgery, obstetrics, ICU) and therefore to the different complexity of the patients studied. A risk adjustment system would be needed to evaluate the results homogeneously. In our experience, for example, there are significant differences in the AEs rate between apparently homogeneous wards belonging to different hospitals. We have not included these tables to be brief. Since it seems to us that this is very interesting, this aspect will be the subject of other publications.

The second reason could be due to changes in the detection protocol. The survey methods are, in fact, different in the different studies. We, for example, have turned our attention only to the triggers observed during admission and not to the triggers present at the time of admission. Therefore the adverse events found do not include those present at the time of admission. This could justify the differences in EAs observed compared to other studies.

Comment

Second, most other IHI GTT reviews show adverse drug events (medication-related events - overdoses, drug-drug interactions, and allergic or idiosyncratic reactions) as the dominant category of AEs detected. In this study, they are a distant number 3. Why?

Response

We do not know why the drug-related AEs observed in our study are infrequent compared to other studies.

In our study the most frequently detected AEs are Health care–associated Infection. In addition, Rutberg et al. Report that Health care–associated Infection accounted for 39% of Aes. (Rutberg H, Borgstedt Risberg M, Sjödahl R, et al. Characterization of adverse events detected in a university hospital: a 4-year study using the Global Trigger Tool method. BMJ Open. 2014; 4: e004879.)

Von Plessen and Coll report that Infections, pressure ulcers procedures-related and gastrointestinal problems were common. (Von Plessen C, Kodal AM, Anhøj J. Experiences with Global Trigger Tool reviews in five Danish hospitals: an implementation study. BMJ Open. 2012; 2: e001324.

Schildmeijer et al report that "*Overall, the level of agreement for detecting AEs is the level of harm for healthcare-related infections, that is, pneumonia, sepsis and urinary tract infection*." (Schildmeijer K, Nilsson L, Årestedt K, et al. Assessment of adverse events in medical care: lack of consistency between experienced teams using the global trigger tool. BMJ Qual Saf. 2012; 21: 307-314.)

Landrigan et al. report that "*Harms that were detected in procedures* (186/588)" (Landrigan CP, Parry GJ, Bones CB, et al. Temporal trends in rates of patient harm resulting from medical care. N Engl J Med. 2010; 363: 2124-2134.)

Mortaro et al. report that the most common types of AEs detected were related to surgical procedures (Mortaro A, Moretti F, Pascu D, et al .: Adverse Events Detection Through Global Trigger Tool Methodology: Results From a 5-Year Study in an Italian Hospital and Opportunities to Improve Interrater Reliability. J Patient Saf. 2017.)

Comment
It appears that the entire IHI GTT program assessed a total of 18,008 records from among 105 hospital wards from June 2015 through June 2018. However, this analysis examines only medicine, surgery, obstetrics, and intensive care wards. Those represented a total 89 wards, with a total of 14,706 records reviewed. This is unclear in both your abstract and your text. It would be very helpful to more clearly explain how you derived the CRs included in your analysis.
Response
We agree that this may be unclear. From June 2015 to June 2018, 18.008 CRs were examined. In this study, we analyzed 14,706 CRs relating to patients discharged from 89 medicines, surgery, obstetrics and intensive care wards. 16 wards and 3,302 CRs concerned pediatrics wards and emergency department.

Comment
You mention that you had a total of 199 individuals who participated in 71 3-person review teams. Obviously, some people participated in more than 1 team. It is not clear how the 2 initial reviewers shared their work, before their findings were submitted to a physician for final validation. Did they both separately review all records? Did they divide their assigned records between them? Please clarify.
Response
The total number of reviewers was 199 divided into 71 teams. In some teams, medical reviewers from another team were the supervisors. The two primary record reviewers should each review all records independently. The third reviewer was always a physician. The physician, who did not review the records, authenticated the consensus of the two primary record reviewers.

Comment
In your introduction, 2nd paragraph, you review alternative methods for detecting AEs, including incident reporting, CR review, and automated administrative data review (e.g., PSIs). Consider adding "prospective (concurrent) clinical trigger systems," that track possible real-time clinical responses to AEs then track back to see if an AE actually occurred. Dr. R. Scott Evans at LDS Hospital in Salt Lake City, Utah; and Dr. David Bates at Brigham & Women's Hospital in Boston, Massachusetts, developed and demonstrated such systems. Such approaches find AEs that never make their way into a traditional clinical record, and may detect far more events.
Response
This is very interesting.
In part, this concept is reported in the conclusions *"The introduction of electronic medical records would allow a quicker process with the automation of the identification of triggers and the possibility of linking together in the identification of adverse events and near misses, especially where there has been more than one trigger "*
We will review the work and attempt to explain this concept better.

