

Kaariina Pertomaa
Mika Kaartinen
Juha Öhman

patient safety monitoring in Finnish hospitals

Current experiences and future trends

REPORT



Report 19/2013

Persephone Doupi, Karolina Peltomaa,
Mika Kaartinen, Juha Öhman

**IHI Global Trigger Tool and patient
safety monitoring in Finnish hospitals**
Current experiences and future trends



NATIONAL INSTITUTE
FOR HEALTH AND WELFARE

© Authors and National Institute for Health and Welfare

Layout: Raili Silius

ISBN 978-952-245-993-0 (printed)

ISSN 1798-0070 (printed)

ISBN 978-952-245-999-2 (online publication)

ISSN 1798-0089 (online publication)

<http://urn.fi/URN:ISBN:978-952-245-999-2>

Juvenes Print – Finnish University Print Ltd
Tampere, Finland 2013

Foreword

Patient safety is gaining growing importance among Finnish health and social care service organisations. An essential component of patient safety work is the ability to monitor the levels of patient safety achieved, identify areas where improvement is needed and follow the impact of implemented interventions. Means to prevent adverse events and/or recognize them in time so as to interfere and possibly mitigate their effect are highly desirable and constitute the future direction of development in patient safety activities.

Trigger tools, both in their paper and automated versions, are among the candidate technologies that can serve the aforementioned goals. This report presents the experiences accumulated in Finland during recent years in implementing trigger tools for organisational patient safety monitoring in the hospital environment, experimentations for further development of such tools, and reviews the research evidence up-to-date regarding trigger tools' strengths and weaknesses.

The report aims to serve healthcare management and clinical staff working with patient safety issues on the organisational level, by providing background evidence with regard to trigger tool methodology and possible implementation technologies, particularly through exploiting the functionalities offered by health IT systems. The report constitutes useful reading also for developers and researchers in the fields of patient safety and healthcare quality.

The report is the product of a fruitful collaboration between THL, healthcare organisations that have pioneered activities in the area of patient safety - TAYS (Tampere University Hospital) department of Neurosurgery and VSSHP (Hospital District of Southwest Finland), as well as SAS Institute, a provider of statistical solutions with potential for implementation also for patient safety purposes.

Persephone Doupi (THL) performed literature reviews and analysis on both paper and automated trigger tools, as well as authored the chapter on future developments and acted as editor of the report. Juha Öhman and his team from TAYS experimented with modified trigger tools in Neurosurgery and together with Mika Kaartinen and the SAS Institute team explored the possibilities of data mining as an implementation approach. Karolina Peltomaa and the team from the Hospital District of Southwest Finland implemented and studied in depth the use of the IHI GTT methodology as part of their patient safety programme.

The authors would like to sincerely thank Chief Physician Erna Snellman (Paijattäme), Head Nurse Marina Kinnunen (Vaasa Central Hospital), and Riikka Vuokko (THL) who reviewed the report and provided useful commentary to improve its contents, as well as Hanna Lehto for proof-reading the manuscript. P. Doupi would also like to gratefully acknowledge the indispensable assistance and exceptional professionalism of the THL librarian services, and particularly Pia Pörfors, without whom the literature review and identification of pertinent material would not have been possible.

Abstract

Persephone Doupi, Karolina Peltomaa, Mika Kaartinen, Juha Öhman. IHI Global Trigger Tool and patient safety monitoring in Finnish hospitals. National Institute for Health and Welfare (THL). Report 19/2013. 95 pages. Helsinki, Finland 2013. ISBN 978-952-245-993-0 (printed); ISBN 978-952-245-999-2 (online publication)

Patient safety is gaining in importance among Finnish health and social care service organisations. An essential component of patient safety work is the ability to monitor the levels of patient safety achieved, in order to identify areas where improvement is needed and follow the impact of implemented interventions. Trigger tools - means for performing focused medical records reviews either manually or automatically, are among the candidate technologies that can serve the aforementioned goals. The IHI Global Trigger Tool (GTT) is a retrospective method for monitoring patient safety levels within a healthcare provider organization. It allows for longitudinal comparisons and assessment of implemented patient safety measures, and enables the identification of target areas for improvement. The method is paper-based, in other words, it does not require or depend on the use of health information systems. This report presents recent experiences in implementing the GTT for organisational patient safety monitoring in the Finnish hospital environment, including experimentations for further development of such tools, and reviews the research evidence up-to-date regarding paper and computerized trigger tools' strengths and weaknesses.

The GTT, as well as the rest of the IHI trigger tools family is a relatively new technology. Our literature review (Chapter 2), which covers studies published up to June 2011, located only nine papers specific to the IHI GTT, mostly published during the last 2-3 years. The articles concerned the tool's development and evaluation, performance features, comparisons with other methods, and examples of utilization either within or across large health systems or in national level programs. Outcomes were reported as number of adverse events/100 patients (used in five studies, with results ranging from 18.1 to 41) and/or as number of adverse events/1000 patient-days (used in four studies, with results ranging from 41.6 to 91). In all reviewed studies severity of the identified events was assessed using the National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) classification. In over half the cases, severity was described as temporary harm to the patient requiring intervention (category E). Preventability of adverse events was assessed in only three of the reviewed studies, and results reported only in two of them. Assessment was performed using a 4-level Likert scale, and the rate of preventable adverse events was reported as 51.7% and 63.1%. There are two main issues of controversy around the GTT: its reliability as a method due to its limited validation, and the suitability of the tool for use as a benchmarking instrument. Compared to full patient record review, GTT places a smaller demand on resources as a result of reviewing a smaller number of records, with a higher probability of containing actual adverse events. In addition to its ability

to detect larger number of events than other assessment methods, comparative studies also indicate that the GTT may identify different types of adverse events. In the light of these observations, we can conclude that use of the trigger tool approach can supplement incident reporting and other interventions when the aim is a comprehensive picture of the level of adverse events in an organisation. The GTT paper method can be seen as a valid alternative for regular use in the place of the resource intensive and research oriented methodology of full patient record review. Indeed, use of the GTT seems to be on the rise - at least in the US, as well as in other Nordic countries, as part of national patient safety initiatives/programmes.

In the Hospital District of Southwest Finland (VSSHP), assessment of adverse events using the IHI GTT has continued since 2009 (Chapter 3). 1335 randomised patient admissions have been sampled, out of which 890 were included into the analysis. A clear trend of reduction of adverse events according to the set target of 30% has been documented through use of the method. Severity and preventability of the identified adverse events has also been assessed. Most of the identified events belong to category E, and 65.5 % of the total was preventable. The intention is to continue with implementation of IHI's GTT in different hospital departments. It is expected that using the methodology in department level will produce more accurate and detailed information. However, this also requires the translation and validation of additional triggers related to e.g. day-surgery, paediatrics and psychiatry.

In the Neurosurgery Department of Tampere University Hospital TAYS, the ability of text mining to detect accurately the same triggers in electronic patient records as manual review was assessed (Chapter 4). The study was performed as a structured retrospective medical record review based on the use of 13 modified IHI GTT screening criteria. Three different ICD-10 based diagnosis groups (head injuries, intracranial haemorrhage and cervical spine pathology as the control group) were analysed through three parallel time representative samples (n = 556) out of a total 1969 neurosurgery admissions between April 2007 and May 2008. Based on manual review, 23.7 % (n = 132) of the 556 admissions had triggers. Compared to manual review the sensitivity of detecting triggers with text mining varied from 60 to 100 % between the triggers. Specificity between triggers varied from 80 to 98 %. The study team concluded that triggers can be found with the text mining tool, and that this method is as reliable and less time and manpower consuming than the conventional manual method.

As the adoption of electronic health records expands, also means for electronic identification of triggers become more feasible, thus reducing the reviewer's task to the identification and assessment of harm. Review of the literature on computerized trigger tools (Chapter 5) highlighted the need for high quality documentation in the patient record and the importance of data integration solutions connecting various hospital information systems. Use of coded and structured text could be the path to real-time use of prospective triggers and is at least in principle feasible based on the Finnish minimum data set, but not further investigated in practice. Prospective trigger systems however, depend heavily on significant changes in workflow and hence face

considerable barriers in implementation. Utilization of free text in the EHR through text mining tools, as in the Swedish MAG application and the Finnish neurosurgery pilot, is also a potential alternative for trigger detection.

Although use of the paper-based GTT method is currently growing, computerized variants will be the future. Automated trigger tools require further research and development, combined with rigorous evaluation. The field is in a state of evolution, and a steady generation of hybrid systems combining different approaches for patient safety event detection can be observed. The resulting applications are still not mature for wide-scale use, and are currently the focus of otherwise pioneering organisations. Nordic collaboration both on the research and on the implementation level is progressing and can result in significant contributions to the area of patient safety monitoring. Nevertheless, it should not be forgotten that collecting data for measurement is only the first step in a process and *not* the end goal in and of itself. For real improvements in patient safety to be achieved, a strategy for disseminating and utilizing the findings of trigger tool methods in order to bring about actual changes in clinical practices must also be formulated as part of the trigger tool introduction.

Keywords: Global Trigger Tool, patient safety monitoring, hospitals, electronic patient record, computerized trigger tools.

Tiivistelmä

Persephone Doupi, Karolina Peltomaa, Mika Kaartinen, Juha Öhman. IHI Global Trigger Tool ja potilasturvallisuuden seuranta suomalaisissa sairaaloissa Terveyden ja hyvinvoinnin laitos (THL). Raportti 19/2013. 95 sivua. Helsingissä 2013
ISBN 978-952-245-993-0 (painettu); ISBN 978-952-245-999-2 (verkkojulkaisu)

Potilasturvallisuuden merkitys kasvaa suomalaisessa sosiaali- ja terveydenhuollossa. Potilasturvallisuustyössä on olennaista, että pystytään seuraamaan saavutetun potilasturvallisuuden tasoa ja siten tunnistamaan kehittämiskohteita ja seuraamaan tehtyjen toimenpiteiden vaikutuksia. Potilasasiakirjojen manuaaliseen tai automaattiseen analysointiin kehitetyt trigger-työkalut ovat yksi mahdollinen tekniikka, jonka avulla potilasturvallisuutta voidaan kehittää. IHI Global Trigger Tool (GTT) -menetelmällä terveydenhuollon toimintayksikön potilasturvallisuuden tasoa arvioidaan takautuvasti. Menetelmän avulla voidaan tehdä pitkäaikaisvertailuja ja –arviointeja toimeenpantujen potilasturvallisuuteen liittyvien toimenpiteiden toimivuudesta sekä tunnistaa kehittämiskohteita. GTT-menetelmä on paperipohjainen, joten sen hyödyntäminen ei edellytä sähköisen potilasasiakirjajärjestelmän käyttämistä. Tässä raportissa kuvataan suomalaisten sairaaloiden viimeaikaisia kokemuksia GTT:n käyttöönotosta organisaation potilasturvallisuuden seuraamiseksi sekä käsitellään menetelmän jatkokehittämistä varten sairaaloissa tehtyjen kokeilujen tuloksia. Lisäksi arvioidaan sekä paperipohjaisen (manuaalisen) että automaattisen GTT:n vahvuuksia ja heikkouksia ajantasaisen tutkimustiedon perusteella.

GTT ja muut yhdysvaltalaisen Institute for Healthcare Improvement (IHI) -laitoksen kehittämät trigger-työkalut ovat suhteellisen uutta teknologiaa. Vuoden 2011 kesäkuuhun mennessä julkaistut tutkimukset kattavassa kirjallisuuskatsauksessamme (kappale 2) löydettiin vain yhdeksän nimenomaisesti IHI GTT:tä käsittelevää julkaisua, joista useimmat olivat ilmestyneet viimeisen 2–3 vuoden aikana. Artikkeleissa käsiteltiin työkalun kehittämistä, arviointia ja toimivuutta sekä vertailtiin GTT:tä muihin menetelmiin ja kuvattiin työkalun käyttöä käytännössä joko laajoissa terveydenhuollon järjestelmissä tai kansallisen tason ohjelmissa. Tuloksia raportoitiin haittatapahtumien määrinä 100 hoitajaksoa kohden (käytössä viidessä tutkimuksessa, ja näissä tuloksien vaihteluväli oli 18.1–41) ja/tai haittatapahtumien määrinä 1000 hoitopäivää kohden (käytössä neljässä tutkimuksessa, ja näissä tuloksien vaihteluväli oli 41.6–91). Kaikissa tarkastelluissa tutkimuksissa havaittuja haittatapahtumia arvioitiin yhdysvaltalaisen National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) -neuvoston menetelmän avulla. Yli puolessa raportoituista tapauksista haittatapahtuman vaikeusaste kuvattiin tilapäisenä häirtana hoitoa tarvitsevalle potilaalle (E-kategoria). Vain kolmessa kirjallisuuskatsauksessa tarkastellussa tutkimuksessa arvioitiin haittatapahtumien ennaltaehkäistävyyttä, ja näistä vain kahdessa raportoitiin tuloksia. Arvioinneissa käytettiin 4-portaista Likert-asteikkoa ja ehkäistävässä olleiden haittatapahtumien tasoksi raportoitiin 51,7 % ja 63,1 %.

GTT-menetelmään liittyvistä kiistanaiheista kaksi tärkeintä koskevat yhtäältä työkalun luotettavuutta menetelmän puutteellisen validoinnin vuoksi ja toisaalta työkalun soveltuvuutta benchmarking työkaluksi. Kattavaan potilasasiakirjojen tarkastukseen verrattuna GTT vaatii vähemmän resursseja, koska sillä tutkitaan pienempi määrä potilasasiakirjoja, joissa todellisten haittatapahtumalöytöjen todennäköisyys on korkeampi. Sen lisäksi, että GTT:llä voidaan havaita suurempia määriä haittatapahtumia kuin muilla arviointityökaluilla, voidaan sillä vertailevien tutkimusten mukaan myös tunnistaa erityyppisiä haittatapahtumia. Näiden havaintojen perusteella on pääteltävissä, että trigger-työkalujen avulla voidaan täydentää haittatapahtumaraportointia ja muita potilasturvallisuuteen liittyviä toimenpiteitä, kun halutaan saada kattava kuva haittatapahtumien tasosta organisaatiossa. Paperipohjainen GTT on perusteltu vaihtoehto säännölliseen käyttöön paljon resursseja vaativan, tutkimuspainotteisen potilasasiakirjojen kattavan tarkastelun sijaan. GTT:n käyttö näyttääkin yleistyneen ainakin Yhdysvalloissa sekä muissa Pohjoismaissa osana kansallisia potilasturvallisuuteen liittyviä aloitteita tai ohjelmia.

Turun yliopistollisessa keskussairaalassa (VSSHP) haittatapahtumia on arvioitu IHI GTT:n avulla jo vuodesta 2009 (kappale 3). Satunnaisotannassa oli 1335 hoitajaksoa, joista 890 analysoitiin. Potilasturvallisuusprojektissa asetettu 30 %:n vähennystavoite haittatapahtumien määrässä voitiin selkeästi todentaa työkalun käytön avulla. Samalla arvioitiin havaittujen haittatapahtumien vakavuutta ja ennaltaehkäistävyttä. Suurin osa havaituista tapahtumista olivat E-kategorian haittatapahtumia, ja kaikista haittatapahtumista 65,5 % oli ehkäistävissä. Tarkoituksena on ottaa GTT käyttöön myös muilla sairaalan osastoilla. Odotettavissa on, että menetelmän käyttö osastotasolla tuottaa täsmällisempää ja tarkempaa tietoa haittatapahtumista. Tämä tosin edellyttää muiden trigger-työkalujen, kuten päiväkirurgian, lastentautien ja psykiatrian alalle suunniteltujen triggereiden suomentamista ja validointia.

Tampereen yliopistollisen sairaalan (TAYS) neurokirurgian osastolla arvioitiin millä tarkkuudella sähköisistä potilasasiakirjoista tunnistetaan tekstinlouhintamenetelmällä (text mining) ne triggerit, jotka on jo tunnistettu manuaalisessa potilasasiakirja-analyysissä (kappale 4). Tutkimus toteutettiin rakenteisena takautuvana potilasasiakirja-analyysinä, joka perustui 13 modifioituun IHI GTT -seulontakriteeriin. Tutkimuksessa analysoitiin kolmea ICD-10-tautiluokitukseen perustuvaa diagnoosiryhmää (pään vammat, aivoverenvuoto ja verrokkiryhmänä kaularangan sairaudet) kolmen ajallisesti rinnakkaisen edustavan otoksen (n = 556) avulla. Otos perustui yhteensä 1969 neurokirurgian hoitajaksoon huhtikuun 2007 ja toukokuun 2008 välisenä aikana. Manuaalisen analyysin perusteella triggereitä oli 23,7 %:ssa (n = 132) kaikista otoksen hoitajaksoista (n = 556). Tekstinlouhintamenetelmällä triggereiden havaitsemisherkkyys manuaaliseen analyysiin verrattuna vaihteli 60 %:sta 100 %:iin. Menetelmän triggerikohtainen tarkkuus vaihteli 80 %:sta 98 %:iin. Tutkimusryhmä totesi, että triggereitä voidaan havaita myös tekstinlouhintamenetelmällä ja että menetelmä on yhtä luotettava ja vähemmän aikaa ja henkilöstöresursseja vaativa menetelmä kuin perinteinen manuaalinen menetelmä.

Sähköisten potilastietojärjestelmien käytön laajentuessa myös triggereiden sähköisen tunnistamisen keinot kehittyvät, jolloin analyysin tekijän tehtäväksi jää potilaalle aiheutuneen haitan tunnistaminen ja arviointi. Automatisoitujen trigger-työkalujen kirjallisuuskatsauksessa (kappale 5) korostuu yhtäältä potilasasiakirjojen korkeatasoisen dokumentoinnin merkitys ja toisaalta organisaation tietojärjestelmien integrointiratkaisujen tärkeys. Koodatun ja rakenteisen tekstin käyttäminen voisi edistää mahdollisten triggereiden tosiaikaista hyödyntämistä. Tämä olisi ainakin periaatteessa mahdollista suomalaiset vähimmäisvaatimukset täyttävän tietoaisteiston pohjalta, mutta asiasta ei ole tutkimustietoa saatavilla. Ennustavat (prospektiiviset) trigger-järjestelmät ovat riippuvaisia merkittävistä muutoksista työnkulussa, ja siten niiden käyttöönotto voi olla haastavaa. Toinen mahdollinen vaihtoehto triggereiden havaitsemiseksi on sähköisten potilasasiakirjojen vapaan tekstin hyödyntäminen tekstintulohintamenetelmän avulla, kuten tehtiin ruotsalaisessa MAG-sovelluksessa ja suomalaisessa neurokirurgian pilottitutkimuksessa.

