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

Kjersti Mevik[1,6], Tonje E Hansen[2], Ellen C Deilkås[3], Alexander M Ringdal[4], Barthold Vonen[5,6]

[1] Department of surgery, Nordland Hospital Trust, Post box 1480, N-8092 Bodø, Norway, [2] Nordland Hospital Trust, PO 1480, N-8092 Bodø, Norway, [3] Unit for Health Service Research, Akershus University Hospital, N-1478 Lørenskog, Norway, [4] Division of informatics, Nordland Hospital Trust, Post box 1480, N-8092 Bodø, Norway, [5] Center for Clinical documentation and Evaluation, North Norway Regional Health Trust, N-9038 Tromsø, Norway, and [6] Institute for community medicine, The Artic University of Norway, N-9037 Tromsø, Norway

Abstract

Objectives: To evaluate a modified Global Trigger Tool (GTT) method with manual review of automatic triggered records to measure adverse events.

Design: A cross-sectional study was performed using the original GTT method as gold standard compared to a modified GTT method.

Setting: Medium size hospital trust in Northern Norway.

Participants: 1233 records selected between March and December, 2013.

Main outcome measure: Records with triggers, adverse events and number of adverse events identified. Recall (sensitivity), precision (positive predictive value), specificity and Cohen's kappa with 95 % confidence interval were calculated.

Results: Both methods identified 35 adverse events per 1000 patient days. The modified GTT method with manual review of 658 automatic triggered records identified adverse events (n=214) in 189 records and the original GTT method identified adverse events (n=216) in 186 records. 110 identical records where identified with adverse events by both methods. Recall, precision, specificity and reliability for records identified with adverse events were respectively 0.59, 0.58, 0.92 and 0.51 for the modified GTT method. The total manual review time in the modified GTT method was 23 hours while the manual review time using the original GTT method was 411 hours.

Conclusions: The modified GTT method is as good as the original GTT method that complies with the GTTs aim monitoring the rate of adverse events. Resources saved by using the modified GTT method enable for increasing the sample size. The automatic trigger identification system may be developed to assess triggers in real-time to mitigate risk of adverse events.

Keywords (2-6): Global Trigger Tool; automatic trigger identification; adverse events; record review, hospital care

Introduction

Identifying and measuring adverse events is important as they entail substantial burden to patients and health providers [1]. In addition, the economic burden of adverse events is considerable [2]. Adverse events have commonly been identified through voluntary incident reporting but this approach significantly underestimates the actual number of adverse events as it relies on health care providers willingness and opportunity to report [3]. Hence, trigger tools, first described by Jicks [4] and refined by Classen et al [5], were developed to identify and measure adverse events. Patient records are screened for specific elements (triggers) in the records. Once a trigger is identified a more in-depth review is performed to determine if an adverse event may have occurred [6]. The trigger search is performed in randomly selected records, usually a limited number that is manageable [7]. The Institute for Healthcare Improvement (IHI) refined the trigger tools further and developed the Global Trigger Tool (GTT) which has successfully been advocated with the aim to monitor adverse events in adult inpatients [8]. The GTT is an easy two-step method of retrospective manual review of record samples: Two primary reviewers (nurses) individually review the records for specific triggers and determine if the triggers represent any adverse events, before reaching consensus (step 1). A secondary reviewer (physician) authenticates their findings (step 2) [8]. In Norway all hospitals are instructed by the commissioning documents from the ministry of health to perform the GTT [9].

Many have considered the GTT as the best method to identify and measure adverse events. Results from the GTT demonstrates that one of five hospitalized patients experience at least one adverse event [10]–[12]. However, the practical disadvantages of the GTT, being resource-intensive due to time and personnel required, limits widespread use and adoption. Automatic identification of triggers in electronic health records (EHRs) provides a digital,

standardized and cost-effective approach to measure adverse events [13]. Rather than a reviewer searches for triggers, algorithms are written to automatically identify triggers. The benefits are promising, ones the algorithms are written, as manual review is only performed in the automatic triggered records [14]–[16]. However, the validity of automatic systems in comparison to other methods measuring adverse events varies [12]–[14].

We developed an automatic trigger identification system that identifies 42 of the GTT triggers. We included the system in a modified GTT method where manual review to identify adverse events was limited to only automatic triggered records, illustrated in figure 1. We considered the original GTT method with all manual review steps as the gold standard. This study aimed to evaluate the modified GTT methods ability to identify and measure adverse events using the original GTT method as a reference standard.

Methods

Study design

The study is an explorative cross-sectional study comparing a modified GTT method to the original GTT method to identify and measure adverse events.