Comment
Under Methods, please change the heading "Sampling selection" to say something like "Sample selection." More importantly, the frame within which "10 inpatient CRs" were selected monthly is not clear. Working the total number of charts reviewed backward, you were probably sampling "10 inpatient CRs" each month for each of the 44 Sicilian public hospitals. Please clarify. How was record selection balanced across the types of wards (medicine, surgery, obstetrics, and ICUs) in each hospital?
Response

We will change the heading "Sampling selection" to "Sample selection."
We randomly selected 10 CRs for each department that participated, not for each hospital. The wards that participated were not homogeneously represented in the different hospitals. In some hospitals only the ICU CRs were analyzed. In some hospitals there were no medical, surgical or obstetrics or ICU wards. In others there were more than one medicine, surgery, obstetrics and ICU wards. In addition, hospitals were recruited at different times.
The CRs sample is therefore not representative. The aim of our study was not to provide a true representation of the patient population present in hospitals during the observation period. Coherently with as reported with the IHI GTT protocol, the main purpose of applying GTT methodology is to produce a sampling approach that is sufficient for the design of safety work in the hospital.

Comment
Under Results, please add clarity regarding the number of patients that had at least 1 AE during their index hospitalization (975); then break out the number of patients who experienced a single AE versus those who had more than 1 AE during their index hospitalization.
Response
We will indicate the number of patients who experienced a single AE versus those who had more than 1 AE during their index hospitalization.

Comment
Table 1: Please clarify – I can't understand what you mean by the phrases "CRs with triggers per inpatient wards" versus "CRs with trigger isolated per wards"; or "CRs with trigger" versus "CRs with trigger isolated".
Response
We will modify the table.

- The second line "CRs with triggers for inpatient wards" indicates the number and percentage of CRs with one or more triggers.
- The third line "CRs with trigger isolated for wards" indicates the number and percentage of CRs with only triggers.
- The fourth and fifth lines are repetitions. It it is a printing error.

Comment
Consider moving the first 2 paragraphs of the section entitled "Rates of triggers" from the Discussion section into the Results section.
Response
We will move the first 2 paragraphs of the section entitled "Rates of triggers" from the Discussion section into the Results section.

Comment
Clarify your text in the first paragraph of your "Rates of triggers" section. For example, it might read: "In this study, 37.9% (n=5,574) of all CRs examined had at least 1 positive trigger. Of those, 2,778 CRs had a single positive trigger (49.8% of all CRs with positive triggers) while 2,796 CRs had more than one positive trigger (51.2% of all CRs with positive triggers).
Response
Thank you for this good advice. We will modify accordingly.

Comment

a) Label each of the subsections in Table 2. For example, in your text you cite the Table relative to "general care triggers". What are "general care triggers"? Presumably you mean the subset of triggers with "C" labels. It would be quite helpful if you put headings on each of the subsections (for C, M, S, P, and I) that labelled the contents of each subsection.

Response

We will put headings on each of the subsections (for C, M, S, P, and I) that labelled the contents of each subsection.

Comment

b) You are (quite appropriately) approaching AE detection as a screening test with 2 main steps. Step 1 involves evaluation of the initial triggers. Step 2 involves following up on positive initial triggers, to see if a validated AE emerges. This is done in the context of a general review by a trained clinical expert, who can (and sometimes does) detect AEs that fall outside of the trigger system.

It would be very useful if you separated the total time spent on CR review into (a) initial assessment of the triggers; versus follow-up analysis of the positive triggers.

Response

We are unclear of the changes required on this point.

Comment

it would similarly be very useful if, in the 3rd column of Table 2 (labelled "Times associated with AEs, n (%)" you showed the proportion of positive triggers that yielded a confirmed AE related to/deriving from the initial positive trigger. The Table, as currently constructed, is quite confusing – in 10 instances (C03, C04, C08, C11, C12, C14, M02, M04, I01, I03) the count in 3rd column is larger than the count in 2nd column. Presumably, this is because you are not connecting the trigger to derivative AEs (but I can't really tell, with any certainty).