Paperipohjaisen GTT:n käytön leviämisestä huolimatta trigger-työkalujen tulevaisuus on kuitenkin tietokonepohjaisissa ratkaisuissa. Automatisoituja trigger-työkaluja on kuitenkin tutkittava ja kehitettävä edelleen, ja niiden käytettävyyttä on arvioitava perusteellisesti. Ala kehittyy jatkuvasti, ja koko ajan tuotetaan uusia hybridijärjestelmiä, joissa yhdistellään erilaisia häirtäpähtumien havainnointimenetelmiä. Nämä sovellukset eivät kuitenkaan ole vielä valmiita laajamittaiseen käyttöön, ja niihin keskittyvät lähinnä muullakin tavoin urauurtavat laitokset. Pohjoismainen yhteistyö etenee niin tutkimuksen kuin käyttöön oton tasolla ja sen panos potilasturvallisuuden seurannan kehittämiseksi voi osoittautua merkittäväksi. On kuitenkin muistettava, että mittaustiedon kerääminen on vasta prosessin ensimmäinen askel, eikä tavoite itsessään. Todellinen potilasturvallisuuden parantaminen edellyttää, että trigger-työkalun käyttöön oton yhteydessä organisaatiossa myös laaditaan toimintasuunnitelma saatavan potilasturvallisuustiedon levittämistä ja hyödyntämistä varten, jotta organisaatiossa saavutettaisiin todellisia muutoksia toimintatavoissa.

Avainsanat: Global Trigger Tool, potilasturvallisuuden seuranta, sairaalat, sähköinen potilasasiakirja, tietokonepohjaiset trigger-työkalut

Sammandrag

Persephone Doupi, Karolina Peltomaa, Mika Kaartinen, Juha Öhman. IHI Global Trigger Tool och övervakning av patientsäkerheten i finska sjukhus. Institutet för hälsa och välfärd (THL). Rapport 19/2013. 95 sidor. Helsingfors, Finland 2013.

ISBN 978-952-245-993-0 (tryckt); ISBN 978-952-245-999-2 (nätpublikation)

Patientsäkerheten ökar i betydelse bland finska hälso- och socialvårdsserviceorganisationer. En viktig del av arbetet med patientsäkerheten är förmågan att övervaka de nivåer av patientsäkerheten som uppnås i syfte att identifiera områden där förbättringar behövs och att följa upp genomslag för utförda ingrepp. Trigger tools – medel för att utföra fokuserade granskningar av vårdjournaler antingen manuellt eller automatiskt – är bland teknikerna som kandiderar för att tjäna ovannämnda syften. IHI Global Trigger Tool (GTT) är en retrospektiv metod för övervakning av nivåer på patientsäkerheten inom en vårdleverantörs organisation. Den medger långsgående jämförelser och bedömning av genomförda patientsäkerhetsåtgärder och möjliggör identifiering av målområden för förbättringar. Metoden är pappersbaserad och med andra ord kräver inte eller är beroende av hälsoinformationssystem. I denna rapport presenteras senare tids erfarenheter av införandet av GTT för organisatorisk övervakning av patientsäkerheten inom den finska sjukhusmiljön, inklusive experiment med vidareutvecklingen av dessa verktyg, och de senaste forskningsfyndigheterna om styrkorna och svagheter hos pappersbaserade och datorbaserade trigger tools granskas.

GTT, liksom resten av IHI:s serie med trigger tools är en relativt ny teknik. Vid vår litteraturgenomgång (kapitel 2), som täcker studier som publicerats fram till juni 2011, kunde bara nio skrifter som specifikt behandlade IHI:s GTT hittas, de flesta publicerade under de senaste 2-3 åren. Artiklarna gällde utvecklingen och utvärderingen av verktyget, dess funktioner, jämförelser med andra metoder och exempel på användning antingen inom eller mellan stora vårdssystem eller i program på nationell nivå. Resultaten redovisades som antalet negativa händelser/100 patienter (använt i fem studier, med resultat som varierade mellan 18,1 och 41) och/eller antalet negativa händelser/1000 patientdagar (använt i fyra studier, med resultat som varierade mellan 41,6 och 91). I alla de granskade studierna bedömdes allvarlighetsgraden hos identifierade händelser med hjälp av National Coordinating Council for Medication Error Reporting and Preventions (NCC MERP) klassificering. I mer än hälften av fallen beskrivs allvarlighetsgraden som tillfällig skada på patienten som kräver ingrepp (kategori E). Förebyggbarheten för negativa händelser bedömdes i bara tre av de granskade studierna och resultaten redovisades bara i två av dem. Bedömningen gjordes med hjälp av en Likert-skala med fyra nivåer, och frekvensen för förebyggbara negativa händelser redovisades som 51,7 % respektive 63,1 %. Det finns två huvudsakliga kontroversiella frågor kring GTT: dess tillförlitlighet som metod på grund av dess begränsade validering och verktygets lämplighet som riktmärkesskapande instrument. Jämfört med fullständig granskning av patientjournaler ställer GTT lägre krav på re-

surser med anledning av att färre journaler, där sannolikheten är större att de innehåller faktiska negativa händelser, granskas. Förutom dess förmåga att upptäcka ett större antal händelser än andra bedömningsmetoder uppger jämförande studier också att GTT kan identifiera olika typer av negativa händelser. Mot bakgrund av dessa observationer kan vi konstatera att användningen av trigger tool-metoden kan utgöra ett komplement till incidentrapportering och andra ingrepp när målet är en omfattande bild av nivån på negativa händelser inom en organisation. GTT:s pappersmetod kan ses som ett gångbart alternativ för regelbunden användning i stället för den resursintensiva och forskningsorienterade metodiken med fullständig granskning av patientjournaler. I själva verket tycks användningen av GTT vara på frammarsch – åtminstone i USA, liksom i andra nordiska länder, som en del av nationella initiativ/program för patientsäkerheten.

I Finlands sydvästra sjukhusdistrikt (TYKS) har bedömning av negativa händelser med IHI GTT börjat redan år 2009 (kapitel 3). Urval har tagits ur 1335 randomiserade inskrivna patienter, varav 890 inkluderades i analysen. En tydlig trend med minskade negativa händelser enligt det fastställda målet på 30 % dokumenterats med hjälp av metoden. Allvarlighetsgrad och förebyggbarhet för de identifierade negativa händelserna har också utvärderats. De flesta av de identifierade händelserna tillhör kategori E och 65,5 % av samtliga händelser var förebyggbar. Avsikten är att fortsätta med införandet av IHI:s GTT på olika sjukhusavdelningar. Det förväntas att användningen av denna metodik på avdelningsnivå ger mer noggrann och detaljerad information. Detta kräver dock också översättningen och valideringen av ytterligare triggers som är relaterade till t.ex. dagkirurgi, pediatrik och psykiatri.

På neurokirurgiska avdelningen på Tammerfors Universitetssjukhus TAYS bedömdes textsökningsens förmåga att korrekt avkänna samma triggers i elektroniska patientjournaler som manuell granskning (kapitel 4). Studien utfördes i form av en strukturerad retrospektiv granskning av patientjournaler grundad på användningen av 13 modifierade screeningkriterier i IHI GTT. Tre olika ICD-baserade diagnosgrupper (huvudskador, intrakraniell blödning och patologi i halsryggraden som kontrollgrupp) analyserades genom tre parallella tidsrepresentativa urval ($n = 556$) ur totalt 1 969 inskrivna neurokirurgipatienter mellan april 2007 och maj 2008. Grundat på manuell granskning uppvisade 23,7 % ($n = 132$) av de 556 inskrivna triggers. Jämfört med manuell granskning varierade känsligheten för att avkänna triggers mellan 60 och 100 % mellan triggers. Specificiteten mellan triggers varierade mellan 80 och 98 %. Studiegruppen drog slutsatsen att triggers kan hittas med textsökningsverktyget och att denna metod är lika tillförlitlig som och mindre tids- och personalkrävande än den konventionella manuella metoden.

Vartefter användningen av elektroniska vårdjournaler ökar blir även medel för den elektroniska identifieringen av triggers mer gångbara och minskar granskarens insats för att identifiera och bedöma skador. Granskning av litteraturen om datoriserade triggerverktyg (kapitel 5) underströk behovet av dokumentering av hög kvalitet i patientjournalen och vikten av dataintegreringslösningar som förbinder olika sjukhus-

informationssystem. Användningen av kodad och strukturerad text kan vara vägen till användningen av presumtiva triggers i realtid och är åtminstone i princip genomförbar grundat på den finska minimidatauppsättningen, men detta har inte undersökts vidare i praktiken. System för presumtiva triggers är dock starkt beroende av signifikanta ändringar i arbetsflödet och möts av betydande hinder för införandet. Användningen av fritext i den elektroniska vårdjournalen av textsökningsverktyg, som i den svenska MAG-tillämpningen och det finska neurokirurgipilotprojektet, är också ett potentiellt alternativ för triggeravkänning.

Även om användningen av den pappersbaserade GTT-metoden för närvarande är på framväxt är datoriserade varianter framtiden. Automatiserade triggerverktyg kräver ytterligare forskning och utveckling kombinerat med rigorös utvärdering. Området befinner sig i ett utvecklingsstadium och jämn framtagning av hybridsystem som kombinerar olika ansatser för avkänning av patientsäkerhetsändelser kan iakttas. De resulterande tillämpningarna är ännu inte mogna för utbredd användning och är för närvarande fokus för organisationer som i övrigt är banbrytande. Nordiskt samarbete framskrider både i fråga om forskning och om införandenivån och kan resultera i viktiga bidrag inom patientsäkerhetsövervakningens område. Det ska dock inte glömmas att insamlingen av data för mätning endast är det första steget i en process och inte i sig det slutliga målet. För att verkliga förbättringar i patientsäkerheten ska uppnås måste även en strategi för spridning och utnyttjande av fyndigheterna från trigger tool-metoderna som syftar till att uppnå verkliga förändringar i klinisk praxis formuleras som en del av införandet av trigger tools.

Nyckelord: Global Trigger Tool, patientsäkerhetsövervakning, sjukhus, elektronisk patientjournal, datoriserade trigger tools.

Contents

Foreword	3
Abstract	4
Tiivistelmä	7
Sammandrag	10
Introduction.....	17
References	18
1 Overview of the IHI Global Trigger Tool: Development and Experiences... 20	
GTT- Description of basic method	21
Review Process	23
Data Collection and Reporting	25
Use of the GTT in national level initiatives: experiences from the other Nordic countries.....	25
Denmark.....	26
Norway	27
Sweden.....	29
References	31
2 GTT: Reviewing the evidence..... 34	
Reviewing methodology	36
Measures of AEs	38
Assessment of severity.....	39
Assessment of preventability	39
Methodological questions: Reviewers' agreement	40
Resources: training and implementation phase.....	42
Implementation questions: the organisational viewpoint	43
Resources needed, Added value of the method	44
Documentation of identified cases – Learning through using the GTT	46
In conclusion	46
References	48
3 Assessment of Adverse Events in the Southwest Hospital District of Finland	50
Preparations before implementation	50
Conducting the analysis.....	51
Preventability of the adverse events	53
Severity of the adverse events.....	53
Results and development from the basis of the analysis.....	54
Conclusions	54
References	55

4	Use of the Global Trigger Tool in neurosurgery: A comparison of manual and text mining detection of triggers	56
	Introduction	56
	Materials and Methods	57
	Review Process	58
	Automatic review and statistical analysis.....	58
	Results	59
	Discussion.....	60
	Conclusions	62
	References	62
5	Computerization of trigger tools: Evidence and possibilities.....	64
	Why use computerized triggers?.....	64
	The Finnish context.....	65
	Literature Review on computerized trigger tools	66
	Methodology	66
	Essential components of computerized trigger tool systems.....	67
	Features of computerized trigger systems.....	67
	Discussion.....	71
	Critical Success Factors.....	72
	The role of data/text mining.....	73
	Future target: Continuum of patient safety monitoring.....	74
	Implications of the review findings for trigger tool activities in Finland	75
	References	76
6	In Conclusion	79
	References	82
	Annex 1. Literature Review Search Strategies.....	83
	Annex 2. Reviewed studies on computerized trigger systems - additional details ..	86
	Annex 3. Triggers and Minimum Data Set (Ydintiedot) for Finnish EPR systems .	91
	Annex 4. IHI GTT Worksheet used in the Southwest Hospital District of Finland patient safety programme	95

Abbreviations

AE	Adverse Event
ADE	Adverse Drug Event
ADR	Adverse Drug Reaction
AHRQ	Agency for Healthcare Research and Quality
AI	Artificial Intelligence
DPS	Danish Patient Safety Database
DRG	Disease Related Groups
EBMeDS	Evidence-Based Medical Decision Support
EHR	Electronic Health Record
EMR	Electronic Medical Record
EPR	Electronic Patient Record
ER	Emergency Room
GTT	Global Trigger Tool
HELP HIS	Health Evaluation through Logical Processing Hospital Information System
HIT	Health Information Technologies
ICD9-CM	International Classification of Diseases 9 – Clinical Modification
IHI	Institute for Healthcare Improvement
LIS	Laboratory Information System
LOS	Length of Stay
OR	Operating Room
PPV	Positive Predictive Value
PSI	Patient Safety Indicators
MAG	Modifierad Automatiserad GTT/Modified Automated GTT
NCC MERP	National Coordinating Council for Medication Error Reporting and Prevention Index
NLP	Natural Language Processing
TAYS	Tampereen Yliopistollinen Sairaala – Tampere University Hospital
VSSHPI	Varsinais Suomen Sairaanhoidopiiri – Hospital District of South-west Finland
SIRO	Finnish Hospital Infection Program
VHA	Veterans Health Affairs

Introduction

Since the publication of the first patient safety strategy in 2009, patient safety activities in Finland have been steadily gaining momentum (1). The coming into force of the long awaited law on Healthcare services (2), including its specific clause on patient safety and quality, and the respective decree have formed a clear legal framework for patient safety activities. At the same time, they have placed healthcare service provider organisations in front of specific requirements to be met. These requirements and the means to achieve them are to be formalized in an organisational patient safety plan, which – among others – involves the follow up and monitoring of patient safety (3).

As a result, in addition to the challenge of drawing up the patient safety plan itself (support for which can be found at the Patient Safety Guide – see Ref.4), organisations and their managers have to urgently come up with a valid answer to the question: how do we measure patient safety?

The answer to this question is neither clear, nor straightforward. A variety of tools are available for the purpose of patient safety measurement and monitoring, each with its own plus and minuses and fitness for use for certain purposes. Among others these include:

- incident reporting systems, either independent ones such as TURPO (Turvalli-suuspoikkeamailmoitukset – Safety deviations reporting) and HaiPro, or built as part of registries such as Hilmo (Hospital Discharge Registry) and SIRO (Hospital infections programme) (5–8)
- measures of patient safety culture (9–11)
- patient safety and quality indicators (12, 13)
- Global Trigger Tool (GTT) and similar tools (14, 15).

The GTT and related ‘trigger’ technologies are the subject of this report. The Global Trigger Tool (14) is a retrospective method for monitoring patient safety levels within a healthcare provider organization. It allows for longitudinal comparisons and assessment of patient safety measures implemented, and it enables the identification of target areas for improvement. The method is paper-based, in other words, it does not require or depend on the use of health information systems. The GTT, as well as the rest of the IHI trigger tools family is a relatively new technology, the use of which seems to be rapidly spreading at least in the US, as well as in other Nordic countries (as is discussed in detail in Chapter 2). Before proceeding with the adoption of a new technology that necessitates specific investments and allocation of resources, it is important to know what evidence exists of its benefits, and what the requirements for successful implementation are. Moreover, it is important to know what the forthcoming technological developments and potential alternatives are that the new technology should be weighed against.

Between 2008 and 2011, the following pertinent experiences have been gained with the IHI GTT in Finland: use of the classic method as part of hospitals' patient safety projects (VSSHP and Vaasa Central Hospital) (16, 17); pilot with adaptation of the GTT in neurosurgery and NICU environments, and experimentation with data mining approaches (18, 19) (Chapter 4); and analysis of the national minimum data set for EHRs to support such applications (20, Annex 3). This report presents and discusses these experiences, against the background of the results emerging from an analysis of the scientific literature on the subject of paper-based and computerized trigger tools and patient safety (Chapters 3, 5 and Annex 2).