Setting

The study was performed at a 524-bed trust with three hospitals in Nordland County, Northern Norway. The trust has approximately 14,000 discharges and 90,000 patients days per year in the somatic adult wards. EHRs (DIPS®, ASA) were implemented in the trust in 1992. The EHRs includes both free text (i.e.: discharge summaries, operative reports, pathology reports, radiology results, transfer of service notes, admission notes, medical progress notes and notes from other healthcare professionals) and indexed variables (i.e.:

laboratory results, admissions and discharge data, diagnosis and procedure codes). In Norwegian hospitals medication administration, prescriber orders and recording of vital parameters are still hand-written and scanned into the EHRs but are being digitalized and indexed within the next two years in clinical information systems. The trust implemented the GTT in 2010 with bi-weekly review of 70 records randomly selected from the seven main units discharge lists [17].

Participants

The records included in the study were original selected for the trusts GTT review in the period March 1 to December 31, 2013. Patient records were excluded if the patient was admitted for less than 24 hours, discharged from psychiatric or rehabilitation units, and was aged 17 years or younger, as the triggers were not developed for these patients [8]. Approval for the study was obtained from the Data Protection Official in Nordland Hospital trust and by the Regional Committee for Medical and Health Research Ethics (ref 2012/1691). The committee approved a waiver for informed consent as the study fulfilled criteria described by Baker et al [18].

Definition of triggers

The Norwegian translation of the GTT includes 57 triggers (supplementary file A) [19]. The triggers are events recorded in the clinical data such as; abnormal lab values, readmission within 30 days, return to surgery, blood transfusion or administration of drugs such as anti-dot or anti-emetic drugs [8]. Some of the triggers are adverse events by its nature, for example 3rd- or 4rd-degree perianal lacerations, pressure ulcer and injury, repair, or removal of organ because of accidental injury. However, most of them are just indicators that an adverse event

may have occurred. A more in-depth review is necessary to decide if the triggers are associated with any adverse events.

Definition of an adverse event

The definition of an adverse event adopted from the GTT was used by both methods when deciding if an adverse event was present when performing manual review of the triggered records [8]: "*Unintended physical injury resulting from or contributed to by medical care that requires additional monitoring, treatment or hospitalization, or that results in death*".

The adverse events were categorized according to severity with the adapted definitions from the National Coordinating Council for Medication Error Reporting and Prevention Index (NCC MERP) for categories E-I [20]:

Category E: Temporary harm to the patient and required intervention

Category F: Temporary harm to the patient and required initial or prolonged hospitalisation

Category G: Permanent patient harm

Category H: Intervention required to sustain life

Category I: Patient death

Review methods

The original GTT method

Two primary reviewers (nurses) reviewed the records individually in a specific order to register any presence of triggers. Once a trigger was identified, a more in-depth review was performed to investigate if the trigger was associated with an adverse event according to the described definition, all performed within a 20-minute time limit. A secondary reviewer (physician) authenticated the primary reviewers' findings. There was no time constraint for

the secondary reviewer. Griffin et al estimated that the secondary reviewer uses two hours per 20 records, confirming or deleting adverse events identified by the primary reviewers [8]. The secondary reviewer reviewed only the relevant parts of the records identified with adverse events.

The modified GTT method

The automatic trigger identification system can only identify triggers, not adverse events. The system identifies triggers based on algorithms. We have included examples used in such algorithms in supplementary file B. The algorithms for indexed variables (e.g., INR> 6, glucose< 2.8 or diagnoses/procedures codes) are based on queries. Algorithms for free text (e.g., patient fall, specialty obstetric consult, induced labour) are based on information extractions and recognitions of text strings and patterns through text mining analysis. All conditions and words representing the actual trigger (e.g., patient fell out of bed, patient slipped in the bathroom) are extracted. In addition, the system omits the information if exclusion criteria are met (e.g., the anastomosis fell in place, the catheter fell out). The automatic trigger identification system included 42 triggers used in the Norwegian GTT (see supplementary file A). Nine triggers were excluded as the information for these triggers are hand written and scanned into the EHR. The automatic trigger identification system cannot identify these triggers; use of anti-dot drugs, use of anti-emetic drugs, vitamin K administration, hypotension and abrupt medication stop. The three triggers labelled “other” related to respectively medication, general and surgical care were not included in the system, as they do not correspond to a specific adverse event but used when reviewers identify an adverse event without finding a corresponding trigger. Finally, we opted to exclude three triggers rarely identified in our previous manual review of 6720 records from 2010 to 2013.

The records, both triggered and non-triggered, were presented in an interface along with information regarding triggers identified (e.g.; type of trigger and which note/lab test/radiology or pathology report the triggers are detected in). One physician performed manual review of the triggered records to decide if the triggers were associated with any adverse events and if so, theirs severity and type. The described definition of an adverse event was applied. The manual review time used in each record was recorded. No time constraint was applied.