Response

In the third column of table 2 we reported the number of times a trigger was found.
In the fourth column the number of times in which an adverse event was related to that trigger.
The title of the fourth column (Times associated, with AE)does not express this clearly.
Therefore in the third column the triggers are indicated, while in the fourth column the adverse events are indicated. We have found some CRswith more than one adverse event. In the third column of table 2 we intended to report that the C03 trigger was found in 279 CRs examined. Instead, in the fourth column of Table 2 we wanted to indicate that in 438 adverse events the C03 trigger was present.

Comment

The section entitled "Rates of AEs", first paragraph, contains the sentence:

"It would appear that the percentage of clinical records with triggers increases with the increase of percentage of CRs examined, compared to patients discharged, while the percentage of CRs with AEs remains more stable."

I can't translate quite what that means. Please simplify and clarify. It may relate to the 2 upper lines in Figure 2, which both appear to be increasing over time. However, you supply no statistical analysis to show that trends over time in the 2 lines are statistically associated, and you offer no reasoning as to why such a relationship may have meaning.

Response

We mean that we did not find a correlation between the number of CRs examined in the period and the number of EAs observed in the same period. We will try to verify if the differences are

statistically significant.

Comment
Table 3 appears to identify high frequency triggers, then count the number of AEs of any sort that occur in CRs that had a high frequency trigger. It would be much more useful if you tracked each trigger to related or derivative AEs, rather than treating AEs generically.
Response
This would be very interesting to describe, but we do not know if it will be possible to provide this kind of detail due to editorial guidelines. It could be the subject of a new study to be published later.

Comment
In the section labeled "Rates of AEs" you have a bulleted list. The last item in that list says: "Intensive-care-related triggers have been identified 1,817 times with an AE in 1,924 (i.e., it is very common for triggers to identify more in the same patient)".
I think I get the gist of what you mean, but (a) this sentence needs more clarity; and (b) the main purpose of identifying AEs is to move towards intervention and prevention (better, safer care). In that circumstance, treating AEs as generic items is not very useful. It is necessary to list them by specific causes. In that framing, associating triggers (which are cause specific) with a total generic AE count is not very informative or useful. You might want to rethink your framing.
Response
This would also be interesting to describe. We have been required to be fit in with te word count and to comply with the editorial guidelines. We will clarify this section

Competing Interests: No competing interests were disclosed.

Reviewer Report 28 March 2019

https://doi.org/10.5256/f1000research.19713.r45400

?

James M. Naessens
Division of Health Care Policy and Research, Mayo Clinic, Rochester, MN, USA

This paper by Parrinello *et al.*, presents the experience of applying the Italian version of the Institute for Healthcare Improvement's Global Trigger Tool (GTT) for screening for adverse events (AEs) to a random sample of medical records for hospital discharges from June 2015 to June 2018 for all 44 public hospitals in Sicily. The authors have done a nice job of describing their overall experience with the GTT and focused on the quantity and type of triggers which helped identify adverse events. There were a few issues where some further explanation would clarify the methods and results, and aid the reader in evaluating the contributions of this paper. In addition, there were a couple of minor editing points:

- In my first reading of the paper, I was confused by Table 2. It is not clear how a trigger could be associated with AEs more often than it is detected. This happened with C03, C04, C08, etc. Table 3 presents some of the same information, but labels the second column as AEs with triggers. To

evaluate the effectiveness of GTT as a screening tool, it is useful to see both what percent of records with a specific trigger had an AE as well as how many AEs were associated with that trigger. The authors should clarify in **Methods** how they identified AEs (role of physician, why record without a trigger was reviewed) and clarify whether results they present are based on the count of AEs or the percent of records with AEs (or both). I'd recommend both, as in Table 1.