References

1. Ministry of Social Affairs and Health, 2009. Finnish Patient Safety Strategy 2009–2013. (In Finnish. Edistämme potilasturvallisuutta yhdessä. Suomalainen potilasturvallisuusstrategia 2009–2013). Available at: http://www.stm.fi/julkaisut/nayta/_julkaisu/1383571.
2. Healthcare Service Act, 30.12.2010/1326. (In Finnish. Terveys- ja hoito- ja lääketieteellisen toiminnan laadunhallinnasta ja potilasturvallisuuden täytäntöönpanon laadittavasta suunnitelmasta http://www.stm.fi/vireilla/lainsaadantohankkeet/sosiaali_ja_terveydenhuolto/laadunhallintasetus
3. Decree of the Ministry of Social Affairs and health for quality management and the implementation of the patient safety plan. Sosiaali- ja terveysministeriön asetus laadunhallinnasta ja potilasturvallisuuden täytäntöönpanon laadittavasta suunnitelmasta http://www.stm.fi/vireilla/lainsaadantohankkeet/sosiaali_ja_terveydenhuolto/laadunhallintasetus
4. Potilasturvallisuus opas – potilasturvallisuus lainsäädännön ja -strategian toimeenpanon tueksi. Opas 15/2011. THL, Helsinki. ISBN 978-952-245-481-2. Available at: <http://www.thl.fi/thl-client/pdfs/b6783c8b-f465-403b-85f7-90f92f4c971f>
5. HILMO – The Finnish Care Register. Information in Finnish: <http://www.stakes.fi/Fl/tilastot/tausta/Rekisteriselosteet/terveydenhuollonhoitoilmoitukset.htm>
6. HaiPro – Reporting System for Safety Incidents in Health Care Organizations. Homepage: <http://www.haipro.fi/eng/default.aspx>
7. TURPO – Basic information available (in Finnish) in: Anneli Pitkänen, Maija-Liisa Lehtonen, Mikko Pilli, Jukka Kuisma et al. Potilaiden väkivallan hallinta Taysin aikuispsykiatrian vastualueella. Pirkanmaan Sairaanhoidopiiri Report, 2/2011, Tampere. ISBN 978-951-667-173-7. Available at: <http://www.tays.fi/download.aspx?ID=16838&GUID=%7B2D222BA3-687C-4B28-B1CE-6AE920078107%7D>
8. SIRO, Finnish Hospital Infection Programme. Homepage (in Finnish): http://www.ktl.fi/portal/suomi/osastot/infe/tutkimus/sairaalainfektioiden_seuranta_siro
9. Patient safety culture measures AHRQ. Information available at: <http://psnet.ahrq.gov/primer.aspx?primerID=5>
10. Singla AK, Kitch BT, Weissman JS, Campbell EG. Assessing Patient Safety Culture: A Review and Synthesis of the Measurement Tools. *J Patient Saf* 2006; 2:105 - 115
11. Patient Safety, Organisational Tools. WHO http://www.who.int/patientsafety/research/methods_measures/human_factors/organizational_tools/en/index.html
12. Patient Safety Indicators. Overview available at: http://www.qualityindicators.ahrq.gov/modules/psi_overview.aspx
13. Health Care Quality Indicators Project – Patient Safety. Information available at: http://www.oecd.org/document/43/0,3746,en_2649_37407_37090539_1_1_1_37407,00.html
14. Griffin FA, Resar RK. IHI Global Trigger Tool for Measuring Adverse Events (Second Edition). IHI Innovation Series white paper. Cambridge, Massachusetts: Institute for Healthcare Improvement; 2009
15. Introduction to trigger tools, Institute for Healthcare Improvement: <http://www.ihl.org/knowledge/Pages/Tools/IntrotoTriggerToolsforIdentifyingAEs.aspx>

16. Peltomaa K, Savunen T, Kantola I, Gröndahl W. VSSHP:n potilasturvallisuusprojekti 2009 – 2011: kokonaisvaltainen kehittämisshanke. In: Honkanen V, Härkönen M, Jylhä V, Saranto K (Eds). Potilasturvallisuustutkimuksen päivät 2011, Esitysten tiivistelmät, Helsinki 2011. p. 33.
17. Kinnunen M, Helovuori A. Kehittämisshanke jatkuvasta ja suunnitelmallisesta potilasturvallisuuden edistämiseksi. In: Honkanen V, Härkönen M, Jylhä V, Saranto K (Eds). Potilasturvallisuustutkimuksen päivät 2011, Esitysten tiivistelmät, Helsinki 2011. p.31.
18. Öhman J, Keski-Kuha T, Kaartinen M, Kujala E. Use of Global Trigger Tool in Detecting Adverse Events in Neurosurgery: Comparison of Manual and Text Mining Detection of Triggers. In: Honkanen V, Härkönen M, Jylhä V, Saranto K (Eds). Potilasturvallisuustutkimuksen päivät 2011, Esitysten tiivistelmät, Helsinki 2011. p. 22.
19. Tammela O. Trigger-työkalu vastasyntyneiden teho-osastolla hättätapahtuman ehkäisemiseksi. In: Honkanen V, Härkönen M, Jylhä V, Saranto K (Eds). Potilasturvallisuustutkimuksen päivät 2011, Esitysten tiivistelmät, Helsinki 2011. p. 20.
20. Doupi P. Ydintietojen soveltuvuus potilasturvallisuuden tukena: tapaus liipaisintyökalut (trigger tools). In: Honkanen V, Härkönen M, Jylhä V, Saranto K (Eds). Potilasturvallisuustutkimuksen päivät 2011, Esitysten tiivistelmät, Helsinki 2011. p. 36.

1 Overview of the IHI Global Trigger Tool: Development and Experiences

*Persephone Doupi, MD, PhD
Karolina Peltomaa, RN, MNsc*

The Institute for Healthcare Improvement built on earlier work on computerized alert detection in order to produce a paper-based (and thus less resource-demanding) means of monitoring patient safety on the organisation level. A distinct feature of the IHI Trigger Tool methodology is its focus on actual harm (restricted to physical injury) inflicted to patients (1, 2). The underlying rationale is that surveillance of events that have led to harm is a more focused and hence more effective approach to developing a strategy for injury reduction (3).

In the course of time, various trigger tools have been developed and tested (see Table 1).

The Global Trigger Tool, just as the other members of the IHI trigger tool family, is a retrospective method for monitoring patient safety levels within a healthcare pro-

Table 1. Overview of IHI Trigger Tools and their development.

Application Area	Timeline	Features
Adverse Drug Events - part of the Idealized Design of the Medication System (4)	1999 (start of development) – 2003 (active use in over 200 organizations)	List of 24 triggers – Supplements incident reports and pharmacy interventions
Use as process specific tool, e.g. warfarin trigger	2002	Helpful for more intensive evaluation of a given process
Customized for use in the outpatient or ambulatory care setting	2002	Focus on 'life events' and their processing by the healthcare system
Pediatric Adverse Drug Event Tool (5)	2002 (collaborative effort)	List of 15 triggers, focused on children's in-hospital adverse drug events
ICU Adverse Event Tool (6)	2002 (collaboration with VHA)	List of 24 triggers, but covers all adverse events, not just medications
General (Global) Trigger Tool (7)	2004 (development start) – 2007 (publication)	Covers the adult inpatient population and is combined with robust review methodology
Surgery Trigger Tool (8)	2006	List of 28 triggers to identify peri-operative adverse events
Neonatal ICU (NICU) Trigger Tool (9)	2006 (collaborative effort)	List of 16 triggers, based on study in North American units

vider organisation. The purpose of the GTT is to establish a baseline level of harm (adverse events) in an organization and then, using statistical process control rules, collect data points over time to determine improvement. In other words, and in accordance with the AHRQ Trigger Terminology Definitions (10) the trigger tools of IHI are **accounting** ones (in the sense that they allow system-level monitoring of patient safety) and **retrospective**, since they identify the adverse event only after the episode of care has been concluded. The results of GTT reviews allow for longitudinal comparisons and assessments of patient safety measures implemented, and enable the identification of target areas for improvement.

During the development and testing of domain-specific trigger tools, one main problem that emerged was the low inter-rater reliability in reviewing patient charts. Hence in the context of developing the GTT, the IHI also invested in producing a more robust and guided approach for the reviewing methodology to be used when implementing the tool (11).

GTT- Description of the basic method

As described in the IHI White paper, the trigger tool methodology is a retrospective review of a random sample of inpatient hospital records using “triggers” (or clues) to identify possible adverse events. Global Trigger Tool requires manual review of closed inpatient records with completed discharge summaries and coding. Note that hospitals should use the IHI Global Trigger Tool as one part of a learning system that includes other component measures, such as voluntarily reported errors, surgical site infections, and other outcome measures.

The review team should consist of, at a minimum, three people: 1) Two primary record reviewers who have clinical backgrounds and knowledge about the contents and layout of the hospital’s record, as well as about how care is generally provided in the hospital; 2) A physician who does not review the records, but authenticates the consensus of the two primary record reviewers. The physician authenticates the findings of the adverse events and the rating of severity, and provides answers to questions the record reviewers have about findings in a specific record.

Using two or more primary record reviewers raises the issue of consistency (intra- and inter-rater reliability) among the reviewers: IHI encourages teams to continually promote consistent, standard record review procedures, use of triggers, and interpretation of events. The review team should remain consistent over time whenever possible. Many hospitals have found that a practical approach is to have a one-year assignment for both primary record reviewers and the physician, with overlapping team members to ensure adequate training of new reviewers.

The recommendation is to sample 10 patient records from the entire population of discharged adult patients (with some exclusions noted below) every two weeks (for example, patients discharged between the 1st and 15th of the month for the first two-

week sample and between the 16th and end of the month for the second two-week sample), for a total of 20 records per month. Some hospitals may elect to review all 20 records at one time monthly, but this generates only one data point per month versus the two data points with the two-week samples. Organizations that have the resources to do so may elect to choose a larger sample size, such as 40 records per month, but reviewing more than 40 records per month accrues little additional benefit. Data from these small samples may show wide variation from sample to sample. However, aggregating these samples over time will increase precision, and plotting this data on control charts will give you useful information about trends and special causes of variation in harm in the organization.

When looking for adverse events from somatic adult patient population, the following guidelines should be used when selecting the records:

- Select 10 patient records for each two-week sampling period, plus a few extra records in case a record is discovered that does not meet review criteria.
- Do not select 20 records for the entire month and divide these 20 records into two 10-record samples; draw each two-week sample independently.

The selection criteria are:

- Closed and completed record (discharge summary and all coding is complete)
- Length of stay at least 24 hours and formally admitted to the hospital (Note: This is a sampling strategy to avoid outpatient cases. Some hospitals include patients with a one-day stay; however, selected records for review should always be patients that have been classified as inpatient.)
- Patient age 18 years or older
- Excluding inpatient psychiatric and rehabilitation patients (Note: Triggers are not defined for these populations in this tool.)

If a record does not meet these criteria, it should not be used; instead, one of the “extras” should be chosen. The team will review only 10 records, so the extra records will not be used unless this occurs.

If possible, records of hospital admissions before and after the index record (i.e., the record selected for review) should be made available. The team can consult these to determine cause for admission or readmission. A complete review with the IHI Global Trigger Tool should not be done on these records, only the record identified in the random sample; the additional records should only be used to investigate the trigger associated with readmission, which should take less than 5 minutes.

Review Process

The two primary record reviewers should each review all records independently. During the review, the physician should be available to answer questions that may arise. Use the following process for the review:

The IHI Global Trigger Tool contains six “modules,” or groupings of triggers. Four of the groupings are designed to reflect adverse events that commonly occur in a particular unit; the Cares and Medication groupings are designed to reflect adverse events that can occur anywhere in the hospital. The six modules are:

- Cares
- Medication
- Surgical
- Intensive Care
- Perinatal
- Emergency Department

All patient records should be reviewed for the triggers in the Cares and Medication modules. The other modules should only be used if applicable; for example, the Intensive Care module should be used when reviewing a record for a patient who spent any part of the hospital stay in an intensive care unit.

The record should only be reviewed to look for the presence of triggers, not to read the record from front page to back page. Experienced reviewers have found the following sections of the record most useful when reviewed in this order:

- Discharge codes, particularly infections, complications, or certain diagnoses
- Discharge summary (look for the specifics of assessment and treatment during the hospital stay)
- Medications administration record
- Laboratory results
- Prescriber orders
- Operative record
- Nursing notes
- Physician progress notes
- If time permits, any other areas of the record (such as History & Physical, Consult notes, or Emergency Department notes)

A 20-minute limit should be set for review of each patient record, once the training period for reviewers has been completed. The “20-minute rule” applies to all records regardless of size.

It is unlikely that all the events in the larger record will be identified since 20 minutes will not be sufficient time to adequately review the entire record using the Trigger Tool technique. It is important to note that the IHI Global Trigger Tool is not meant to identify every single adverse event in a record. The review time limitation and random

selection of records are designed to produce a sampling approach that is sufficient for the design of safety work in the hospital.

A “positive trigger” indicates only the presence of a trigger, not necessarily an adverse event. When a positive trigger is found, review only the pertinent portions of the record. The focused review will determine whether an adverse event has occurred. If no adverse event is found, the reviewer should then move on and look for other triggers. Reviewers will find many positive triggers, but will identify many fewer adverse events. Occasionally, reviewers will find adverse events with no antecedent trigger. Include these events.

When a reviewer identifies a positive trigger, the reviewer should check other relevant portions of the record such as progress notes and orders that were documented in close proximity to the occurrence of the trigger. Documentation that the patient experienced harm from medical care should be present for an adverse event. For example, an INR level greater than 6 would be a positive trigger. The reviewer should look for documentation of bleeding or decreased haemoglobin with need for transfusion and other adverse events that can result from over-anticoagulation. In determining whether an adverse event has occurred, consider that an adverse event is defined as unintended harm to a patient *from the viewpoint of the patient*.

An adverse event that is present on admission to the hospital should be included, provided that it meets the definition of being harm related to medical care. All such adverse events are counted because the measure is what the patient experienced, not what happened within the hospital. Field experience has shown that fewer than 10 percent of all harms that are detected by the IHI Global Trigger Tool will be present on admission. It is useful to keep track of which events occurred outside the hospital so that this can be noted when reporting data. Such data may indicate an opportunity to collaborate with other organisations to improve patient safety, even if the events did not result from hospital care itself.

Once reviewers have determined that an adverse event has occurred, assign a category of harm as follows (Based on National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) Index for Categorizing Errors.):

- Category E: Temporary harm to the patient and required intervention
- Category F: Temporary harm to the patient and required initial or prolonged hospitalization
- Category G: Permanent patient harm
- Category H: Intervention required to sustain life
- Category I: Patient death

These categories are not progressive (i.e., an event does not have to first meet the definition of E and F before it can be categorized as G). For category E, some intervention is required. As an example from category F is a patient, who has had surgery and after having been discharged returns back to the hospital with a postoperative wound infec-

tion. The criteria for category G would be filled if during an elective operation patient suffers from permanent nerve damage due to a laceration.

Data Collection and Reporting

The two-week data collections should initially be presented in three ways:

- Adverse events per 1,000 patient days;
- Adverse events per 100 admissions; and
- Percent of admissions with an adverse event.

Each method has certain advantages. “Adverse events per 1,000 patient days” is the traditional measure and is the recommended measure to track the harm rate over time. Data should be presented in a run chart with “adverse events per 1,000 patient days” on the Y-axis and time in two-week increments on the X-axis. “Adverse events per 100 admissions” is an alternative presentation of rate. It provides a more easily understood representation of harm for leadership. Data should be presented in a run chart similar to “adverse events per 1,000 patient days.” Note that the conversion from “adverse events per 1,000 days” to “adverse events per 100 admissions” simply entails a switch from number of patient days (1,000) to records reviewed (admissions). “Percent of admissions with an adverse event” is a convenient way to present the information to lay leadership, although it diminishes the number of events because some patients may have more than one adverse event during a hospital stay.

In addition to the run chart representations, the team should present categories of harm in a bar chart depicting the volume of harm in each category (E through I). Data is also often presented by type of adverse events. The types of events have commonly been defined as infections, medications, and procedural complications. Hospitals have found this categorization to be useful in prioritizing areas for improvement work. It is also helpful to include a category in the bar chart of those events that occurred prior to admission and were present on arrival. These should not be excluded from the rate or percentage data in run charts, but it will be important to leadership to see the occurrence of these events.

Use of the GTT in national level initiatives: experiences from the other Nordic countries

Experiences with the actual implementation of the GTT have grown during the last two to three years. In addition to its implementation in the assessment of various US health service systems (results of which are presented as part of the literature review in Chapter 2), the GTT has also been one of the harm detection methods employed to estimate the national incidence of adverse events for hospitalized Medicare bene-

ficiaries (12). It has also been used in Scotland as part of the National Patient Safety Programme¹ (13). First the tool was used locally in Ninewells Hospital near Dundee, (which achieved a reduction of patient harm by 60% during three years). Later on, and as part of the five-year goals of the Scottish Patient Safety Programme, the aim was set for a 30% reduction of hospital adverse events, as measured by the GTT.

The GTT has also been tested for applicability in Thailand (see details of study in Chapter 2), in the context of the National Patient Safety Programme monitoring. We present here in more detail the experiences from Nordic countries other than Finland – namely Denmark, Norway and Sweden, which in the last five years have taken a very active role in extensive GTT implementation as part of national patient safety programmes.

Denmark

In Denmark, first experiences with the GTT method were gained in a project by the Danish Cancer Society which aimed to assess the risks of hospitalized cancer patients in the country. The researchers used a combination of two methods: a GTT-based review of 527 patient records and analysis of patient safety events sent to the Danish Patient Safety Database. They found that each method captured different types of adverse events and concluded that combination of different approaches is needed in order to get as full as possible a picture of causes of harm (14, 15).

A much larger project was undertaken in 2008 (16) with hospital-level implementation and piloting of the tool. At that point, several years since the start of adverse event reporting to the Danish Patient Safety Database (DPSD), it was not yet possible to assess the extent by which patient safety promotion efforts had actually resulted in a reduction of the number of patient injuries. The GTT was viewed as a validated tool that could be utilized to illustrate the magnitude of iatrogenic injuries. The Danish version of the method was produced through translation and adaptation to Danish conditions of the IHI original paper and its Swedish version. A clinical expert customized triggers to reflect more appropriately areas such as Danish laboratory values and clinical practices. A GTT learning kit was sent to all hospitals in January 2009.

The Danish hospital study reached the following main conclusions (16):

- Experiments should be made both with internal and external reviewers
- Reproducibility of results can be achieved through good training, use of the same review team, being true to the methodology, and maintaining a log of harm cases and discussions

¹ NHS Scotland (formed in 1999) has 158,000 staff, incl. more than 47500 nurses, midwives, visiting health professionals and more than 3800 specialists. In addition, there are more than 12000 doctors, GPs and allied health professionals (incl. dentists and pharmacists).

- It would be advisable to departments and/or hospitals to receive assistance to the analysis task by an experienced statistician (perhaps through establishment of a hotline function at hospital or regional level to assist departments with the final processing of data)
- It is very important that the tool is continuously developed at the national level, validated and assessed in an international environment - Nordic Cooperation on GTT is central in this context.
- As the electronic health record is implemented, triggers could be found electronically, so that the reviewer's task would involve only identification and assessment of any harm.

In the framework of the Safer Hospital Initiative, which is a collaboration between the Danish Society for Patient Safety, the TrygFonden Foundation and the IHI, targets of 15% reduction in 30-day mortality and 30% reduction in unintended harm (as measured by the GTT) were set in 2010 (17). Five geographically distributed hospitals are participating to the initiative. As part of the quality strategy for 2011-2014 the Center for Quality in South Denmark made the decision to systematically apply the GTT in all hospital units. Presently, the GTT material is undergoing revision in a collaboration between the Danish Society for Patient Safety and the Southern regions, while an automatic trigger search (ATS) project has been launched by the Region of Southern Denmark and SAS Institute Denmark (18).

Norway

In Norway, use of the GTT begun in a small scale from Akershus hospital, where it was combined with patient safety culture measurements. During the period of January-May 2007 the Akershus University Hospital's Quality Department checked the records of a random sample of 481 patient journals in four of the hospital's departments by using the IHI GTT method. The hospital has 500 somatic (and 200 psychiatric) beds, 4200 employees, and an annual budget of 2.500.000.000 NOK (approximately 450 million US\$). It serves a population of 280 000 people, treats 53.000 in-patients and provides 150.000 out-patient consultations annually. Most in-patients (85%) are unscheduled emergency admissions. The departments' percentage of patients whose records documented that they experienced an adverse event during hospitalization correlated strongly (except for the factor 'Stress recognition') with the departments' average staff SAQ-factor scores (SAQ: Safety Attitudes Questionnaire). Due to the very low number of departments studied, only one of the correlation coefficients was significant at the 0.05-level (19).

The National patient safety campaign

Patient safety activities in Norway really took off when the Norwegian patient safety campaign, *In Safe Hands*, was launched in January 2011 by the Ministry of Health (20). The three-year campaign aims to reduce patient harm, and involves both specialist and primary healthcare services. A steering group, led by the leader of the National Health Directorate, is responsible for all key decisions in the campaign. The campaign secretariat forms part of the National Unit for Patient Safety, positioned in the National Knowledge Center for the Health Services.