Statistics

1400 records from the trusts GTT review of a 10-month period were selected. 167 records were excluded as the data from the automatic trigger identification system was missing for these records, leaving a total of 1233 records included in the study.

The objective of this study was to evaluate the modified GTT method. Paired t-test was used to compare the number of triggered records, number of records with identified adverse events and number of identified adverse events between the methods. A p value < 0.05 was regarded significant. We calculated recall (sensitivity), precision (positive predictive value) and specificity with their respective 95 % confidence intervals (CI) to evaluate the validity of the modified GTT method using the original GTT method as gold standard:

$$Recall = \frac{No.\,of\ correct\ positive\ records\ identified\ by\ the\ modified\ GTT\ method}{No.\,of\ positive\ records\ identified\ by\ gold\ standard}$$

$$Precision = \frac{No.\,of\ correct\ positive\ records\ identified\ by\ the\ modified\ GTT\ method}{Total\ no.\,of\ positive\ records\ identified\ by\ the\ modified\ GTT\ method}$$

$$Specificity = \frac{No.\,of\ correct\ negative\ records\ identified\ by\ the\ modified\ GTT\ method}{No.\,of\ negative\ records\ identified\ by\ gold\ standard}$$

Recall represents the proportion of "correctly" identified records with adverse events by the modified GTT method. Precision represents the proportion of records with adverse events identified by the modified GTT method that also were identified by the original GTT method. Specificity represents the proportion of "correctly" identified records with no identified adverse events by the modified GTT method. For reliability, we used Cohen's Kappa to measure agreement of the results (inter-rater reliability) between the methods, taking into account the agreement occurring by chance. The following interpretations from Landis and Koch were used for the Cohen's Kappa coefficient: poor (<0.0), slight (0.00–0.20), fair (0.21–0.40), moderate (0.41– 0.60), substantial (0.61–0.80) and almost perfect (0.81–1.00) [21]. A 95 % CI was set. The CI for recall, precision, specificity was calculated using the Wilson score method [22]. CI for Cohen's kappa was $\kappa \pm 1.96 * SE$. All analyses were performed using SPSS (version 22.0; SPSS Chicago, IL).

Results

58 % (716) were women and average age was 58 years (range; 18-102, standard deviation (SD); 22). Mean length of stay was 5 days (range; 1-65, SD; 6).

The modified GTT method identified a total of 1216 triggers in 658 records while the original GTT method identified a total of 1267 triggers in 626 records. The number of the individually triggers identified by each method are included in supplementary file C. In 110 identical records, both methods identified adverse events. In 79 records, the modified GTT method identified adverse events alone and vice versa in 76 records (figure 2). The recall, precision, specificity and Cohen's kappa with their respective 95 % CI of the modified GTT method are

presented in table1. Figure 3 displays the types of adverse events identified by the two methods which differed between the methods. Number of records identified with adverse events and number of identified adverse events according to severity are presented in table 2.

The modified GTT method identified 34.7 adverse events (n=214) per 1000 patient days by manual review of 658 automatic triggered records for the 10-month period. Adverse events were identified in 28.7 % (n=189 records) of the automatic triggered records (n=658 record). Mean manual review time used per record was 2 minutes (range 0.2- 21.5) and the total manual review time was 23 hours. The original GTT method identified 35.0 adverse events (n=216) per 1000 patient days of 626 manual triggered records in the same 10-month period. Adverse events were identified in 15.1 % (n=186 records) of the records reviewed manually for triggers and adverse events (n=1233 records). Total manual review time of 1233 records was 411 hours.

Discussion

The aim of this study was to evaluate a modified GTT method with automatic trigger identification to identify and measure adverse events using the original GTT method as gold standard. We found that the modified GTT method is a valid, reliable and efficient method to monitor the rate of adverse events. The modified GTT method demonstrated major decrease in review time compared to the original GTT method. Both methods identified a rate of 35 adverse events per 1000 patient days. There was no significant difference between the methods regarding the severity of the identified adverse events. The modified GTT method comply with the GTTs aim to monitor the rate of adverse events over time consistently, but not completely.

The values of a “new” measure are related to values from a reference measure performed at the same time and are defined as the concurrent validity of the measure. Concurrent validity is evaluated by recall (sensitivity), precision (positive predictive value) and specificity, which we calculated for the modified GTT method. A review of current literature did not find any reference to evaluation of the validity of the GTT [23] but studies have demonstrated that the GTT identifies more adverse events than other methods [10], [11]. The purpose of the GTT is, with an easy method, to select those patients that may have experienced an adverse event by the use of triggers as screening criteria. We adopted this purpose when we evaluated the modified GTT method. We recorded therefore only the unique number of identified triggers in the triggered records and did not considered excessive testing of the individually triggers as this was beyond the scope of the study.