- The first paragraph of the results states that 18,008 records were reviewed, but only 14,706 were analyzed. The authors should explain why the records were excluded, to show they did not bias the results. A flow chart might help.
- They also state that 7 CRs had AEs without triggers. My understanding of the GTT method is that only records with triggers are reviewed for AEs. Please explain how these AEs were found.
- The first paragraph under "Rates of triggers" in the discussion, which should be moved to the **Results** section, refers to significant differences. The method of testing these statistics should be added to the "Statistical analysis" section of the **Methods.**

There were several typos or minor errors:

- P3, Methods, Review team: 71-person team (?).
- Table 1, lines 2 vs. 4 and lines 3 vs. 5: not clear what the differences are. How do we get different %'s?
- Figure 1: style of sensitivity line in graph does not match that of label.
- Table 5: Procedure complication has an asterisk, but no explanation.
- Figure 2: typo (EA) in 3rd label.
- P9: surgery-elated (?).

After reading the manuscript, it appears that there are two main questions that can be answered from the results:

- How does patient safety in Sicily, based on the GTT, compare to other published studies?
- Can we learn anything from the analysis of the triggers that can help us improve the review process?

The first question is addressed in the discussion (page 4). Adding a confidence interval to the reported rate from this study (6.6%) would clarify that these hospitals are as good or better than other reported rates.

The most important conclusions I would draw from the trigger analysis is that few AEs were identified by isolated triggers and many isolated triggers are not associated with AEs. Although not highly useful for identifying AEs, some isolated triggers may be direct measures of "near misses". I would restructure the results and discussion to emphasize these points.

Overall, this is a good description of a broad-based standardized implementation of the Global Trigger Tool. The authors have presented suggestions for making improvements to the method.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
No source data required

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: Dr. Naessens was an invited speaker at the initial meeting of Sicilian hospital personnel when this project was being implemented. His travel expenses were reimbursed, but no honorarium was necessary.

Reviewer Expertise: Health Services Research, Patient Safety and Quality of Care Measurement, Biostatistics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 03 Apr 2019

vincenzo parrinello, Vincenzo Parrinello, U.O. Quality management and patients safety - Azienda Ospedaliero-Universitaria, Catania, Italy

We appreciate the review by Professor Naessens of our paper, the accuracy of his observations and the relevance of his suggestions. We are honored that he considers our paper "*a good description of a broad standardized implementation of the Global Trigger Tool*" and that our work "*presented suggestions for improving the method.*"

We are aware that some of the points require further clarity, which may have been lost as we limited the word count.

Comment
In my first reading of the paper, I was confused by Table 2. It is not clear how a trigger could be associated with AEs more often than it is detected. This happened with C03, C04, C08, etc. Table 3 presents some of the same information, but labels the second column as AEs with triggers. To evaluate the effectiveness of GTT as a screening tool, it is useful to see both what percent of records with a specific trigger had an AE as well as how many AEs were associated with that trigger. The authors should clarify in Methods how they identified AEs (role of physician, why record without a trigger was reviewed) and clarify whether results they present are based on the count of AEs or the percent of records with AEs (or both). I'd recommend both, as in Table 1.
Response
Table 2
In the third column of this table we reported the number of times a trigger was found.
In the fourth column the number of times in which an adverse event was related to that trigger.
The title of the fourth column (Times associated, with AE)does not express this clearly.
Therefore in the third column the triggers are indicated, while in the fourth column the adverse events are indicated. We have found some CRswith more than one adverse event. In the third

column of table 2 we intended to report that the C03 trigger was found in 279 CRs examined. Instead, in the fourth column of Table 2 we wanted to indicate that in 438 adverse events the C03 trigger was present.

Table 3

We have tried to elucidate this concept in Table 3.

This table lists the triggers that, in our experience, require more focus on. Finding one of the triggers on this list, linked to other triggers, could indicate the presence of more than one adverse event.

Effectiveness of a GTT

We agree that to evaluate the effectiveness of the GTT as a screening tool, it is useful to see both what percentage of CRs with a specific trigger have an AE and how many AEs have been associated with that trigger. We have a table to explains this, but due to space constraints it was not included and we can add it if the editorial staff agree.

Definitions and process

Adverse events were identified based on the IHI protocol definition: "*unintended physical injury resulting from or contributed to medical care that requires additional monitoring, treatment or hospitalization, or that results in death.*" The physician reviewed the consensus with the two records and reached a final agreement on the type, number, and severity of events. The physician did not review the CRs, only the summary sheet in IT-platform.

Table 1

Table 1 shows the number of CRs examined, the number of CRs with triggers and the percentage of discharged patients. One line is missing which provides the number of discharged patients.