In the context of the campaign, specific measures are to be introduced in selected areas throughout the health care services, with the goal of reducing the number of patient injuries. The areas are selected by experts on the basis of their potential for improvement, and include e.g. measures for safer drug use and measures to combat infections. Leadership is a recurrent theme. Some of the campaign's priorities have been piloted and are ready for wider distribution in health and care services. As of 2012, the health and care services have a statutory duty of systematic work with quality improvement and patient safety.

The responsibility of the hospitals with regard to serious events has increased, and a special emergency unit of the Board of Health has been established, in an effort to improve learning from adverse events. The government will also follow up with a white paper on quality and patient safety, to be presented in the course of 2012 (21).

One of the missions of the campaign is to uncover the extent of patient harm in Norwegian healthcare. The first step is a national review of patient records in order to achieve an overview of patient harm in the country. Throughout the campaign, all hospital trusts will continue to conduct review of patient records using the Global Trigger Tool (GTT), as a means of detecting patient harm. The figures will be used to monitor the improvement of each individual healthcare provider organisation, rather than compare hospitals (22).

First preliminary results reported in the fall of 2011– after data was submitted from 11 out of 19 health authorities– indicated that 14.5 percent of patient stays in 2010 was associated with an injury (23). The causes of the most common injuries had not yet been analysed. Over six percent of the injuries has led to prolonged hospital stay, and an estimated 0.6 percent of patients died as a result of injury. It is important to note, that the results indicate the number of patients harmed, regardless of whether there has been error or not. This means that the survey did not distinguish between injuries that are preventable, and those that cannot be prevented.

The official report was published in December 2011², presenting the final official results of the first year of national GTT use (24). 18 out of 19 trusts and five pri-

2 The second national GTT analysis report, which covered the data of 2011 and arrived at very similar results, was published in March 2013 (Deilkås, ET. Rapport for Nasjonal Journalundersøkelse med Global Trigger Tool 2011. Rapport fra kunnskapssenteret. Oslo: Nasjonalt kunnskapssenter for helsetjenesten, 2013).

vate hospitals submitted results. A total of 39 GTT teams reviewed the medical records from minimum 200 randomly selected hospital admissions of patients that had been discharged between March 1st and December 31st of 2010. Records of 7819 admissions were reviewed.

- 16 % of the hospital admissions included at least one adverse event (min 3, 5 %-max 38 %).
- 7 % of the hospital admissions included at least one adverse event that led to prolonged hospitalization (min 2 %-max 18 %).
- 1 % of the hospital admissions included at least one adverse event that gave the patient permanent harm (min 0 %- max 3 %).
- 0,66 % of the hospital admissions involved patient harm that led to death (min 0% - max 2 %).
- A total of 8,9 % of the admissions involved an adverse event that led to prolonged hospitalization or more serious consequences (F to I categories) (min 2,5 %- max 21 %).

Sweden

In Sweden, a large scale retrospective medical record review was undertaken in 2008, to establish the national rate of adverse events in hospitalized patients (25). The methodology was primarily based on the Harvard Medical Practice Study and its subsequent modifications. A type of ‘triggers’ was used in the form of screening criteria by which nurses performed the first stage identification of records for further, in-depth review. The researchers concluded that in order to utilize this criteria list for clinical purposes, there would probably need to be a revision.

Under an initiative of the regions of Östergötland, Kalmar and Jönköping, a Swedish translation of the IHI Global Trigger Tool was made available for the first time in 2005. The Swedish version of the tool includes the evaluation of injury preventability (which is not recommended by IHI or included in the Danish version, since it was found difficult and unnecessary) (26).

Later, in 2008–2009, and as a part of the national patient safety initiative, specific practical steps were taken to broaden the use of the GTT (27):

- Assignment of budget share for GTT put into contracts with health care providers
- Acute hospitals were required to start using the manual GTT method during the period 2009-2011
- In 2010, training in use of the manual GTT method was given to all acute care hospitals
- Proposal for audit in all surgical specialties as of 2011
- In the work of the period 2012–2015 tightened requirements for record-reviewing.

By the beginning of 2012, the situation had evolved as follows (28 - 30): out of 21 country councils, 11 had already taken up monitoring using the classic GTT method, and another 5 were planned to begin within 2012. Specifically in the area of Stockholm, all acute hospitals use GTT as of 2012, and they are also offered support (in the form of coverage of the software/technological investments) for adoption of the computerized version (referred to as MAG, more details below).

Automation of the GTT: the Swedish MAG

- Semi-automated method, combining computerized identification of triggers and generation of a patient's risk profile with manual selection & review of the record.
- Used first in analysing cases of patient deaths (all cases in Karolinska hospital in 2008, and a subset of Neurosurgery records in 2009)
- All acute care hospitals were to obtain by early fall 2011 support for the integration of the automated tool in their electronic patient record systems.
- Special emphasis to be placed on the follow-up of tool use in paediatrics

According to information provided by members of the Swedish MAG project team (29):

The functional model of the system took 6 months to develop, and another 9 months for improvements.

A wide and knowledge support group was overseeing and assisting the development process. The participants of the support group were:

- National reference group (for clarifying the definition of triggers, in order to make them machine understandable)
- Stockholm quality reference group
- Stockholm's healthcare administration (HSF)
- Karolinska University Hospital department of Quality and Patient Safety
- IT department of the hospital
- BearingPoint (project management)
- SAS (analysis and data mining software).

A particular feature of the MAG application (in addition to the combination of trigger overview, adverse event reporting and injury follow up) is the creation of a risk profile on the basis of historical data; the calculation of probability of injury or preventable injury for a certain patient case.

The system is at present used retrospectively, but the vision is that it would have the potential to also provide real time assessment of patient data and thus alert the care team of potential high risk of an incident.

Data is made available to the reviewing physician through the MAG portal, which integrates the three functions mentioned earlier (trigger overview, adverse event re-

porting and injury follow up). There are different levels of reporting – department, hospital and county level.

It is important to know that Karolinska University Hospital, which has 8 divisions and 69 departments, has its own (developed in house) electronic medical record system (named TakeCare), as well as a data warehouse, where all data from the hospitals various departments are collected. Still, there have been problems, primarily due to the quality and availability of data, e.g. anaesthesiology services still do all their documentation on paper, hence triggers that require anaesthesiology data cannot be detected.

Karolinska University Hospital undertook training and roll out of the paper GTT in 2010, and at present there are GTT teams in 31 out of 35 hospital departments. During 2011, all surgical departments used the automated model for the review of 20 patient records per month. Roll out continues with the focus on patient safety improvement activities, such as healthcare associated infections.

Lessons learned:

- Thorough preparation and analysis are needed
- Involve and validate on all levels of care
- Clear and frequent communication with all actors
- Early and detailed resource planning
- Time for testing
- Availability of technical environment and resources
- National focus (in order to produce a solution generic enough so that it can be more easily transferred to other locations; technically that would require -among others- the development of ‘adapters’ to interface with every location’s EHR systems).

At present, the system does not cover outpatients or day surgery patient cases (although it is acknowledged that their numbers will be increasing in the future).

References

1. Griffin FA, Resar RK. IHI Global Trigger Tool for Measuring Adverse Events (Second Edition). IHI Innovation Series white paper. Cambridge, Massachusetts: Institute for Healthcare Improvement; 2009
2. Improvement Tip: Focus on Harm, Not Errors. Institute for Healthcare Improvement. Available at: <http://www.ihl.org/knowledge/Pages/ImprovementStories/ImprovementTipFocusonHarmNotErrors.aspx>
3. Resar RK, Rozich JD, Classen D. Methodology and rationale for the measurement of harm with trigger tools. *Qual Saf Health Care* 2003;12:39-45.
4. Adverse Drug Event Tool
5. Takata, GS, Mason W, Taketomo C, Logsdon T, Sharek PJ. Development, Testing, and Findings of a Pediatric-Focused Trigger Tool to Identify Medication-Related Harm in US Children’s Hospitals, *PEDIATRICS* 2008; 121(4), pp. e927-e935 (doi:10.1542/peds.2007-1779)
6. Intensive Care Unit (ICU) Adverse Event Trigger Tool, VHA/IHI, January 2002. Available for download at: <http://www.ihl.org>

- ihi.org/IHI/Topics/CriticalCare/Intensive-Care/Tools/ICUAdverseEventTriggerTool.htm
7. Griffin FA, Resar RK. IHI Global Trigger Tool for Measuring Adverse Events. IHI Innovation Series white paper. Cambridge, MA: Institute for Healthcare Improvement; 2007.
 8. Surgical Trigger Tool for measuring Perioperative Adverse Events. IHI, August 2006. Available for download at: <http://www.ihl.org/IHI/Topics/PatientSafety/SafetyGeneral/Tools/SurgicalTriggerTool.htm>
 9. Sharek PJ, Horbar JD, Mason W, Bisarya H, Thurm CW et al. Adverse Events in the Neonatal Intensive Care Unit: Developing, Testing and Findings of a NICU-Focused Trigger Tool to Identify Harm in North American NICUs. *Pediatrics* 2006;118:1332-1340.
 10. Triggers and Targeted Injury Detection Systems (TIDS) Expert Panel Meeting. Conference Summary. AHRQ, July 2008. Available at: <http://www.ahrq.gov/qual/triggers/triggers.pdf>
 11. Classen DC, Lloyd RC, Provost L, Griffin FA, Resar R. Development and Evaluation of the Institute for Healthcare Improvement Global Trigger Tool. *J Pat Safety*, 2008; 4(3):169-177
 12. Adverse events in hospitals: national incidence among Medicare beneficiaries. Department of Health and Human Services. Office of inspector general, OEI-06-09-00090, November 2010.
 13. Haraden C, Leitch J. Scotland's Successful National Approach to Improving Patient Safety in Acute Care. *Health Affairs* 2011; 30(4): 755-763
 14. Knudsen, JL. Patient Safety in Danish Cancer Treatment, Presentation at the IS-QuA 2009 Conference. Available at: http://www.isqua.org/Uploads/Conference/Abstracts/Microsoft_PowerPoint_-_C10_JL_Knudsen_594_isqua_2009_patient_safety_%5BCompatibility_Mode%5D.pdf
 15. Nielsen ML, Medicine room project and introduction to Global Trigger Tool in Denmark. Presentation at Meeting no. 2 of the Network for the Prevention of Medication Errors, 30 October 2008, Danish Medicines Agency. Available at: <http://www.dkma.dk/1024/visUKLSArtikel.asp?artikelID=15093>
 16. Center for Quality, Region Southern Denmark. Report: Systematisk journal-gennemgang med IHI's GLOBAL TRIGGER TOOL. Erfaringer og resultater fra afprøvning på danske sygehuse. December 2008 (In Danish). Available at: <http://www.centerforkvalitet.dk/wm293462>
 17. Lilja B. What is happening in patient safety in Denmark? Presentation at the 1st Nordic Patient Safety Conference, Stockholm, May 2010.
 18. GTT in Denmark. Presentation at the Nordic meeting, 10.1.2012, Stockholm.
 19. Deilkås ET, Hofoss D. Psychometric properties of the Norwegian version of the Safety Attitudes Questionnaire (SAQ), Generic version (Short Form 2006). *BMC Health Services Research* 2008, 8:191 doi:10.1186/1472-6963-8-191.
 20. The National Patient Safety Campaign of Norway 'In Safe Hands'. Homepage at; <http://www.pasientsikkerhetskampanjen.no/no/I+trygge+hender>
 21. Flesland Ø. Patient Safety and IT in Norway. Presentation, Round Table on Patient Safety and Health IT, September 2011, Helsinki, Finland.
 22. GTT baseline mapping in Norway 2010. Presentation at the National Patient Safety Conference in September 28-29 2010). Available at: <http://www.helse-nord.no/getfile.php/RHF/Konferanser/Ellen%20Deilk%C3%A5s%20-%20110928%20%20Bod%C3%B8%20pasientsikkerhetskonsferansen.pdf>
 23. Preliminary results of national baseline measurements in Norway.. Available at: <http://www.pasientsikkerhetskampanjen.no/no/I+trygge+hender/Kampanjenytt/Over+14+prosent+av+pasientene++skades.493.cms>
 24. Deilkås, ET. Report from a national medical record review using the Global Trigger Tool. Report from Kunnskapsenteret no. [XX]–2010. Oslo: Norwegian Knowledge Centre for the Health Services.
 25. Soop M, Fryksmark U, Köster M, Haglund B. The incidence of adverse events in Swedish hospitals: a retrospective medical record review study. *Int J Q Health Care* 2009;21(4):285-291
 26. Strukturerad journalgranskning för att identifiera och mäta förekomst av skador i vården enligt metoden Global Trigger Tool. Measuring Harm in Healthcare according to the GTT method – Swedish Translation and adaptation GTT Handbook, 2007.

27. Presentations in GTT seminar, Stockholm, December 2010
28. Rutberg H. Patient Safety in Sweden. National Initiatives and new legislation. Keynote, 2nd National Patient Safety Conference, September 2011, Helsinki, Finland.
29. Yhde P, Bodin N. Case: Karolinska University Hospital. Presentation at GTT seminar, Finnish Medical Association Conference January 2012, Helsinki, Finland.
30. Personal communication with Dr. H. Rutberg, February 2012

2 GTT: Reviewing the evidence

Persephone Doupi, MD, PhD

The Global Trigger Tool is a relatively new methodology. The White Paper of IHI presenting the GTT method was published in 2007 and updated in 2009 (1), while the first scientific publication on the development and evaluation of the tool appeared in 2008 (2). Publications around the experiences gained from the practical implementation of the GTT in everyday practice are limited, and have mostly appeared in the last 2–3 years. In addition, the tool’s developers have played a key role in the execution and reporting of many of the studies, which in turn raises the issue of adequate objectivity (3).

Our review of the literature used a combination of formal searches of Ovid Medline and Evidence Based Medicine databases, with other methods and channels of materials identification, and covers literature published until mid-2011 (Further details regarding the search strategy can be found in Chapter 5 and Annex 1). Although at first glance the amount of trigger tool-related literature appeared to be substantial, our focus was exclusively on the GTT. Of 122 candidate publications³, only the ones that dealt specifically with the testing, use and/or evaluation of the Global Trigger Tool were analysed in more detail – in total 9 papers (2–10). These GTT-specific publications concerned the tool’s development and evaluation (1 study), performance features (1 study – focus on different type of reviewer teams), comparisons with other methods (namely AHRQ Patient Safety Indicators and organisations’ voluntary incident reporting systems – 2 studies) and examples of utilization either within or across large health systems or on national level programmes (4 studies). Excluded were: white papers, as well as papers that focused on the development, testing and implementation of other trigger tools, such as those for paediatric, surgical, adverse drug events (ADE) or primary care use.

Our results are compatible with those of two other studies reviewing trigger tools at large (11, 12). It can be hence concluded that the body of evidence regarding the use of the GTT methodology remains still rather limited.

³ Combined results of 2010 and 2011 searches, including the results on computerized triggers

Table 2. Characteristics of reviewed studies concerning the IHI Global TriggerTool

Article	Focus	Comparison	Country	Setting/Population
Classen 2008	GTT-refinement, assessment of improved inter-rater reliability after training	None	USA	15 training records from 2 hospital systems in the Midwestern United States containing 22 adverse events (standardized training exercise); 50 records from the same 2 hospital systems, containing 49 adverse events (testing phase).
Naessens 2009	GTT - Comparative assessment	PSI, Provider reporting	USA	Mayo Clinic Rochester hospitals. Retrospective cross-sectional study. All patients discharged in 2005 (n=60599, incl. deaths) Additional information on patient demographics, LOS, diagnoses and procedures was obtained from administrative data systems. 235 patient records reviewed with GTT.
Asavaroengchai 2009	Evaluation of GTT's effectiveness in identifying adverse events	None	Thailand	Cross-sectional medical record review of patients hospitalized at King Chulalongkorn Memorial Hospital during January 2008. Total of 576 records reviewed.
Landrigan 2010	Temporal trends in patient harm assessed by the GTT	None	USA	Retrospective comparison of 10 hospitals in North Carolina. Randomly selected medical records of patients discharged between January 2002 and December 2007. Internal reviewer team completed 2341 of 2400 reviews planned (97,5%). External teams completed 2374 of the 2400 record reviews planned (98,9%).
Naessens 2010	Determination of GTT's inter-rater reliability in a practice setting; exploration of the value of individual triggers	None	USA	Prospective assessment of application of the GTT to monthly random samples of hospitalized patients at four hospitals across three regions in the USA. Mayo Clinic campuses are in Minnesota, Arizona and Florida. A total of 1138 non-pediatric inpatients from all units across the hospital.
Carter 2010	GTT use in Luton & Dunstable Hospital	None	UK	Random audits from January 2005 to July 2009 (20 case-notes per month) after 50 baseline case-notes. Total of 1130 case-notes.
Good 2011	GTT deployment	None	USA	Integrated Baylor Healthcare System in North Texas: eight general acute care hospitals, two inpatient cardiovascular hospitals and two rehabilitation/long term acute care hospitals between July 2006 and June 2007, for a total of 2369 admissions reviewed.
Classen 2011	GTT – Comparative assessment	PSI, voluntary reporting	USA	Inpatients in three leading hospitals (part of large health systems) with advanced patient safety programs. Randomly selected study patients from all adult inpatients admitted during October 2004. Total of 795 patient records reviewed.
Sharek 2011	Assessment of GTT inter-rater reliability and trigger performance	Different type of reviewer teams (hospital-based versus externally hired)	USA	A retrospective study in a stratified random sample of 10 North Carolina hospitals. Acute care for adult inpatients. 10 randomly selected records per hospital in each quarter from January 2002 to December 2007. Almost 2400 patient records reviewed.

Objective

The focus of the reviewed studies varied, and covered the following areas: development and evaluation of the tool (1 study), performance features (2 studies), comparisons with other methods (2 studies) and examples of utilization either within, or across large health systems or on national level programmes (4 studies).

Design

All reviewed studies concerned retrospective, cross-sectional studies of medical records of discharged patients. The duration of the analysis period varied from one to six years except in the study from Thailand, where follow up concerned only one month of hospitalizations.

Setting and Population

All the reviewed studies, with the exception of one study in Thailand and one in the UK, originate from the United States and concern adult acute and/or long term in-patient care (number of participating hospitals per study ranging from 1 to 10). The number of records reviewed by the GTT method varied from 65 to about 2400.

Reviewing methodology

In all 9 studies the basic pattern of the two-stage review was followed, according to the tradition of the Harvard Medical Practice Study (13). Typically, the primary (first stage) reviewers – those who scan the selected sample of patient records for presence of triggers – are not physicians, but mostly nurses and pharmacists. Physicians act then as the secondary (second stage) reviewers, who make the final decision as to the presence or absence of an adverse event, its severity and potential preventability. The size of the reviewing team may vary, as well as the way of recording and presenting the results. Teams may consist of internal reviewers, i.e. staff members of the organisation being studied, or external reviewers, - clinicians not related to the organisation whose data is being analysed.