The modified GTT method demonstrated an efficient method to identify and measure adverse events with a total of 23 hours to complete the manual review of 658 automatic triggered records compared to 411 hours of review of 1233 records with the original GTT method. The modified GTT method reviewed only the triggered records thereby reducing the number of records to be manual reviewed by 50 %. This reduction enables for increasing the sample size without applying further resources. Critics have argued that the recommend sample size, 10 records bi-weekly, in the GTT is too low to estimate the rate of adverse events for an institution. Thus, sampling size should correspond to the hospital size [17]. Extrapolation, which is used when estimates are made on small samples, increases the random variability. Infrequent adverse events can also be missed when only samples of records are reviewed [24]. Increasing the sample size makes the results regarding the rate of adverse events more valid [17], [25].

The manual review processes differed somewhat between the two methods. Only one reviewer, a physician, performed the subsequent manual review of automatic triggered records in the modified GTT method. The original GTT method included two primary reviewers and a secondary reviewer authenticating their consensus findings. Reviewers in both methods were experienced reviewers of the GTT. The aim of the study was to assess if the rate of adverse events altered when we modified the GTT with manual review of only automatic triggered records. Hence, we do not consider the differences of the manual review processes as a bias.

Poor to moderate agreement between reviewers and between review teams have been demonstrated [26], [27].We believe the agreement can be improved by using an automatic trigger identification system. First, automatic identification of triggers in the EHR excludes the variability of manual identified triggers as triggers based on index information (i.e.; blood transfusion and dialysis) have demonstrated higher agreement than triggers derived from free text (i.e.; pressure ulcers, patient fall) [28]. Second, the manual trigger identification could suffer from the time constraints excluding possible triggers causing adverse events to be missed [27], [29]. Automatic trigger identification does not have a time constraint and all present triggers are identified. These issues make the identified adverse events based on automatic trigger identification more standardized and comparable than adverse events identified by manual trigger identification. Moreover, with further development, the automatic trigger identification system can provide a platform to identify patients at risk of adverse events in real-time. Such systems could be used to improve clinical outcome, optimize treatment, reduce the financial burden of patient harm and most importantly; reduce the suffering of the patients due to adverse events [24], [30]. However, the development of such methods requires both technical and economic inputs.

Strength and limitations

The main strength of the study is that we demonstrated a valid and efficient method to identify and measure adverse events.

Our study has some limitations. First, fifteen of the original 57 triggers were excluded in our automatic trigger identification system, but nine of them can be included when all patient data are digitalized and indexed in clinical information systems. Second, record reviews depend on that the necessary data are documented in the EHR. Records could be incomplete regarding documentation of adverse events causing adverse events to be missed. Third, this study has been performed in one hospital only. Modification of the automatic trigger identification system must be applied before adoption.

Conclusions

Our study demonstrated that the modified GTT method with automatic trigger identification is a valid, reliable and efficient method to identify and measure adverse events to comply the aim of the GTT in respect to the original GTT method. We therefore recommend that the modified GTT method should be preferred as it offers an efficient alternative to the common costly and time-consuming approaches mainly used to identify and measure adverse events. The resources saved by using the modified GTT method are considerable and this enables for increase of the sample size.

Acknowledgement

The authors would like to thank the review teams at our trust who conducted the manual reviews. We also thank the SAS Institute® for contributing in developing our automatic trigger identification system implementing the algorithms. A special thanks to Christian Hardahl in SAS® reviewing the manuscript regarding logics, to Tom Wilsgaard for help with the statistics, to Fran Griffin reviewing the manuscript and to Laila Bjølgerud for making the graphical illustration.

Funding

This work was supported by The Northern Norway Regional Health Authority.

References

[1] A. Jha and P. Pronovost, “Toward a Safer Health Care System: The Critical Need to Improve Measurement.,” *JAMA*, vol. 315, no. 17, pp. 1831–2, May 2016.

[2] T. B. Agbabiaka, M. Lietz, J. J. Mira, and B. Warner, “A literature-based economic evaluation of healthcare preventable adverse events in Europe,” *International Journal for Quality in Health Care*, vol. 29, no. 1. pp. 9–18, 2017.

[3] C. Macrae, “The problem with incident reporting.,” *BMJ Qual. Saf.*, vol. 25, no. 2, pp. 71–5, Feb. 2016.

[4] H. Jick, “Drugs — Remarkably Nontoxic,” *N. Engl. J. Med.*, vol. 291, no. 16, pp. 824–828, Oct. 1974.

[5] D. C. Classen, S. L. Pestotnik, R. S. Evans, and J. P. Burke, “Description of a computerized adverse drug event monitor using a hospital information system.,” *Hosp. Pharm.*, vol. 27, pp. 774, 776–779, 783, 1992.