If the editorial staff allows, we can integrate Table 1 with the number of patients discharged and the number of AEs compared to the CRs examined and with triggers.

Comment

The first paragraph of the results states that 18,008 records were reviewed, but only 14,706 were analyzed. The authors should explain why the records were excluded, to show they did not bias the results. A flow chart might help.

Response

From June 2015 to June 2018, 18.008 CRs were examined. In this study, we analyzed 14,706 CRs relating to patients discharged from 89 medicine, surgery, obstetrics and intensive care wards. 3.302 CRs concerned pediatrics and emergency department wards.

Comment

They also state that 7 CRs had AEs without triggers. My understanding of the GTT method is that only records with triggers are reviewed for AEs. Please explain how these AEs were found.

Response

Examination of the CRs was performed in accordance with the IHI protocol:

- Discharge codes, particularly infections, complications, or certain diagnoses
- Discharge summary
- Medications administration record
- Laboratory results
- Prescriber orders
- Operative records (operational report and record anesthesia, if applicable)
- Nursing notes
- Physician progress notes
- If time permits, any other areas of the record

In seven CRs the reading of the Discharge codes and of the Discharge summary has allowed us to identify directly the adverse events represented by two cases of invasive procedures complications, two cases of surgical complications, a case of hypoglycemia and a case of adverse reaction to the administration of a drug. Therefore the reviewers, having identified the adverse event, did not look for triggers.

Comment
The first paragraph under “Rates of triggers” in the discussion, which should be moved to the **Results** section, refers to significant differences. The method of testing these statistics should be added to the “Statistical analysis” section of the **Methods**
Response
The term "significantly" used in the first paragraph under "Rates of triggers" in the discussion, is misleading. We did not intend to claim that there was a statistically significant difference.

Comment
Style of sensitivity line in graph does not match that of label of Figure 1.
Response
if possible we will correct the figure.

Comment
P3, Methods, Review team: 71-person team (?).
Response
The total number of reviewers was 199 divided into 71 teams.
In some teams, medical reviewers from another team were the supervisors.

Comment
Table 1, lines 2 vs. 4 and lines 3 vs. 5: not clear what the differences are. How do we get different %’s?
Response
In lines 2 and 4 of table 1 the number of CRs with triggers are shown.
In line 2, the % refers to the total number of patients discharged; in line 4 the % refers to CRs with triggers. For example, from June 2015 to June 2018, 1,571 medical cases with triggers were found in the medical departments, equal to 34.3% of all patients discharged in the same period (n = 4,527) and 33.4% of all triggered CRs (n = 5,574). I think it would be clearer if we put a line with the number of patients discharged.

Comment
Table 5: Procedure complication has an asterisk, but no explanation.
Response
The explanation is Central Venous Catheter (CVC), Continuous Veno-Venous Hemofiltration (CVVH), Endoscopy, Orotracheal Intubation, Endoscopic Retrograde Cholangiopancreatography (ERCP).
If possible we will correct the table.

Comment
Figure 2: typo (EA) in 3rd label.
Response
Figure 2: typo (EA) in 3rd label. is a translation error (AE). In Italian it is EA,

Comment
P9: surgery-elated (?).
Response
it would be surgery-related.

Comment
After reading the manuscript, it appears that there are two main questions that can be answered from the results:
- *How does patient safety in Sicily, based on the GTT, compare to other published studies?*
- *Can we learn anything from the analysis of the triggers that can help us improve the review process?*

The first question is addressed in the discussion (page 4). Adding a confidence interval to the reported rate from this study (6.6%) would clarify that these hospitals are as good or better than other reported rates.
Response
Regarding the addition of a confidence interval at the rate reported by this study (6.6%) to clarify if these hospitals are equal or better than other reported rates, we think that the methodologies used in the various studies are too heterogeneous, protocols are often locally adapted to the local context, the populations studied are different, and the skills of the reviewers vary. For these reasons we decided not to explore this topic from a statistical point of view.

Comment
The most important conclusions I would draw from the trigger analysis is that few AEs were identified by isolated triggers and many isolated triggers are not associated with AEs. Although not highly useful for identifying AEs, some isolated triggers may be direct measures of "near misses". I would restructure the results and discussion to emphasize these points
Response
We agree and we will emphasize this point.

Competing Interests: No competing interests were disclosed.