Table 3. Overview of review methodology in studies using the IHI Global Trigger Tool

Study	Team Composition	Review Process
Classen 2008	4 primary reviewers (2 pharmacists, a respiratory therapist, and a nurse) and 2 secondary reviewers (a pulmonologist and a family physician)	2-stage review approach with initial review by 4 clinicians followed by a second review by 2 physicians and then a consensus process.
Naessens 2009	Qualified nurses. All adverse events confirmed by a physician	Qualified nurses determine whether a trigger word or condition was present, and, if so, whether an event with harm was associated with the trigger. Then level of harm resulting from the event was assigned. All adverse events were confirmed by a physician. No attempt was made to assess whether the adverse event was preventable.
Asavaroengchai 2009	Two registered nurses. Physician consultant provided to assist the reviewers' decision	Records were reviewed for the presence of triggers. If a trigger was found, the focus review would determine whether an adverse event had occurred. In each record, it was possible to find more than one trigger and also more than one adverse event.
Landrigan 2010	Both by a team of hospital-based (internal) reviewers, who worked in the hospitals where they reviewed charts, and a team of external reviewers, who worked elsewhere (hired and supervised by Batelle). Primary reviewers, typically nurses. Secondary reviewers: physicians with expertise in hospital care	Independently conducted two-stage reviews of the same records in each hospital. Within each team, a primary reviewer conducted a review of each record using the trigger tool. Primary reviewers prepared one- to two-paragraph summaries of all suspected harms, which were presented in a second stage to two independent physician reviewers, who were likewise unaware of dates of hospitalization. The physician reviewers made final determinations about the presence, severity, and preventability of any suspected harms identified. Cases in which physician reviewers disagreed were discussed and consensus was achieved. Inter-rater reliability was calculated from pre-discussion ratings.
Naessens 2010	Two registered nurses, with physician review for confirmation. Web-based data capture tool (Mayo Global Trigger Toolkit) developed not only to allow the uniform collection of data between each of the three sites, but also to streamline the review workflow, reconciliation process and data capture of all evaluations performed by reviewers	Brief review lasting no more than 20 minutes was conducted independently by each of two trained nurses available from a dedicated pool, for the presence of any of 55 'triggers' using the IHI GTT. The occurrence of an adverse event was determined and in cases where an adverse event was identified, the level of harm to the patient was assessed. A physician reconciled the independently recorded triggers and adverse events by each nurse, after assessing the presence of the identified triggers, occurrence of adverse events and the level of harm. Complete results for chart reviews by two nurse reviewers and a final reconciliation including a physician were required for inclusion in the study.
Carter 2010	Triggers, harm events and severity scores written on GTT forms by anesthetist reviewer, subsequently checked by nurse reviewer	Disagreements between reviewers discussed before data entry on an Excel sheet, which included: LOS, triggers, harm events, severity harm scores.
Good 2011	External professional nurse auditors	MS Access Tool developed for this initiative. A structured narrative description of each identified adverse events facilitated text mining to further characterize adverse events.
Classen 2011	Four non-physician reviewers ("primary reviewers") and two physician reviewers ("secondary reviewers")	A physician-led local review team independently conducted a complete detailed review of all hospital records for patients included in the study at one hospital (hospital A) and a review of all clinical, financial, administrative, electronic, and longitudinal health history information on those same patients. Two-stage review process, refined from the Harvard Medical Practice Study's methodology. No attempt to evaluate the preventability or ameliorability (whether harm could have been reduced if a different approach had been taken).
Sharek 2011	Hospital-based (internal) and contract research organization-hired (external) reviewers Primary reviewers: nurses Confirmation: Physicians	For each record reviewed, the primary reviewer completed an electronic "chart review" form on a secure encrypted website. This form was used to record demographic information on the patient and initial information on any triggers or adverse events identified. For each "suspected" adverse event identified, the primary reviewer completed an "adverse event" electronic form, including study ID, patient demographic data, adverse event date, hospital location where the adverse event occurred (including before admission), primary hospital service when the adverse event occurred, free text description of the adverse event, determination whether the adverse event was identified by a GTT trigger, severity level of the adverse event, immediate response to the adverse event, category of the adverse event, and a specific adverse event code.

Measures of AEs

The findings of the studies regarding rates of adverse events detected are presented in a variety of ways, as demonstrated in Table 4. The measures proposed by the IHI for reporting results, i.e. the number of adverse events/100 patients (encounters) and the number of events/1000 patient days are still the most common reporting style. In five of the nine reviewed studies, the number of adverse events/100 patients and the number of AEs per 1000 patients are both used, and they range from 18,1 AEs/ 100 patients to 49/100 patients and 50,4/1000 patient-days to 91/1000 patient days, respectively.

Table 4. Outcomes and measures used in reviews with the IHI Global Trigger Tool in the included studies

STUDIES	total # of triggers	Adverse Events		events/100 patients	events/1000 patient-days
		n	%		
Classen 2008	NA	NA	NA	NA	NA
Naessens 2009*	-	65	27,7	-	-
Asavaroengchai 2009	776	236	24,0 (20,5 – 27,5)	41,0 (32,3 – 49,6)	50,4 (40,7 – 60,0)
Landrigan 2010**	-	588	18,1	25,1 (23,1 – 27,2)	56,5 (52,0 – 61,2)
Naessens 2010	3361	307	27,0	-	-
Carter 2010		-	-	-	- †
Good 2011***		736	-	31,1	41,6
Classen 2011	-	354	33,2	49 †	91 †
Sharek 2011	-	588	-	25,1 (23,1 – 27,2)	-

NA: Not applicable; - : not reported

Numbers in parentheses indicate 95% Confidence Interval range

* only 11 AEs (17%) were also detected as provider-reported events or PSIs, while only 3 cases detected by other methods (2 by provider reporting – 1 by PSI) were missed by the GTT review

** Numbers shown concern the findings of the internal reviewer teams

*** Numbers presented concern only those identified AEs which occurred during hospitalization

† Results of the hospital's AE rates/1000 patient days are presented as a graphic overview of a 4,5-year period

† as detected by all three methods combined, and not exclusively by GTT review

Assessment of severity

Assessment of adverse event severity has been undertaken in all studies, using the National Coordinating Council for Medication Error Reporting and Prevention Index (NCC MERP), categories E to I, which represent harm to the patient as follows:

- Category E: contributed to or resulted in temporary harm to the patient and required intervention
- Category F: contributed to or resulted in temporary harm to the patient and required initial or prolonged hospitalization
- Category G: contributed to or resulted in permanent patient harm
- Category H: required intervention to sustain life
- Category I: contributed to or resulted in the patient's death

Consistent with earlier studies on adverse event measurement, the majority of cases (over 50%) fall in the less grave E and F categories. Also, category E is mentioned as the most problematic to identify, particularly for less experienced reviewers.

Assessment of preventability

Preventability assessment has been undertaken in only three of the nine studies, and actually reported in two of them (the third study refers to a future publication regarding preventability and types of ADEs identified). Preventability has been assessed using subjective judgement of the physician reviewers and variations of a 4-level Likert scale, with the following categories:

- Level 1: unpreventable,
- Level 2: unlikely to prevent (less than 50% chance),
- Level 3: likely to prevent (more than 50% chance),
- Level 4: preventable.

Table 5. Assessment of severity and preventability in the reviewed GTT studies

Severity NCC MERP	Asavaroeng- chai 2009		Landrigan 2010		Naessens 2010		Good 2011		Classen 2011		Sharek 2011*	
	n	%	n	%	n	%	n	%	n	%	n	%
Category E	125	53,0	245	41.7	156	50.8	-	62,2	204	-	246	10,51
Category F	102	43,2	251	42.7	126	41.0	-	24,5	124	-	258	11,02
Category G	6	2,5	17	2.9	25	8.1	-	3,3	8	-	18	0,77
Category H	1	0,4	50	8.5			-	7,2	14	-	52	2,22
Category I	2	0,8	14	2.4			-	2,9	4	-	14	0,6
Preventabil- ity assess- ment	YES		YES		NO		NO		NO		YES**	
	n	%	n	%	n	%	n	%	n	%	n	%
Preventable AEs	122	51.7	364†	63.1	-	-	-	-	-	-	-	-

* The provided numbers concern the results of internal reviewers; the study reports also those of the external reviewing team and the combined result

** Details to appear in forthcoming publication

† The large majority of identified preventable harms were classified as category E (144) or category F (163); 13 caused permanent harm (category G), 35 were life-threatening (category H), and 9 caused or contributed to a patient's death (category I).

Methodological questions: Reviewers' agreement

As mentioned earlier, concerns have been raised regarding the reliability of the tool because of poor levels of agreement among reviewers which had been observed in several studies. Degree of agreement between reviewers using the GTT has been assessed using the kappa or weighted kappa co-efficient (14).

Table 6. The kappa co-efficient and its categories

Kappa Statistic	Strength of Agreement
<0.00	Poor
0.00 - 0.20	Slight
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Substantial
0.81 – 1.00	Almost perfect

Consistency in reviewing practices and inter-rater reliability issues can be at least in part (and usually have been) addressed through training of the reviewers as a team. An important aim of the training is to reduce variation by providing a commonly shared understanding of the definition of an AE, and corresponding ability to identify it (15), as well as a shared view of AE's severity and preventability. The advice and practice of using consistently the same review team (at least for one year period at a time) is also a common one (16).

In the reviewed studies, inter-rater reliability was variable, depending on the object of review (presence of an AE, severity, preventability) and the type of reviewers compared (nurses vs. physicians, internal vs. external reviewer teams etc.).

Generally, the levels of inter-rater agreement were in the moderate to substantial area of agreement, but there were also cases with stronger agreement, as for example in the case of the internal reviewers in the study of Sharek et al (3). The results of this study are presented separately in Table 8, due to the detailed comparisons undertaken in this work.

Table 7. GTT reliability: Reviewer agreement results by Cohen's kappa co-efficient

Classen 2008	Naessens 2009	Asavaroengchai 2009	Landrigan 2010	Naessens 2010	Good 2010	Classen 2011
0.007 -0,512 (before training) 0,164 – 0.703 (after training)	NA	0,86 (95% CI: 0.81-0.90)	0.64 – 0.93 (internal reviewers) 0.40 – 0.72 (external reviewers) k for preventability 0.83 (internal) 0.54 (external)	0.53-0.73 (nurses on triggers) 0.40-0.60 (nurses on adverse events) 0.65-0.77 (nurses and physicians on adverse events)	NA	NA

Table 8. Reviewer comparisons and agreement levels in the study of Sharek et al. (3).

AGREEMENT COMPARISONS – weighted kappas	INTERNAL TEAM	EXTERNAL TEAM
Inter-rater reliability within each team		
physician reviewers consensus vs. primary reviewers for severity	SUBSTANTIAL mean about 0,65	MODERATE mean about 0,5
physician reviewers between them for:		
presence of AE	ALMOST PERFECT mean about 0,85	MODERATE mean about 0,55
total number of AEs	ALMOST PERFECT mean about 0,9	SUBSTANTIAL mean about 0,7
Severity	ALMOST PERFECT mean about 0,85	MODERATE mean almost 0,6
Preventability	0,83	0,54
Inter-rater reliability between same type teams	Substantial mean almost 0,70	Moderate mean almost 0,4
Agreement between different team types (internal vs. external)	MODERATE mean 0,40	
number of AEs suspected number of AEs confirmed Severity	0,40 0,43 0,43	
Internal and external team vs. experienced team presence of AEs	FAIR to MODERATE (overall)	FAIR TO MODERATE (overall)
presence of AE	0,45	0,30
total number of AEs	0,35	about 0,30
Severity	about 0,57	0,25
Intra-rater reliability (primary reviewers) number of AEs	Almost perfect mean 0,90	Moderate mean 0,50

Resources: training and implementation phase

The resources required for utilization of the GTT were not reported in a clear or consistent manner across the studies. Most often, it is reported that the review time for each record was kept to a maximum of 20 minutes, as instructed by IHI. On the basis of that information, the actual amount of total reviewer time used can be calculated based on the number of reviewers and the size of the reviewed sample. There is no concrete information on the amount of time required to resolve and reconcile differences in assessment, neither is the time of the second stage review explicitly stated.

Another element in resource use is that of training, which was reported in five of the eight studies.

In the study of Classen (2), the training phase included reading the GTT White Paper, and then undertaking a review process (of 20 min each) for 15 records. Next, physician reviewers conducted a 2h formal training session with the primary reviewers

on the GTT. Total participation time for primary reviewers in this study is estimated to be approximately 20 hours, including record reviews, training session, and planning sessions; physician reviewers spent additional hours selecting training records and reviewing primary reviewer data submission.

Landrigan et al (8) report training organised for both internal and external reviewing teams through standardized series of Web-based seminars and provided by patient-safety experts and experienced reviewers. These seminars included didactic sessions, practical review exercises, and debriefing sessions.

In the Naessens study (4) the first 5 months of reviews were treated as training cases to establish a working process for on-going reviews.

Good et al (2011) report that the initial training lasted approximately 12 hours, composed of face-to-face didactic sessions, data entry and review of standardized training cases, and discussions of case evaluations. This training programme was developed by a team consisting of a physician, a nurse and the GTT project manager, and was delivered by the project manager. Additional training, in the form of face-to-face sessions and conference calls, has been provided as needed after revisions to the data collection tool and audit or data management processes

Sharek et al (2011) describe several phases of training. First an orientation phase consisting of reading the GTT White Paper combined with 1 hour of Web-based review. Subsequently, a phase of reviewing a standardized set of training medical records (sample of 10 HIPAA-compliant records for primary reviewers, and sample of 5 for physician reviewers) again combined with 1 hour Web training. Finally, review of 10 site-specific hospital medical records, plus 1 hour Web training. Training on the web application and data collection instruments was conducted by a Contract Research Organisation via a webinar.

Finally, yet another category of resource use is that of development and testing of data collection tools. In one of the studies (8) a Microsoft Access tool was developed for review data collection and entry, and pilot tested by the lead nurse reviewer for the project and the GTT initiative project manager by reviewing the same 10 charts from a single facility. After testing, the tool was used by the external nurse auditors who performed the GTT-based review of the sampled records.

Implementation questions: the organisational viewpoint

On the basis of the reviewed studies, as well as the evidence produced and reported by the implementation projects in other Nordic countries, we summarize here some practical matters for the consideration of organisations planning to embark on the use of the GTT as a patient safety monitoring instrument.

In spite the fact that the GTT methodology is standardized, in practice there are a number of issues that can be (and often have been) implemented in ways deviating

from the original guidance of the IHI. As a result, there are a number of choices to be made prior to actual use of the tool. These include:

- Number of records to be sampled, frequency of sampling
- Population of patients included
- Triggers to be used – possible modifications
- Duration and content of training the reviewing team
- Size of the reviewing team
- Professional background of the reviewers
- Choice of internal or external primary reviewers
- Frequency of meetings between primary and secondary reviewers
- Method of resolving differences of opinion on each review stage
- Decision on inclusion or exclusion from review for adverse events ‘present on admission’
- Decision on inclusion or exclusion of errors of omission and commission (in the original methodology only the latter are included)
- Decision on whether preventability of confirmed AEs will be assessed and if yes, method of assessment
- Documentation practices of the review team (Will the nurses produce narrative descriptions of adverse events? Will the findings of the review, plus additional details be documented for future use?)
- Means for analysing and reporting the results and selection of targeted groups for dissemination

Resources needed, Added value of the method

The resources an organisation needs to invest for using the GTT should be distinguished to the following categories:

1. Resources required for modification or fitting of the method to local context (selection or modification of triggers, translation to local language and adaptation to local practices);
2. Resources required for training the original reviewing team, and for regular updates either because of changes in methodology and/or turn- over of personnel;
3. Resources required for regular implementation and follow up with the GTT. According to IHI’s assessment, review of 10 patient records every two weeks (which equals one data point) takes approximately three to four hours of mid-level staff time for each reviewer and about 30 minutes of physician time to accomplish;
4. Yet a separate category in terms of amount of resources required is that of development, testing and validation of a tool for application in a previously unexplored clinical domain. As an example, the most widely tested Ad-

verse Event Trigger Tool, required collection of data from 86 hospitals where more than 2800 charts were reviewed, covering more than 268000 medication doses (17).

In the Finnish experiences, the corresponding resource demands were documented as follows:

In VSSHP training took approximately 1,5 days, and the first samples of records from a two weeks' randomized sample covering ten patient journals were analysed without any time limit. Conducting the first phase of the analysis took approximately 72 hours/year for one person (9 working days) – which is to be multiplied by two as there are two reviewers. The second phase took from the whole team of four persons 8 hours/year. The team meeting once every three months was easy to integrate with chief medical officers' programme.

It should be noted that the hospital district already had personnel working specifically with patient safety and quality improvement (who were then also the ones to perform the first stage analysis).

In the Vaasa Central Hospital (22), four nurses and one doctor were first trained to use the GTT method. In addition, two more physicians were named as members of the team, but did not participate in the training. All team members received in advance the Finnish translation of the IHI GTT manual. In addition, all nurses on the team acted as patient safety co-ordinators, so they had received patient safety education earlier, and their understanding of system approach to analysis was already good. After training, the nurses practiced the review process in pairs – first jointed review by a pair, and then consensus with another pair. In the phase of regular use, the GTT method was used as indicated in the manual, i.e. first each nurse reviewed the sampled records separately, and then they jointly wrote down their observations. Subsequently, review of findings was undertaken by the physicians of the team. The GTT has been in use at the hospital for over two years. The participating nurses, in the context of their role as patient safety coordinators, have one day per week allocated to patient safety work. From that time, about half (4 hours) goes to GTT review of records. The remaining half is then utilized to benefit from the knowledge received, as well as from other patient safety evidence. The physicians have met at a maximum of three month intervals, and those meetings have lasted about 2 hours, which means that on an annual basis a maximum of 8 hours of physician work is needed for the task.

The resource demands reported in the neurosurgery pilot in TAYS (reported in Chapter 4) cannot serve directly as an example of regular use of the GTT, but rather they are indicative of the resources and effort required for the adaptation and application of the tool to a new clinical field.

In return, the added value of the method is seen in using patterns of harm emerging over time, as a means for the organisation to identify areas where resources and patient safety improvement efforts should be focused. The true added value is strongly

connected with the next theme discussed, i.e. the utilization of the review process and its results as a learning opportunity.

Documentation of identified cases – Learning through using the GTT

Using the GTT within an organisation's patient safety programme is also an excellent learning opportunity that appears in several distinct occasions and various forms. Through the literature, and through the practical examples in the Finnish hospitals, we have identified that the following learning opportunities arise:

- Learning from the clarification/final assessment of reviewed cases. Reviewers state and exchange their views on presence and type of triggers, severity of case, preventability or not of an adverse event. There is a clear need to utilize and capitalize better on the rich material generated through assessment sessions, by documenting the reasoning supporting the final decisions made, so that it is available as future reference.
- Learning in terms of identifying target areas for development
- Dissemination of findings to the whole organisation
- Dissemination of findings to collaborating partner organisations – connected to identification of cases where the adverse event happened before admission. The means and the most appropriate channel for sharing this knowledge are a subject of further innovation.