[6] R. K. Resar, J. D. Rozich, and D. Classen, “Methodology and rationale for the measurement of harm with trigger tools.,” *Qual. Saf. Health Care*, vol. 12 Suppl 2, p. ii39-i45, 2003.

[7] J. Garrett, R. Paul, and C. Sammer, “Developing and Implementing a Standardized Process for Global Trigger Tool Application Across a Large Health System,” … *J. Qual.* …, vol. 39, no. 7, 2013.

[8] F. Griffin and R. Resar, “IHI Global Trigger Tool for measuring adverse events (Second Edition),” *IHI Innov. Ser. white Pap. Cambridge, Massachusetts Inst. Healthc. Imrovement;*, pp. 1–44, 2009.

[9] E. Deilkås, G. Bukholm, J. Lindstrøm, and M. Haugen, “Monitoring adverse events in Norwegian hospitals from 2010 to 2013,” *BMJ Open*, 2015.

[10] P. D. Hibbert *et al.*, “The application of the Global Trigger Tool: a systematic review,” *Int. J. Qual. Heal. Care*, vol. 28, no. 6, pp. 640–649, Sep. 2016.

[11] D. C. Classen *et al.*, “‘Global trigger tool’ shows that adverse events in hospitals may be ten times greater than previously measured.,” *Health Aff. (Millwood).*, vol. 30, no. 4, pp. 581–9, Apr. 2011.

[12] J. M. Naessens *et al.*, “A comparison of hospital adverse events identified by three widely used detection methods.,” *Int. J. Qual. Health Care*, vol. 21, no. 4, pp. 301–7, Aug. 2009.

[13] S. N. Musy *et al.*, “Trigger Tool-Based Automated Adverse Event Detection in Electronic Health Records: Systematic Review.,” *J. Med. Internet Res.*, vol. 20, no. 5, p. e198, 2018.

[14] D. C. Stockwell, E. Kirkendall, S. E. Muething, E. Kloppenborg, H. Vinodrao, and B. R. Jacobs, “Automated adverse event detection collaborative: electronic adverse event identification, classification, and corrective actions across academic pediatric institutions,” *J Patient Saf*, vol. 9, no. 4, pp. 203–210, 2013.

[15] A. K. Jha *et al.*, “Identifying Adverse Drug Events,” *J. Am. Med. Informatics Assoc.*, vol. 5, no. 3, pp. 305–314, 1998.

[16] V. Lemon and D. C. Stockwell, “Automated Detection of Adverse Events in Children,” *Pediatric Clinics of North America*, vol. 59, no. 6. pp. 1269–1278, 2012.

[17] K. Mevik, F. A. Griffin, T. E. Hansen, E. T. Deilkås, and B. Vonen, “Does increasing the size of bi-weekly samples of records influence results when using the Global

Trigger Tool? An observational study of retrospective record reviews of two different sample sizes," *BMJ Open*, vol. 6, no. 4, 2016.

[18] D. W. Baker and S. D. Persell, "Criteria for Waiver of Informed Consent for Quality Improvement Research," *JAMA Intern. Med.*, vol. 175, no. 1, p. 142, Jan. 2015.

[19] E. Deilkås, "Gjennomføring av journalundersøkelse med Global Trigger Tool (GTT) i den norske pasientsikkerhetskampanjen," 2011.

[20] S. C. Hartwig, S. D. Denger, and P. J. Schneider, "Severity-indexed, incident report-based medication error-reporting program.," *Am. J. Hosp. Pharm.*, vol. 48, pp. 2611–2616, 1991.

[21] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data.," *Biometrics*, vol. 33, no. 1, pp. 159–74, Mar. 1977.

[22] R. G. Newcombe, "Two-sided confidence intervals for the single proportion: Comparison of seven methods," *Stat. Med.*, vol. 17, no. 8, pp. 857–872, 1998.

[23] M. Hanskamp-Sebregts, M. Zegers, C. Vincent, P. J. van Gurp, H. C. W. de Vet, and H. Wollersheim, "Measurement of patient safety: a systematic review of the reliability and validity of adverse event detection with record review.," *BMJ Open*, vol. 6, no. 8, p. e011078, Aug. 2016.

[24] C. Sammer *et al.*, "Developing and Evaluating an Automated All-Cause Harm Trigger System," *Jt. Comm. J. Qual. Patient Saf.*, vol. 0, no. 0, Feb. 2017.

[25] V. Good and M. Saldana, "Large-scale deployment of the Global Trigger Tool across a large hospital system: refinements for the characterisation of adverse events to support patient safety," *BMJ Qual.* …, 2011.

[26] K. Schildmeijer, L. Nilsson, K. Arestedt, and J. Perk, "Assessment of adverse events in

medical care: lack of consistency between experienced teams using the global trigger tool," *BMJ Quality & Safety*, vol. 21. pp. 307–314, 2012.