In conclusion

When compared to the traditional method of full patient record review for the identification of adverse events, the Global Trigger Tool is a worthwhile alternative, particularly when the context of implementation is regular follow up within an organisation, and not research work. Use of the GTT requires less resources, due to the fact of reviewing a smaller, but condensed (in the sense of having higher probability of containing adverse events) sample of records. Nevertheless, the need to look into the record in more detail once a trigger has been located still remains.

The GTT tends to identify a larger number of adverse events when compared to other detection methods used in the US, such as voluntary reporting systems of sentinel events and the AHRQ Patient Safety Indicators. This difference has been attributed to the broad definition of adverse events used by the GTT, which includes also events present on admission, as well as less serious than sentinel events.

An important finding that emerged from the reviewed studies is that the Global Trigger Tool detects events, which would have gone unnoticed by other standard methods (e.g. incident reports, pharmacy interventions etc.). In addition, there is little overlap in the types of events detected through GTT vs. other methodologies used during the same period of follow up. In the light of these observations, we can con-

clude that use of the trigger tool approach can supplement incident reporting and other interventions when the aim is a comprehensive picture of the level of adverse events in an organisation.

This echoes the position of the IHI (White Paper) that “... hospitals should use the IHI Global Trigger Tool as one part of a learning system that includes other component measures, such as voluntarily reported errors, surgical site infections, and other outcome measures” (1). The necessity for utilizing a palette of methods to monitor and improve patient safety has been echoed in the publication of both scientists and organisations in the field (18 – 20).

There are two main issues of controversy around the GTT: its reliability as a method due to its limited validation, and the suitability of the tool for use as a benchmarking instrument, to assess performance of different hospitals. Consistency in reviewing practices and inter-rater reliability issues can be at least in part (and usually have been) addressed through using consistently the same review team (at least for one year period), already at the training phase. Classen et al have demonstrated that training improves inter-rater level of agreement, and generally in most studies there seems to be at least a moderate level of agreement achieved (higher when internal reviewers are used).

On the other hand, every implementation of the GTT among the ones reviewed seems to be an own, local variant, with the only elements that are truly stable across studies being the two staged review approach and the method of severity assessment (although there is not necessarily agreement on its implementation). This in turn raises the question as to whether the goal of reducing variation in inter-rater reliability and achieving generalizability has been attained – particularly if the target would be to allow cross-organisation comparison as in benchmarking.

In the early days of trigger tools development, the IHI team had stated clearly that the tools should not be used as a benchmarking instrument across institutions, since they had not been validated. In addition, they felt that comparison of ADE rates across organisations would be counterproductive and instead would cause either unnecessary anxiety or, conversely, a false sense of security (16). Later on, more emphasis was placed on the use of the tool for large scale assessments, but still not in the context of benchmarking.

The study published in April 2011 in *Health Affairs* (9), took the first big step towards comparative use of the tool, by applying it to comparison of specific adverse event rates of different hospitals. The publication has drawn a lot of publicity, but it has also received a lot of criticism, including the observation that the definition of adverse events used by different methods can be a significant part of the explanation of the results (21).

Further yet, by focusing on patient harm, IHI methodology approaches the subject of patient safety from a viewpoint closer to the patient/subject of care. However, this happens on the expense of preventability – the method does not in itself differentiate between injuries caused by error or substandard care and those that were una-

voidable. IHI's view on preventability with regard to the GTT is clear (1): preventability should not be an inclusion/exclusion criterion for a patient record, precisely because of preventability's constant change over time (23). That view, however, does not mean that assessment of preventability of confirmed events should not be undertaken, in order to identify areas (and perhaps also means) of improvement. Without it, the benefits of measuring adverse events remain limited, and the opportunity to learn and improve is lost.⁴

With regard to learning, attention should also be paid to the inherent limitation of the GTT: namely that it explicitly excludes near-misses, as well as errors of omission – both of which are very important sources of learning and advancing towards prevention of adverse events.

In spite its limitations, the GTT paper method can be seen as a valid alternative for regular use in the place of the much more resource intensive and generally research oriented methodology of full patient record review. In addition to the tool's ability to detect larger number of events than other assessment methods, comparative studies also indicate that the GTT may identify different types of adverse events. In the light of these observations, we can conclude that use of the trigger tool approach can supplement incident reporting and other interventions when the aim is a comprehensive picture of the level of adverse events in an organisation. Indeed, use of the GTT seems to be on the rise - at least in the US, as well as in other Nordic countries, as part of national patient safety initiatives/programmes.

The original inspiration for the current generation of trigger tools was work on automated triggers in the early '90s. Now, after almost a decade of development, IHI and the developers of GTT are placing again their hopes for future success and more widespread adoption of the tool on the automatization of medical records (9). Many of the groups reporting their experience with the paper-based GTT also refer to the need of a tool integrated with electronic patient record systems (16). Since the topic of computerized or automated trigger tools is most relevant also for the Finnish context, it is addressed in depth in a separate chapter of this report (Chapter 5).

References

1. Griffin FA, Resar RK. IHI Global Trigger Tool for Measuring Adverse Events (Second Edition). IHI Innovation Series white paper. Cambridge, Massachusetts: Institute for Healthcare Improvement; 2009
2. Classen DC, Lloyd RC, Provost L, Griffin FA, Resar R. Development and Evaluation of the Institute for Healthcare Improvement Global Trigger Tool. *J Pat Safety* 2008. 4(3):169-177

⁴ The matter of preventability is receiving now more attention by the IHI, as demonstrated in the interview of David Classen (24), where he is also proposing to enlarge the concept by including mitigability and ameliorability – aspects which become relevant as automation of the GTT progresses.

3. Sharek PJ, Parry G, Goldmann D, Bones K et al. Performance Characteristics of a Methodology to Quantify Adverse Events over Time in Hospitalized Patients. *Health Services Research* 2011; 46(2):654-678. DOI: 10.1111/j.1475-6773.2010.01156.x
4. Naessens JM, Campbell CR, Huddleston JM, Berg BP et al. A comparison of hospital adverse events identified by three widely used detection methods. *Int J Qual Health Care* 2009; 21 (4): 301–307
5. Asavaroengchaia S, Sriratanabana J, Hiran-suthikula N, Supachutikulb A. Identifying adverse events in hospitalized patients using global trigger tool in Thailand. *Asian Biomedicine* 2009; 3 (5): 545-550
6. Naessens JM, O’Byrne TJ, Johnson MG, Van Such MB et al. Measuring hospital adverse events: assessing inter-rater reliability and trigger performance of the Global Trigger Tool. *Int J Qual Health Care* 2010; 22 (4): 266–274
7. Landrigan CP, Parry GJ, Bones CB, Hackbarth AD, et al. Temporal Trends in Rates of Patient Harm Resulting from Medical Care. *N Engl J Med* 2010; 363:2124–34
8. Good VS, Saldaña M, Gilder R, et al. Large-scale deployment of the Global Trigger Tool across a large hospital system: refinements for the characterisation of adverse events to support patient safety learning opportunities. *BMJ Qual Saf* 2011 20: 25-30. doi: 10.1136/bmjqs.2008.029181
9. Classen DC, Resar R, Griffin F, Federico F et al. 'Global Trigger Tool' Shows That Adverse Events In Hospitals May Be Ten Times Greater Than Previously Measured. *Health Affairs* 2011; 30 (4):581-589. doi: 10.1377/hlthaff.2011.0190
10. Carter M. Measuring harm levels with the Global Trigger Tool. *Clin Risk*, 2010;16:122-126
11. Health Foundation Report. Trigger Tools. <http://www.health.org.uk/public/cms/75/76/313/2601/global%20trigger%20tools.pdf?realName=InzqMF.pdf>
12. Hutchinson A, Coster JE, Cooper KL, McIntosh A, Walters SJ, Bath PA, et al. Comparison of case note review methods for evaluating quality and safety in health care. *Health Technol Assess* 2010;14(10), DOI: 10.3310/hta1410
13. Brennan TA, Leape LL, Laird NM, Hebert L et al. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. *N Eng J Med*; 1991: (324), 370–6
14. Landis RJ, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. 1977;33:363Y374.
15. Resar RK, Rozich JD, Classen D. Methodology and rationale for the measurement of harm with trigger tools. *Qual Saf Health Care* 2003;12:39-45
16. Rozich JD, Haraden CR, Resar RK. Adverse drug event trigger tool: a practical methodology for measuring medication related harm. *Qual Saf Health Care* 2003;12:194-200.
17. Adverse Event Analysis - Chapter 6. In: Patient safety : achieving a new standard for care / Committee on Data Standards for Patient Safety, Board on Health Care Services ; Philip Aspden ... [et al.], editors. National Academy Press, 2004, ISBN 0-309-09077-6 (hardcover)
18. Rosen AM. Perspective – Measuring Patient Safety. Available at: <http://www.webmm.ahrq.gov/perspective.aspx?perspectiveID=94>
19. Battles JB. Commentary. *Qual. Saf. Health Care* 2005;14;225-226
20. Ferranti J, Horvath M, Cozart H, Whitehurst J, Eckstrand J, Pietrobon R, Rajgor D, Ahmad A. A Multifaceted Approach to Safety: The Synergistic Detection of Adverse Drug Events in Adult Inpatients. *Journal Pat Safety*. 2008; 4(3):184-190.
21. Campione JR. Comparing Global Adverse Events in Hospitals. North Carolina Center for Hospital Quality and Patient Safety.
22. Use of the GTT in Vaasa Central Hospital – Personal communication by M. Kinnunen, Patient Safety & Quality Coordinator for the hospital
23. Scanlon MC, Mistry KP, Jeffries HE. Determining pediatric intensive care unit quality indicators for measuring pediatric intensive care unit safety. *Pediatr Crit Care Med* 2007; 8 (2)(Suppl.). S3-S10
24. AHRQ Morbidity and Mortality Rounds on the Web. In Conversation with David C. Classen. April 2012. Available online at: <http://webmm.ahrq.gov/printviewperspective.aspx?perspectiveID=119>

3 Assessment of Adverse Events in the Southwest Hospital District of Finland

Karolina Peltomaa, RN, MNSc, Quality Manager, Vaasa
Central Hospital

Weronica Gröndahl, RN, MNSc, Clinical nurse specialist,
Department of Surgery

Ilkka Kantola, MD, Chief Medical Officer,
Department of Internal Medicine

Timo Savunen, MD, Chief Medical Officer (retired),
Department of Surgery

A Patient Safety Project was implemented in the Southwest Hospital District of Finland as an internal development project for years 2009 - 2011. One of the aims was to decrease the number of adverse events in the whole hospital district with 30 % by the end of the year 2011. The IHI Global Trigger Tool (1) was chosen as an instrument for assessment of the number of adverse events, first by creating a baseline measurement and eventually for assessing the outcomes of the project. The tool was translated into Finnish with the permission and help from the authors of the IHI White Paper (Frances A. Griffin, Roger K. Resar). Descriptions of relevant patient safety concepts were included, in addition to a tool evaluating the preventability of the adverse events, which is not included in the original White Paper. The unofficial translation was conducted by the patient safety manager. Three senior medical officers and a clinical nurse specialist checked both the translation and the triggers. The first edition of the tool was introduced in 2009, and the second edition with updated triggers was implemented in 2010. GTT has been conducted from 2009 until today – nearly four years.

Preparations before implementation

A research plan was prepared before the implementation, and the plan went also through evaluation by an ethical board. Before the implementation, the project manager participated into a course regarding retrospective patient record analysis methodology in Jönköping, Sweden. Subsequently, the project manager educated the rest of the team. In addition to the project manager, the team consisted of a clinical nurse specialist, both MNSc with broad clinical experience, and two chief medical officers (representing surgery and internal medicine). When choosing the team members close attention was paid to their clinical experience and to the possibility for commitment for several years, or as long as possible.

Conducting the analysis

Retrospective patient record analysis has been conducted in Southwest Hospital District of Finland according to the White Paper published by IHI. Randomisation of the patient records was performed by the unit responsible for electronic patient records. Randomisation was performed using the following inclusion criteria:

- Time from discharge date at least two months
- Length of stay at least 24 hours and formally admitted to the hospital
- Patient age 18 years or older
- Excluding inpatient psychiatric and rehabilitation patients
- Including all the deceased.

The population from which the randomisation was performed consisted of all the adult, somatic patients from the whole hospital district, as presented in the original GTT white paper. To form one sample, 15 patient records from the entire population of discharged adult patients were randomised every two weeks (for example, patients discharged between the 1st and 15th of the month for the first two-week sample and between the 16th and end of the month for the second two-week sample), for a total of 30 records per month. 10 patient records were analysed twice a month. The sample size proved to be practical, because despite the clear intake criteria presented to the responsible unit, in several samples one or two patient records did not meet the inclusion criteria (ex. day-surgery patients). In such situations the next patient record was chosen from the list of randomised cases. All the data used in the analysis (Excel sheets, results from the randomisation and other material) was saved into a internal web disc / file. Only the members of the analysis team had the access codes to the files.

During the training phase nurses went through several samples together in order to validate the methodology and concepts used. The first sample was analysed without the time limit. This proved to be a clear asset - it was useful to go through all the triggers and discuss the different connotations related to each trigger and how they were presented in the patient records. In addition, a lot of productive discussion developed in connection with the evaluation of severity and preventability. After the training sessions the nurses started to perform the analysis separately, using the method's 20 minutes timeframe per patient.

Analysis was conducted by looking at one admission. From the patient records nurses went through the discharge codes, particularly infections, complications, or certain diagnoses (ICD 10), the discharge summary, physician progress notes, nursing notes, medications and laboratory results. Also the previous and upcoming admissions were checked, so the possible re-admissions would be identified as well. From the admission in question nurses were looking for triggers (Annex 4) which are 51 in total in the original White Paper. After identifying a trigger a conclusion was made whether the trigger pointed to an adverse event. If an adverse event was identified, also the severity was assessed. All the findings and the progress of the analysis were documented into an electronic form. Very often there was no adverse event, even if a positive trig-

ger was found, ex. when a patient returned for an unplanned check up in 30 days after discharge, but there was no evidence of a complication or an adverse event.

The first phase of the analysis was performed separately by the two nurses. The electronic patient record system allowed the analysis to take place where ever there was a computer available, so there was no need to organise a separate office space. After analysis of a sample the nurses met and discussed their findings, either face to face or by phone. If there were discrepancies between the two analyses – in the number of adverse events, their severity or primary assessments of preventability, the nurses went back to the records to reach consensus over the matter. There were rarely any differences in the adverse events found by the two nurses, and those were talked over easily. The broad clinical experience was crucial. Those performing the analysis must know their organisation, the documentation procedures and how clinical care is organised. Being familiar to the documentation styles is especially relevant in a situation where documentation seems not to be systematic and the quality varies.

During the second phase the whole team, including the chief medical officers, gathered to go through the primary analysis. The aim of these meetings was to confirm the results from the first phase and make the final assessment regarding severity and possible preventability of the adverse events, and identify areas needing development and further analysis. During those meetings the team looked at the statistics and in some cases the actual patient records to make the right judgement. In some more complex cases nurses wanted to consult doctors before making any final decisions on the presence of an adverse event, and these cases were also discussed in more detail during the second phase. In a deviation to the instructions of the original manual, the whole analysis team met only every three months instead of twice a month for the second phase. These meetings made it possible to go through multiple adverse events simultaneously with more analytic discussions. On the other hand, in some cases it was difficult to remember all the details of some patient cases, because the primary analysis was conducted several months before, which then forced the team to go back to the patient records. A second, but significant difference to the original methodology was that the analysis team made a decision to include both omissions and commissions as adverse events. In the original White Paper adverse events as a result of omissions were not included, but the team was unanimous in including omissions as well. This will affect the final statistical variables in a profound way, which is to be recognised when evaluating the results of the hospital district.

Preventability of the adverse events

As a third difference to the original method, the preventability was also assessed when an adverse event was found. In Sweden there is a rough guideline, based on subjective assessment in order to help the evaluation of preventability:

1. No evidence of preventability
 2. Small or moderate evidence of preventability
 3. Preventable with likelihood of almost 50 % or less
 4. Preventable with likelihood of almost 50 % or more
 5. Strong evidence of preventability
 6. Clear evidence of preventability
1. – 3. Adverse event was not preventable
 4. – 6. Adverse event was preventable

According to this definition preventable adverse events are “vårdskador” (adverse events as a result of medical treatment). Assessment of the preventability of an adverse event is based on clinical experience, knowledge and subjective reflections on the reasons of the adverse event. Thus also the broad and extensive experience of the chief medical officers was extremely significant. In spite of the subjective nature of the assessment of preventability, the decision was almost without exceptions unanimous.

Severity of the adverse events

In the White Paper the National Council for Medication Error Reporting and Prevention (NCC MERP) Index for Categorization Errors is used for assessment of the severity of the adverse events:

Category E:	Temporary harm to the patient and required intervention
Category F:	Temporary harm to the patient and required initial or prolonged hospitalization
Category G:	Permanent patient harm
Category H:	Intervention required to sustain life
Category I:	Patient death

Assessing the severity of the adverse events using the categorisation was clear. However, adverse events in category E proved to be difficult to find because of the heterogeneity of the documentation, and in the early stages of the analysis this was the category where some discrepancies in the analysis were also found.

Results and development from the basis of the analysis

Assessment of adverse events using IHI Global Trigger Tool has continued in the Southwest Hospital District of Finland for nearly four years. Because analysis is performed retrospectively, there are 89 samples ready at the time of the publication of this report. The results include 1335 randomised patient admissions, out of which 890 were included into the analysis. Based on the analysis conducted during year 2012, from patient journals were identifiable 33 AE's / 1000 patient days, 12,9 AE's /100 patient admissions and 11,2 % of all admissions included one or more AE's. With this method it has been possible to identify several issues in clinical care and organisational processes that required development and further evaluation. In addition to the development areas it has been possible to acquire knowledge of the general level of patient safety in adult, somatic patients and identify the general progress over time. As a result of the analysis it was possible to confirm that the 30 % reduction in adverse events, the goal set in the beginning of the patient safety project, was reached within 2 years.

Conclusions

After using the IHI Global Trigger Tool for nearly four years it is possible to recommend the methodology. It produces knowledge and information which is usable for development and evaluation of patient safety in an organisation. It is clear that on organisational level the information produced is more beneficial in terms of statistical variables in relation to time, as finding the development areas from such a broad and heterogenic population is demanding. When transferred to department level the benefits will increase, because the results will indicate more in detail the problems in a limited speciality.

Finding the resources for the analysis was not demanding in the hospital district, as the persons conducting the first phase were already working with patient safety and quality improvement. In addition the meeting every three months could be relatively easily fitted into the schedule of the chief medical officers. Training took approximately 1,5 days, and the first samples were analysed without any time limit. Conducting the first phase of the analysis took approximately 72 hours / year for one person (9 working days). The second phase took for the whole team 8 hours / year. When considering the information IHI Global Trigger Tool is providing, the method did not appear as very resource intensive on the organisational level. However, if it is to be conducted in smaller departments, the resourcing is likely to become an issue, especially if there are no personnel responsible for quality and patient safety in place.

There were some challenges in the search for a team with the possibility to make the commitment for several years, in this case for three years. In the beginning it was important to discuss about the different meanings and connotations of the triggers, as well as severity and preventability of the adverse events in order to find a shared un-

derstanding and thus raise the level of internal validity. This covers also the discussions on the adverse events during team meetings.

Conducting the analysis “manually” from electronic patient records proved to be the only effective way at the moment as a result of the variety and heterogeneity of the quality of documentation. Especially the documentation of medication records was problematic in the hospital district. Most of the adverse events in categories E and F were found from nursing notes, but categories G, H and I were seen in doctor’s notes also. In ICD coding the categories E, F and G were also present rarely.

The progress of using IHI GTT in Southwest Hospital District will continue with implementation to different departments, Surgical Departments as one of the firsts. Using the methodology on department level will produce more accurate and detailed information. However, this requires the translation and validation of additional triggers related to day-surgery, pediatrics and psychiatrics. In addition, in the beginning the analysis was performed with Excel sheets, but the accumulating number of patient records required the development of an electronic, web-based solution, which is currently in use.

References

1. Griffin FA, Resar RK. IHI Global Trigger Tool for Measuring Adverse Events (Second Edition). IHI Innovation Series white paper. Cambridge, Massachusetts: Institute for Healthcare Improvement; 2009

4 Use of the Global Trigger Tool in neurosurgery: A comparison of manual and text mining detection of triggers

Juha Öhman, MD, PhD
Teemu Keski-Kuha, M.Sc
Henna Henttonen, MD
Mika Kaartinen, M.Sc
Erkki Kujansuu, MD, PhD

Introduction

Adverse events during hospitalization affect nearly one out of 10 patients (1-4). These constitute a major problem with serious consequences, time and monetary loss (1, 5, 6). The Global Trigger Tool is a method for monitoring patient safety within a health-care provider organisation (7-9) It enables longitudinal comparisons and assessment of patient safety measures implemented (8). It is performed in retrospect from randomly selected patient records by an independent reviewer (7, 10). Since nearly all patient records in Finnish healthcare settings are in electrical form, a computer-based trigger detection approach is eventually of more interest than the paper-based method (8). In addition, only by the use of computer based approaches is it possible to achieve real-time monitoring and hence preventive use of patient safety triggers, by incorporating them into everyday use. Two complementary methods were studied in parallel in this study: testing the suitability and applicability of an adapted trigger tool for detecting triggers in neurosurgery patients through the use of electronic patient records with manual review and with text mining.

Table 1. Final list of triggers modified to neurosurgery.

Trigger #	Definition
T1	Deterioration of neurological condition (e.g. decrease of 3 or more points on GCS)
T2	X-ray imaging of any kind and not routine for the procedure postoperatively in ICU or normal ward
T3	Unplanned return to surgery or endovascular procedure
T4	Complication caused by any treatment procedure during ICU or normal ward stay
T5	Wound problem (e.g. CSF leak, poor wound edge adhesion)
T6	CSF circulation disturbances (e.g. hydrocephalus)
T7	Infection (e.g. any nosocomial infections: central line infection, surgical site infection, or urinary tract infection either serious or mild)
T8	Readmission to ICU or readmission to treating hospital within 30 days
T9	Prolonged ICU treatment due to other than primary cause (e.g. infection, ventilator treatment)
T10	Intubation or reintubation
T11	Fluid balance disturbances needing treatment
T12	Electrolyte balance disturbances needing treatment
T13	Blood pressure disturbances needing treatment

GCS, Glasgow Coma Scale; ICU, intensive care unit; CSF, cerebrospinal fluid

Materials and Methods

This is a retrospective study in single hospital with three different ICD-10 based diagnosis groups from the Department of Neurosurgery, Tampere University Hospital, Finland. Three parallel time representative samples ($n = 556$) were formed from the total neurosurgery admissions between April 2007 and May 2008 ($n = 1969$) to assess adverse events. Triggers were identified using two different methods: a conventional manual review of electronic patient records and a text mining procedure using the same electronic patient text data.

Thirteen triggers relevant to neurosurgery were developed (Table 1). The final list of triggers was based on the original IHI Global Trigger Tool, and the subsections of the Surgery and ICU Trigger Tools.⁷ Triggers were modified to be suitable for neurosurgery according to adverse events identified through prior clinical studies in areas of intracranial haemorrhage and traumatic head injury.¹¹⁻¹²

A sample of 506 records of patients with head injuries (ICD-10 diagnosis codes S06, $n = 195$) and intracranial haemorrhage (ICD-10 diagnosis codes I60 - I62, $n = 211$) who had been hospitalized in the period of between April 2007 and May 2008 were retrieved. In addition, the records of 150 patients hospitalized with cervical spine pathology (e.g. spondylarthrosis, cervical disc disease) (ICD-10 diagnosis codes M50.0 - M51.1 and M47.1 - M47.2) from the same time period were also retrieved in order to serve as the control group.

Review Process

A trained study nurse reviewed the records for triggers. A review of each record was performed with the use of the trigger tool in a standardized fashion from the screen in about 20 minutes. The order of record review was randomized (i.e. reviews were not conducted in order of admission date or diagnosis) to prevent any distortion in the results over time by the reviewer's gradual accumulation of experience with the trigger tool. All data was collected into an Excel program (version 11, 2003, Microsoft) and after cross-checking the data for typographic errors the triggers were identified. The final decision on whether the cases identified by the study nurse were indeed patient safety incidents was made by an experienced neurosurgeon who reviewed the cases blind to the nature of the triggers. Moreover, the same neurosurgeon additionally reviewed 20 % of the files randomly selected from each group. The results were compared with the study nurse's review results and the data, after agreement, was entered into the file.

The study was designed only to detect triggers from the electronic patient records. No demographic data or clinical assessment of the severity of the underlying disease was taken into account.

Table 2. Total number of triggers found in manual review of different patients groups.

Patient group	No. of patients	No. of triggers found	Patients with trigger(s)	Percent of patients with trigger(s) within the group
Spinal	150	14	14	9.3 %
HI	195	137	44	22.6 %
ICH	211	315	74	35.1 %

Spinal, ICD-10 diagnosis codes M50.0 - M51.1 and M47.1 - M47.2

HI, ICD-10 diagnosis codes S06

ICH, ICD-10 diagnosis codes I60 - I62

Automatic review and statistical analysis

For the purposes of the text mining review the SAS® Text Miner (SAS Institute Inc.) program was used. The patient text data was extracted by ICD-10 diagnosis number from the electronic patient record system (Uranus, Logica, Finland) and transferred to the SAS system for processing. Patient data was cleansed with SAS Enterprise Guide™ program. Two alternative versions of patient record source data were produced, one containing the full patient text and another with only summaries. To prevent over fitting of the model and to obtain an unbiased assessment of text mining effectiveness, text mining began by splitting of the data randomly into 3 data mining sections, training (for model fitting, 40 %), validation (for model selection, 30 %) and test sets (for unbiased effectiveness estimate, 30 %). Selection of training validation and test set was random and independent for each trigger. Words in the patient records were stemmed

with synonym and acronym detection. Words were then transformed into a smaller set of linear combinations of words using Singular Vector Decomposition (SVD) of the document-term matrix of patient records. Alternatively, some key words associated with triggers were used and SVD generation bypassed. Finally the logistic regression model was fitted using the resulting words or linear combinations of words as independent variables and the existence of trigger (as detected in the manual review) as a target variable. For each trigger, a combination of best possible source data (summaries vs. whole record) and data reduction method (SVD or keywords) was selected as the final model.

The nurse entered the data of each patient record review in an Excel sheet for further analysis. The final decision on whether the cases identified by the nurse reviewer were indeed patient safety incidents was made by an experienced physician.

Results

Manual review revealed a total number of 446 triggers as follows: in the head injury group (HI), 137 triggers were identified in 44 patients, in the intracerebral haemorrhage group (ICH) 315 triggers were identified in 74 patients, while in the cervical spine pathology group (Spinal) only 14 triggers were identified in 14 patients (Table 2). These results were in accordance with the original hypothesis that there should be only few triggers in the spinal group. The ICH group had almost twice the number of triggers as the HI group. The incidence of individual triggers found in the manual review is shown in Table 3.

Although only a test set of patient records (i.e. 30 % of the total number of patient records) was used in the text mining review, the results were in good accordance with the manual review.

Because incidences consistent with the triggers described in the patient records varied from single words or numbers to the entire content of the text there were notable differences between various trigger detections. The incidence of detection of triggers yielded a sensitivity from 60 to 100 % and specificity from 80 to 98 % (Table 4). Triggers described with exact wording in the text were found more often than broadly described triggers e.g. blood pressure disturbances with a sensitivity of 87.5 % vs. deterioration of the neurological condition with a sensitivity of 69.2 %. The manual review of the electronic patient records took a total of four months' worth of working days for a trained study nurse and data checking five full working days for a neurosurgeon. Testing the text mining for detecting the triggers required about 20 working days by an experienced statistician. Once the model was ready for analysis, the processing of the electronic patient record forms took only a few minutes.

Table 3. Incidence of individual triggers found in manual review (percentage of triggers per total number of patients, n = 556).

Trigger	Incidence in group Spinal n = 150 (% total)	Incidence in group HI n = 195 (% total)	Incidence in group ICH n = 211 (% total)
Deterioration of neurological condition	0	22 (14.7)	35 (16.6)
X-ray imaging	7 (3.6)	30 (20.0)	61 (28.9)
Unplanned return to surgery or endovascular procedure	5 (2.6)	24 (16.0)	40 (19.0)
Treatment complication	1 (0.5)	1 (6.7)	9 (4.3)
Wound problem	0	3 (2.0)	6 (2.8)
Cerebrospinal fluid circulation disturbances	0	0	23 (10.9)
Infection	1 (0.5)	28 (18.7)	57 (27.0)
Readmission	0	15 (10.0)	15 (7.1)
Prolonged intensive care unit treatment	0	0	8 (3.8)
Intubation or reintubation	0	0	9 (4.3)
Fluid balance disturbances	0	11 (7.3)	3 (1.4)
Electrolyte balance disturbances	0	1 (6.7)	27 (12.8)
Blood pressure disturbances	0	2 (13.3)	22 (10.4)
Total no. of triggers found	14	137	315

Discussion

This was a single-centre study to assess the suitability and applicability of an adapted trigger tool in neurosurgical patients through the use of electronic patient records with two different approaches.

Studies based on retrospective assessment of information in medical records may underestimate the actual rate of adverse events (13). Moreover, some adverse events, such as those related to pharmaceutical therapy, may be difficult to detect, as the side effects of drugs may be very similar to the symptoms of diseases. Neurosurgery was selected for this study because adverse events related to invasive procedures and op-

erations are easier to identify and, conceivably, more consistently documented in the medical records than other types of adverse events (13,14).

Structured case note review, when carried out by trained professionals, has been shown to reliably detect adverse events (13, 15). Recent studies have shown that the Global Trigger Tool has very high specificity, high reliability, and high sensitivity, especially when used by trained reviewers, to achieve high levels of agreement on the presence and severity of adverse events (AEs) (16,17). While manual chart review is effective, it is too labour-consuming and costly for routine use (18). Natural Language Processing (NLP) in its various forms has been used to pick up association statistics for pharmacovigilance purposes from Electronic Health Record (EHR) (19). Moreover, NLP proved to be an effective technique for detecting a wide range of adverse events from discharge summaries and outperformed previous automated adverse event detection methods (18).

Text mining of medical records using an existing software product, SAS® Text Miner (SAS Institute Inc.), was used in Mayo Clinic Rochester hospitals to detect whether the elements of follow-up appointment arrangements were documented in a large volume of hospital discharge summaries (20). It was shown that the text mining could accurately and rapidly identify individual appointment elements and thus save considerable resources required for manual abstraction for performance assessment and quality-related research.

Our study showed that the detection of triggers using text mining almost reached the level of manual review, which was considered the gold standard. The study was designed to assess only unanalysed data, i.e. data that was not used for model fitting or model selection. This was done to avoid any study bias that would have occurred if the same data had been used in a repetitive manner. However, this was also a weakness of our study because the occurrence of individual triggers remained low in the remaining data material. Although there was a modest reduction in the rate of detection of the triggers it seemed clear that with a more fine-tuned text mining process and with a greater number of patient records detection would be further improved.

Table 4. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of individual triggers found in text mining review.

Trigger	Sensitivity (95 % CI)	Specificity (95 % CI)	PPV (95 % CI)	NPV (95 % CI)
Deterioration of neurological condition	69.2 (42.4 – 87.3)	80.0 (72.8 – 85.7)	23.7 (13.0 – 39.2)	96.7 (91.7 – 98.7)
X-ray imaging	75.0 (53.1 – 88.8)	80.4 (73.0 – 86.2)	35.7 (23.0 – 50.8)	95.7 (90.3 – 98.2)
Unplanned return to surgery or endovascular procedure	66.7 (41.7 – 84.8)	87.4 (81.0 – 91.9)	35.7 (20.7 – 54.2)	96.2 (91.3 – 98.4)
Treatment complication	75.0 (30.1 – 95.4)	81.9 (75.1 – 87.2)	9.7 (3.4 – 24.9)	99.2 (95.7 – 99.9)
Wound problem	100.0 (34.2 – 100.0)	98.1 (94.5 – 99.3)	40.0 (11.8 – 76.9)	100.0 (97.5 – 100.0)
Cerebrospinal fluid circulation disturbances	71.4 (35.9 – 91.8)	93.5 (88.4 – 96.4)	33.3 (15.2 – 58.3)	98.6 (95.1 – 99.6)
Infection	60.0 (38.7 – 78.1)	89.2 (83.0 – 93.4)	44.4 (27.6 – 62.7)	93.9 (88.5 – 96.9)
Readmission	77.8 (45.3 – 93.7)	81.9 (74.9 – 87.2)	20.6 (10.4 – 36.8)	98.4 (94.3 – 99.6)
Prolonged intensive care unit treatment	100.0 (34.2 – 100.0)	84.7 (78.3 – 89.5)	7.7 (2.4 – 24.1)	100.0 (97.2 – 100.0)
Intubation or reintubation	100.0 (34.2 – 100.0)	89.2 (83.3 – 93.1)	10.5 (2.9 – 31.4)	100.0 (97.3 – 100.0)
Fluid balance disturbances	75.0 (46.8 – 91.1)	82.3 (75.4 – 87.6)	25.7 (14.2 – 42.1)	97.6 (93.1 – 99.2)
Electrolyte balance disturbances	100.0 (20.7 – 100.0)	86.7 (80.5 – 91.1)	4.6 (0.8 – 21.8)	100.0 (97.3 – 100.0)
Blood pressure disturbances	87.5 (52.9 – 97.8)	80.5 (73.5 – 86.1)	19.4 (9.8 – 35.0)	99.2 (95.5 – 99.9)

Conclusions

This study showed that the modified Global Trigger Tool is capable of finding triggers that may predict the occurrence of adverse events in neurosurgical patient material. Automation of the method is technically possible, which will reduce the costs and the use of manpower. Moreover, automation will reduce reliance on human judgement and enable earlier identification and avoidance of potential harm.

References

1. Kohn L, Corrigan J, Donaldson M. To err is human building a safer health system. Washington, D.C: National Academy Press, 2000.
2. Wilson RM, Harrison BT, Gibberd RW, Hamilton JD. An analysis of the causes of adverse events from the Quality in Australian Health Care Study. *Med J Aust* 1999;170:411-5.
3. 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;324:377-84.
4. Thomas EJ, Studdert DM, Runciman WB, et al. A comparison of iatrogenic injury studies in Australia and the USA. I: Context, methods, casemix, population, patient and hospital characteristics. *International Journal for Quality in Health Care* 2000;12:371-8.
5. Bates DW. The costs of adverse drug events in hospitalized patients. Adverse Drug Events Prevention Study Group. *JAMA* 1997;277:307.
6. Classen DC. Adverse drug events in hospitalized patients. Excess length of stay, extra costs, and attributable mortality. *JAMA* 1997;277:301.
7. Griffin FA, Resar RK. IHI Global Trigger Tool. IHI Innovation Series white paper. Cambridge, Massachusetts: Institute for Healthcare Improvement; 2007. (Available on www.ihl.org)
8. Naessens JM, Campbell CR, Huddleston JM, et al. A comparison of hospital adverse events identified by three widely used detection methods. *Int J Qual Health Care* 2009;21:301-7.
9. Sharek PJ, Parry G, Goldmann D, et al. Performance Characteristics of a Methodology to Quantify Adverse Events over Time in Hospitalized Patients. *Health Serv Res* 2010 August 16 (Epub ahead of print).
10. de Wet C, Bowie P. The preliminary development and testing of a global trigger tool to detect error and patient harm in primary-care records. *Postgrad Med J* 2009;85:176-80.
11. A multicenter trial of the efficacy of nimodipine on outcome after severe head injury. The European Study Group on Nimodipine in Severe Head Injury. *J Neurosurg* 1994;80:797-804.
12. Macdonald RL, Kassell NF, Mayer S, et al. Clazosentan to overcome neurological ischemia and infarction occurring after subarachnoid hemorrhage (CONSCIOUS-1): randomized, double-blind, placebo-controlled phase 2 dose-finding trial. *Stroke* 2008;39:3015-21.
13. Soop M, Fryksmark U, Köster M, Haglund B. The incidence of adverse events in Swedish hospitals: a retrospective medical record review study. *Int J Qual Health Care* 2009;21:285-91.
14. Resar RK, Rozich JD, Simmonds T, Haraden CR. A trigger tool to identify adverse events in the intensive care unit. *Joint Commission Journal on Quality & Patient Safety* 2006;32:585-90.
15. Sari AB, Sheldon TA, Cracknell A, Turnbull A. Sensitivity of routine system for reporting patient safety incidents in an NHS hospital: retrospective patient case note review. *BMJ* 2007;334:79.
16. Classen DC, Lloyd RC, Provost L, Griffin FA, Resar R. Development and Evaluation of the Institute for Healthcare Improvement Global Trigger Tool. *Journal of Patient Safety* 2008;4:169-77.
17. Resar RK, Rozich JD, Classen D. Methodology and rationale for the measurement of harm with trigger tools. *Quality & Safety in Health Care* 2003;12:39-45.
18. Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. *J Am Med Inform Assoc* 2005;12:448-57.
19. Wang X, Hripcsak G, Markatou M, Friedman C. Active computerized pharmacovigilance using natural language processing, statistics, and electronic health records: a feasibility study. *J Am Med Inform Assoc* 2009;16:328-37.
20. Ruud KL, Johnson MG, Liesinger JT, Grafft CA, Naessens JM. Automated detection of follow-up appointments using text mining of discharge records. *Int J Qual Health Care* 2010;22:229-35.