[27] P. J. Sharek *et al.*, "Performance characteristics of a methodology to quantify adverse events over time in hospitalized patients.," *Health Serv. Res.*, vol. 46, no. 2, pp. 654–78, Apr. 2011.

[28] J. M. Naessens, T. J. O'Byrne, M. G. Johnson, M. B. Vansuch, C. M. McGlone, and J. M. Huddleston, "Measuring hospital adverse events: assessing inter-rater reliability and trigger performance of the Global Trigger Tool.," *Int. J. Qual. Health Care*, vol. 22, no. 4, pp. 266–74, Aug. 2010.

[29] M. Unbeck and K. Schildmeijer, "Is detection of adverse events affected by record review methodology? an evaluation of the' Harvard Medical Practice Study' method and the" Global Trigger," *Patient Saf …*, vol. 7, no. 1, p. 10, 2013.

[30] K. Jensen *et al.*, "Analysis of free text in electronic health records for identification of cancer patient trajectories," *Sci. Rep.*, vol. 7, p. 46226, 2017.

Table 1 Validity and reliability of the modified GTT method versus the original GTT method (gold standard)

Variable	Recall[i] (CI)[ii]	Precision[iii] (CI)[ii]	Specificity[iv] (CI)[ii]	Cohen's Kappa[v] (CI)[ii]
Triggered records	0.83 (0.80-0.86)	0.79 (0.76-0.82)	0.78 (0.74-0.81)	0.61 (0.56-0.66)
Records with adverse events	0.59 (0.52-0.66)	0.58 (0.51-0.65)	0.92 (0.91-0.94)	0.51 (0.44-0.57)
Records with adverse events within the common triggered records	0.71 (0.63-0.77)	0.61 (0.54-0.68)	0.81 (0.77-0.85)	0.50 (-0.31-1.30)

[i]Recall represent the proportion of "correctly" records identified with triggers or adverse events by the modified GTT method.

[ii]95 % confidence interval (CI).

[iii]Precision represent the proportion of records with triggers or adverse events that were confirmed by the original GTT method.

[iv]Specificity represents the proportion of "correctly" records with no identified adverse events by the modified GTT method.

[v]Cohen's Kappa is the inter-rater reliability of the modified GTT method and the original GTT method evaluated by a 2 x 2 table.

Table 2 Number of adverse events and records with adverse events according to severity identified by the modified GTT method and the original GTT method

Severity category	Original GTT method		Modified GTT method		Records with adverse events: Modified GTT method vs. Original GTT method		Number of adverse events: Modified GTT method vs Original GTT method	
	Adverse events	Records with adverse events	Adverse events	Records with adverse events	P*-value	CI 95%	P*-value	CI 95%
E	120	109	95	90	0.08	-0.032-0-002	0.045	-0.04-0.00
F	87	80	97	91	0.29	-0.008-0.026	0.38	-0.01-0.03
G	5	5	12	12	0.09	-0.001-0.012	0.09	-0.001-0.01
H	1	1	1	1	1.00	-0.002-0.002	1.00	-0.002-0.002
I	3	3	9	9	0.03	0.000-0-009	0.03	0.00-0.01
Total	216	198***	214	203**	0.81	0.01- -0.02	0.90	-0.03-0.024

Notes: Severity level according to the NCC MERP index

*P-value of Paired sample T-test

**14 admissions with two more adverse events with different severity level

***12 admissions with two or more adverse events with different severity level

Figure 1: The modified GTT method

Figure 2: Records identified with triggers and adverse events by the modified GTT method and the original GTT method

Figure 3: Adverse events identified according to types

Figure 1

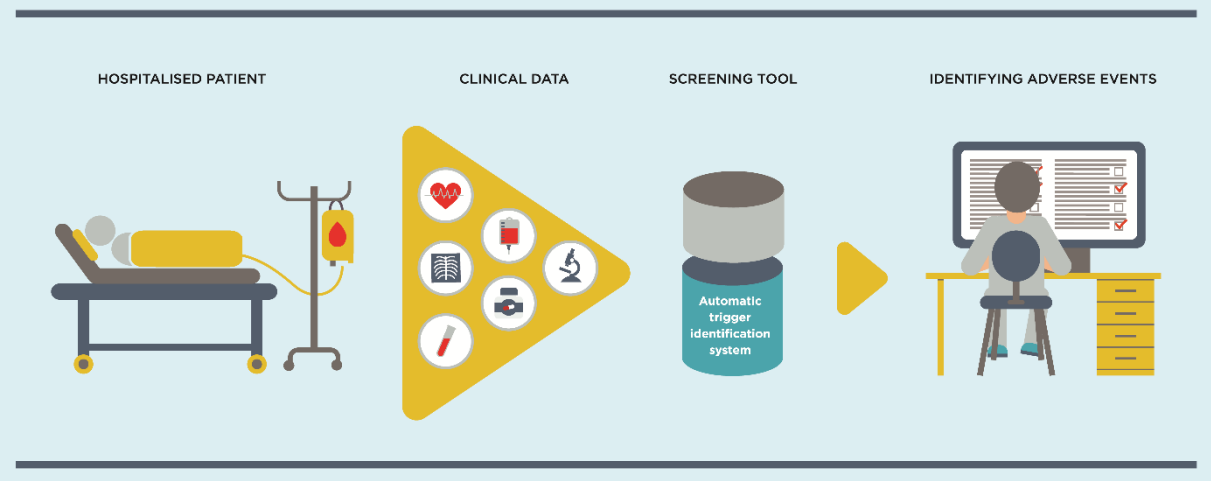

HOSPITALISED PATIENT
CLINICAL DATA
SCREENING TOOL
IDENTIFYING ADVERSE EVENTS
Automatic trigger identification system

Figure 2

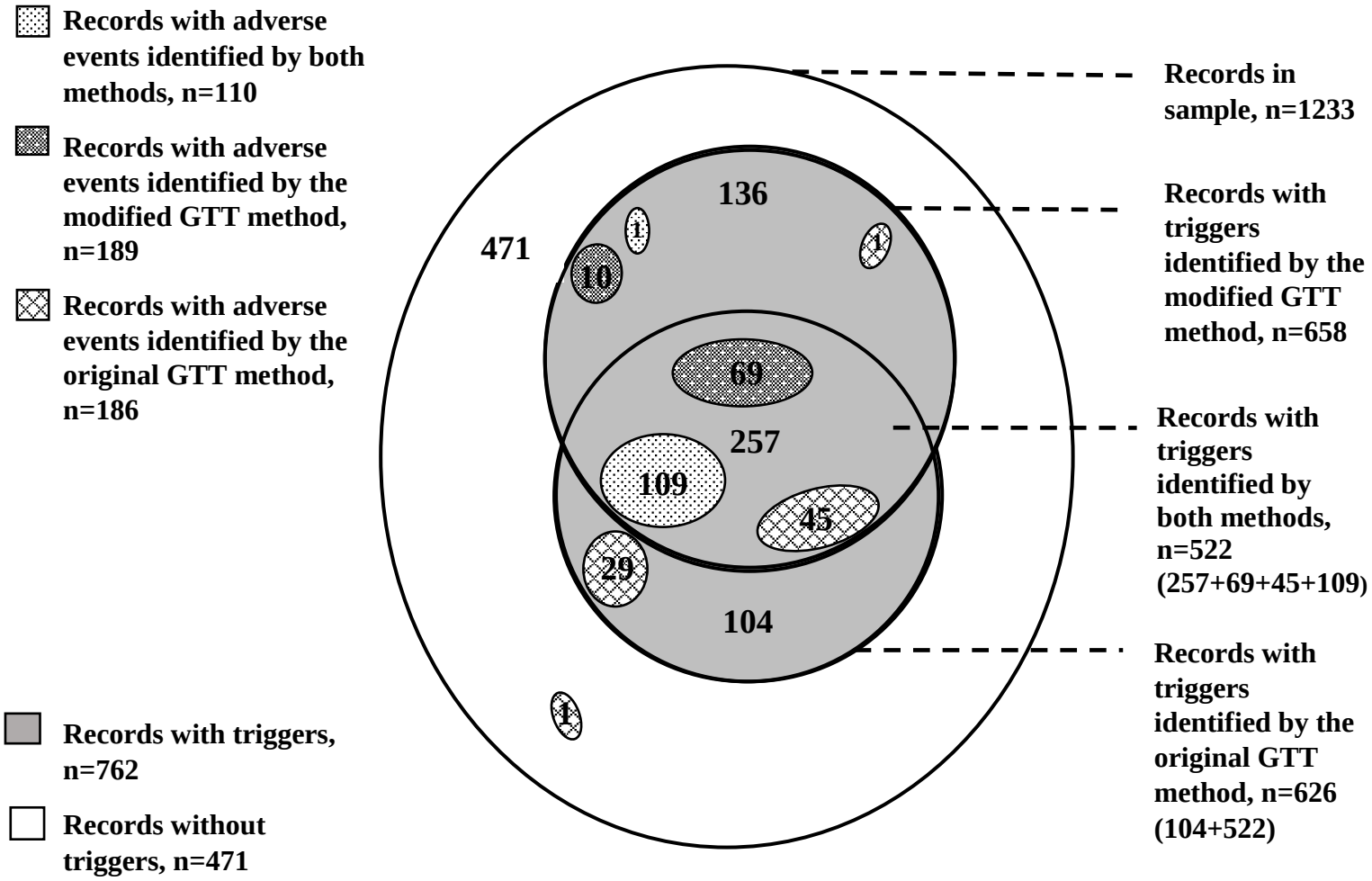
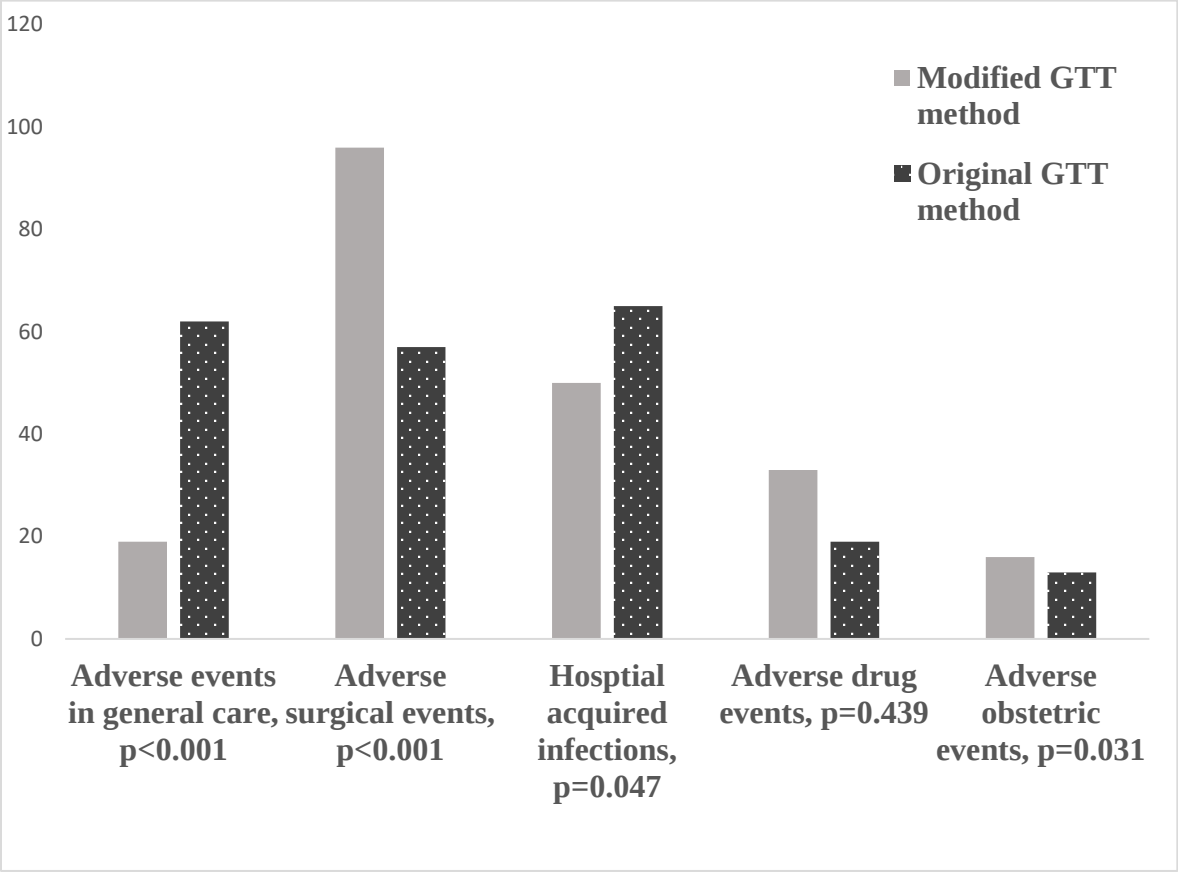

Records with adverse events identified by both methods, n=110
Records with adverse events identified by the modified GTT method, n=189
Records with adverse events identified by the original GTT method, n=186
Records with triggers, n=762
Records without triggers, n=471
471
136
1
1
10
69
257
109
45
29
104
1
Records in sample, n=1233
Records with triggers identified by the modified GTT method, n=658
Records with triggers identified by both methods, n=522 (257+69+45+109)
Records with triggers identified by the original GTT method, n=626 (104+522)

Figure 3



Adverse events in general care: allergy, bleeding, patient fall, fracture, medical technical event, thrombosis/embolism, deterioration of chronic disease and other events.

Adverse surgical events: infection after surgery, return to surgery, injury or removal of an organ by accident, bleeding after surgery, respiratory complication after surgery, switch in surgery and any other surgical complication.

Hospital acquired infection: urinary tract infection, lower respiratory infection, ventilator-associated infection, central vein catheter associated infection and other infections.