5 Computerization of trigger tools: Evidence and possibilities

Persephone Doupi, MD, PhD.

Why use computerized triggers?

The idea of utilizing triggers for the identification of adverse events in patient records has its roots and origins in healthcare information systems. The pioneers in the field were the team of Classen et al. from the Later Day Saints (LDS) Hospital in Utah, who developed in the early '90s a system for the automated detection and characterisation of adverse drug events based on the HELP hospital information system (1,2).

The GTT and the other members of the IHI trigger tool family were developed a decade later, transferring the idea of the automated triggers to a paper-based environment (3-6). It would be safe to assume, that this sort of regression step was necessary, in order to address the limited availability of and know-how on health information systems in American health care provider organisations at the time⁵. The purpose of the IHI trigger tools is to establish a baseline level of harm (adverse events) in an organization and then, using statistical process control rules (7), collect data points over time to determine improvement. In other words, and in accordance with the AHRQ Trigger Terminology Definitions (8) the trigger tools of IHI are **accounting** ones (in the sense that they allow system-level monitoring of patient safety) and **retrospective**, since they identify the adverse event only after the episode of care has been concluded. Such trigger systems do have their place and role in the continuum of monitoring and improving patient safety. On the other hand, paper-based trigger tools also present certain limitations when contrasted to the possibilities opened up by automated or computer-enabled trigger systems.

At a minimal level of computerization, which is also the view presented by IHI (3), triggers - particularly medications and laboratory values – can be directly captured from a patient information system (once the random selection of records has happened), thus speeding up the review process. This is also the idea (at least in part) underlying the Swedish MAG system⁶ (9). Such a trigger system can be viewed as only the 'first generation' of computerized trigger systems, since the objective is still the

5 Similar challenges remain in present time US, but the clear political commitment to promote HIT uptake is currently trying to change that.

6 At least in theory, the MAG could also be applied to all the records of a healthcare organization, rather than a small random sample, but that option presents both technical and ethical challenges.

post-hoc identification of harm. As the sophistication and capacity of electronic systems improves, the closer the implementation of trigger systems moves to the possibility of intervening with an adverse patient safety event before it causes harm to the patient (**concurrent and real-time systems**) or even before it even happens (**interventionist trigger system**). In practice, the necessary infrastructure of EPRs and trigger-based modules can be utilized for a dual purpose – “classic” triggers for post-event assessment as a metric of patient safety measures impact and preventive triggers as a means of patient safety promotion and support. The latter can focus either on the management/organisational perspective or on the care of individual patients, or both.

The Finnish context

In 2008, a small pilot project (partly reported on in Chapter 4) was launched in the context of the Nordic Council Minister’s project on Healthcare Quality Indicators. The purpose was to explore the applicability of the Global Trigger Tool to the Finnish setting. The pilot was a collaboration between TAYS Neurosurgery department and (former) STAKES Information Division (since 2009, THL Information Department). The level of computerized applications uptake in Finnish healthcare settings has reached complete coverage already for some time (10). Therefore at the planning phase of the pilot it became apparent that a computer-based or automated trigger detection approach would be (at least in the long run) of more interest rather than the paper-based method of IHI. In addition, only through the use of computer-based approaches it is possible to achieve real-time monitoring and hence a preventive use of patient safety triggers, by incorporating them into everyday clinical practice and use of patient record systems.

At the time of the pilot project start, trigger tool approaches to patient safety monitoring had not yet been tested in Finland. The patient safety project in the Southwest Hospital District begun in 2009 (11), and piloting/testing of the GTT in the Hospital District of Vaasa started a year later, in 2010 (12). These hospital pilots of the GTT method in Finland are already computer-assisted: even though the remaining process is manual, the sample of patient records to be reviewed is derived from each hospital’s databases, and presented to reviewers electronically. Also, at least part of the assessment process is executed with the help of IT-applications.

In order to explore the requirements and potential barriers for the implementation of computerized trigger tools in Finnish hospitals we undertook the following activities:

- a literature review, focusing on the use of electronic patient record systems in the development and implementation of computerized trigger tools. The aim was to collect and analyse existing experiences and approaches utilized thus far, particularly from the point of view of their implications for, and dependence on structured electronic patient record data.

- an assessment of the readiness of Finnish EPR structures to support trigger-based approaches to patient safety, through the cross-mapping of GTT triggers vs. the trigger set proposed by the neurosurgery pilot and vs. the Finnish minimum data set elements (ydintiedot, in Finnish)⁷.

The literature review methodology and results are presented and discussed in the forthcoming paragraphs, while the materials concerning the assessment of the Finnish EPR structures (particularly the minimum data set) are presented in Annex 3.

Literature Review on computerized trigger tools

Methodology

Collection of references on computerized trigger tools was performed in three distinct phases, both in terms of their timing and purpose, as well as their methodology. The first phase of reference identification (2008-2009) had the purpose of scoping the field, as well as supporting methodologically the neurosurgery pilot. As a result, a broader scope of publications concerning both paper trigger tools and information technology enabled solutions were identified and reviewed⁸. The starting point for identifying publications concerning trigger tools was the website of their parent organisation, the IHI (13) as well as other pertinent patient safety organisations, such as the AHRQ, the Health Foundation in the UK etc (14, 15). In addition, publications were sought through PubMed Medline as well as through Google, with the terms ‘trigger tools’ and ‘patient safety AND trigger tools’.

In the second and third phase, publications on trigger tools, now with explicit emphasis on the role of the electronic patient record, were retrieved through a search of Ovid MEDLINE and EBM databases. During these phases, the same search strategies (available in Annex 1) were applied with an interval of a year, in June 2010 and June 2011.

Throughout the three stages of materials identification, we also used what is known as a snow-balling technique. Relevant publications on trigger tools that were present in the reference list of articles already reviewed were also included for consideration. In addition, the review material was complimented by publications identified through the ‘Related articles’ function of PubMed Medline, and publications identified through other on-going research work on the relation of patient safety and health information technology. Articles were included for analysis if they were available in English, and concerned computerized trigger tools in the hospital (inpatient) setting,

⁷ Ydintiedot: the Finnish minimum data set for Electronic Health Records, which needs to be recorded in structured and coded form (16).

⁸ The publications concerning the GTT and other IHI trigger tools are referenced and reviewed in Chapter 2 of this report.

hence excluding primary care and ambulatory patients. Publications concerning data mining solutions were handled separately, since they were not the primary focus of the work (and thus not appropriately reflected in the applied search strategy) and are discussed later on in this chapter.

From a total of 122 publications potentially relevant for review, a total of 15 articles were eventually included for analysis after full text review (see Annex 2 for details of the included studies).

Essential components of computerized trigger tool systems

Computerized trigger systems may vary in the clinical area they target and the rules they employ, but are generally built using the same basic components:

- **a repository of patient data**, the extent of which may vary depending on the type of adverse events targeted and the stage of development of the particular hospital information and electronic patient record system;
- **a rules database**: starting from the original LDS Hospital system rules, followers have expanded and modified the set that seems to be steadily growing in size to 52, 69 and even 130 rules (not all of which however are in use).
- **a notification system and procedure for responding to an alert or trigger**. This element concerns a combination of technology and human resources in the form of workflow planning. The electronically generated notification of a possible adverse event is usually delivered (through various communication means and settings) to a team of primary reviewers (often nurses and/or pharmacists) and the final decision on the presence and management of an adverse event is made by an experienced physician (who may also be the treating physician of the patient in question).

Features of computerized trigger systems

The majority of the reviewed studies on automated trigger systems, 11 studies out of 15 in total (1-2, 17-30, see Table 12 for details), concerns the detection of adverse drug events, mostly in general adult populations, as well as in paediatrics (2 studies, concerning the same system). Other application areas are: respiratory care (2 studies), cardiovascular procedures (1 study) and general adverse events (1 study).

Table 5. Overview of reviewed studies on computerized trigger tool systems

Article	Focus	Setting	System description
Classen 1991	ADEs New method for detection and characterization	LDS Hospital, Salt Lake City, Utah, USA - 520-bed tertiary facility. 18-month evaluation period	HELP HIS (clinically operational for over 15y). Computerized medical record with integrated database from various sources and interactive modular knowledge base for data analysis. Combined automated detection and computer-supported voluntary reporting.
Classen 1992	ADE detection	LDS Hospital Salt Lake City, Utah, USA. Part of a 23 hospital network with a computerized Master Patient Index, containing permanent, online EPR abstract of all patients admitted to any of the hospitals. ADE stored, used for partial prevention of future occurrences	HELP - HIS. Development over 20years, in use in several hospitals. Patient information (except physician histories, physical examinations and progress notes) stored in coded format, allowing for easy data retrieval and access to knowledge base for AI applications (DRG, ICD9-CM). Modular knowledge base containing algorithms that can be data driven or time driven.
Greenway 1993 ⁹	Respiratory care	LDS Hospital, Salt Lake City, Utah, USA - 520-bed trauma center, associated with Univ. of Utah Medical School.	HELP system + integrated database (info from all medical & nursing department, plus financial data). Utilization in respiratory care since 1983 - Functions (incl. alerting system) supported by active Medical Informatics Department nationally recognized in medical computing and logic application.
Camacho 1998	Cardiovascular procedures - comparison of three detection methods: UCDCSS, HealthCare Financing Administration's generic screens and HMPS screening criteria	Johns Hopkins hospital, large teaching hospital specializing in tertiary care. A random sample of 451 medical records of inpatients who underwent cardiovascular procedures were screened by trained nurses and also examined by physicians masked to screening results.	The version of Uniform Clinical Data Set System (UCDCSS) used was a computer-assisted method in which the nurses entered clinical data abstracted from the medical chart. Algorithms built into the software identified possible deviations from standard care requiring medical audit.
Jha 1998	ADEs – development and assessment of computer-based monitor (comparison with intensive record review & stimulated voluntary reporting)	Brigham and Women's Hospital (tertiary hospital), Boston, USA. Prospective cohort study (all patients admitted to 9 medical and surgical units) during 8 months	Internally developed information system - Brigham Integrated Computer System; includes a computer-based event detection application using rules to detect a variety of clinical events. Daily report of alerts, first review by trained reviewer followed by classification by physician - evaluation for severity and preventability
Levy 1999	ADE's computerized detection: Implement and measure the effects of automatic computerized laboratory signals (ALS) as a detection support tool of ADRs in a hospital	Prospective observational study of 192 patients (199 sequential medical admissions) during a 2-month period (April to May 1997) in a 34-bed medical ward at the Hadassah University hospital, Jerusalem, Israel.	Daily distribution to staff physicians of lists of automatic signals generated from computerized laboratory data as potential indicators of ADR's. Patient charts were reviewed by the clinical pharmacology team for ADRs and to see whether those were recognized by the staff physicians.
Lederer 2005	ADEs - Anticoagulation treatment with warfarin	Novant Health System, seven hospitals, three nursing homes etc. serving 3,4 mil residents in western North Carolina, USA. Trial data collection begun April 2002, baseline data established starting July 2002 - data collection continued for the rest of 2002 and 2003.	Each month a 7-day sampling period (first seven working days of the month), HIS generated triggers used by clinical pharmacists to identify patient charts for review. Harm classification using MERP.

⁹ After full-text review of this article, also the related publication of Elliot (20) was retrieved and reviewed (although not originally identified by the search strategy), because it contained more specific information on the system in question.

Article	Focus	Setting	System description
Kilbridge 2006	ADEs - hospital setting (comparison between academic & community hospital)	Duke University Health System; Durham Regional Hospital (16000 patients/year) & Duke University Hospital (36000 patients/year)	Durham Regional Hospital - Siemens Medical System components; Duke Univ. Hospital - combination of custom-built and commercial products, incl. McKesson pharmacy system and locally developed core HIS. Components common to both hospitals: single Cerner LIS and common clinical data repository receiving real-time data from principle patient care IS (lab, radiology, other diagnostic reports). Custom-built computerized ADE surveillance system. DB2 database residing on mainframe computer that operates Duke UH core HIS. Rule engine written in PL/I and Web-based evaluation application in C++. System queries data daily to identify potential ADEs based on 69 rules. Rules reviewed for PPV - low-yield rules removed from operation, rules added or modified based upon early experience prior to the study. Duke UH has an online incident reporting system since 2004 (about 40% of reports medication related).
Szekendi 2006	AE and medical errors occurring during hospitalization (of at least 48 hours and still not discharged at time of review)	725 bed academic medical center in Chicago, Illinois, USA. Excluded: neonates, OR, labor and delivery, ER. No pediatric department. Target: review of 300 records over a 3-month period - approx. 8% of eligible triggers reviewed (493 triggers in total)	Combined use of paper (physicians' progress notes, medication administration record, nursing flow sheets) and electronic records (lab results, radiology results, some nursing documentation): Record not fully electronic at time of study - text searches of progress notes or discharge summaries not possible. Already in use in the hospital: Voluntary incident reporting system & retrospective administrative data mining (AHRQ Patient Safety Indicators). Adaptation of list of triggers developed during pilot project, incl. abnormal lab values, high risk medications, and medications used as antidotes.
Seger 2007	ADE's computerized detection using a combination of medication and laboratory data	6 months (July 1 st to December 31 st 2002) of retrospectively collected medication and laboratory data from a 140-bed community hospital in Boston, MA, USA, and rules from a computerized knowledge base. AIM: assess whether resulting alerts would have allowed a clinician to prevent or lessen harm related to medication toxicity. Randomly selected 11% of high-or critical priority charts (58, of which 56 were available); determined frequencies of ADEs and preventable ADEs.	System application based on commercially available components. (Dynamic Pharmaco-Monitoring System applied through Microsoft products).
Kilbridge 2008	Pediatrics - preliminary evaluation of ADEs	SLCH (St. Louis Children's Hospital, St. Louis, USA) – principal pediatric teaching hospital for the Washington University School of Medicine. Follow up for 22 consecutive days on February 2008 - oncology patients excluded at original evaluation.	Expert (rules based) computer program was used to analyze combinations of data (see earlier ref). Rule set was constructed during previous work in adult hospital and modified to the pediatric environment.

Article	Focus	Setting	System description
Ferranti 2008	Comparison of voluntary ADE reporting (SRS) and computerized surveillance (ADE-S) in large academic hospitals	Adult, in-patient ADE from a retrospective 7-month period (1.12.2006 - 30.6.2007) evaluated and scored using standardized methodology. Labor and delivery, as well as pediatric units excluded. Duke University Hospital (DUH) - large, tertiary academic medical center Analysis: ADE's per 1000 patient days	Voluntary reporting (SRS): home-grown web based application - wide capture of event types Computerized surveillance: The DUH ADE-S system was deployed on November 1st, 2004 by an internal team of technical and safety experts. The system delivers an electronic daily report to a web-based surveillance application that details all triggers fired by the system.
Hera-sevich 2009	Acute Lung Injury (ALI) electronic surveillance	A total of 3,795 consecutive critically ill patients in nine multidisciplinary intensive care units (ICUs) of the Mayo Clinic in Rochester, Minnesota, a tertiary care teaching institution with two hospitals comprised of 1,900 inpatient beds and 201 ICU beds. Evaluation for a period of 4 months (July–November 2006)	Use of a near-real time copy of electronic medical records. A custom ALI electronic alert (ALI "sniffer") based on the European-American Consensus Conference Definition was developed and validated. It is a tool combining a near real-time query of the chest X-ray reading with arterial blood gas values ("ALI sniffer"). The multidisciplinary Epidemiology and Translational Research in Intensive Care (METRIC) database (METRIC Datamart) was used. This SQL-based integrative database accumulated data within 1 h from its entry into the EMR - linked demographic, monitoring, laboratory, intervention, and outcome data required for this study.
Kilbridge 2009	Pediatrics - implementation and evaluation - ADEs	SLCH- 250-bed hospital specializing in care of acutely ill pediatric patients. Principle pediatric teaching hospital for Washington Un. School of Medicine. Member of BJC Health-Care, 13-hospital integrated delivery system. 14,500 admissions annually, average LOS 3.4 days. Study population: All patients admitted between Feb.1 and July 31 2008, excl. oncology	Modified rules-based computer program performing real-time surveillance of patient data from SLCH clinical systems. Automated Guideline Monitor module evaluates data against the rules - alerts generated are displayed on Web-based user interface for evaluation by pharmacists.
Ramirez 2010	ADE's computerized detection through Automatic Laboratory Signals	La Paz University Hospital in Madrid, Spain, 1,365 bed tertiary care facility providing a wide range of services. From July 2007 to June 2008, all admissions to all wards were monitored by the Prospective Program from Laboratory Signals at Hospital - PPLSH (incl. data of patients who died in the emergency ward).	Prospective pharmacovigilance program based on systematic detection of predefined abnormal laboratory values, using the LIS of the hospital (PPLSH). At the time of the study, EMRs in the hospital included: all laboratory data, images and other exploratory results, previous medical reports, and discharge summaries. A specific database application was developed within the Integrated Laboratory System (in use in the hospital since 2003) to detect predefined ALSs. All ALSs were retrieved systematically.

All of the studies concerned systems developed and implemented in American hospitals, with the exception of two: one in Israel (22) and one in Spain (30). The majority of computerized trigger tool systems have been developed on the basis of customised, locally developed hospital information systems, in certain leading healthcare institutions. These are healthcare provider organisations which often combine at least two features: long standing tradition in medical informatics, represented by an in-house development team, and large scale programmes on patient safety.

The type of triggers used by each application varies, depending on the targeted patient population and the focus of the system, but generally trigger tools rely on the combination of data from various systems, most importantly laboratory values and pharmacy data, combined with patient demographics/basic clinical care data (see Table 13, provided in Annex 2). Of the systems reviewed, six were implemented in (almost) real-time, allowing for possible interventions in the care of individual patients, in case an adverse event was identified.

Discussion

Based on the articles we have reviewed on computerized trigger tool systems, it is apparent that the area of implementation where the most experience has accumulated is that of adverse drug events (ADEs). The choice of ADEs as a target area for patient safety improvement efforts makes sense, given the frequency of this type of events and their contribution to the overall burden of patient safety incidents (31). At the same time, this also means that developing systems with a broader focus will eventually present even more challenges (starting from the type of variables that need to be followed and their standardized representation). Pharmacy and laboratory data are by nature easier to handle automatically, since they have an inherent structure, are often numeric, and also coded. However, certain types of adverse events (such as those related to procedures or diagnostic errors) rely on symptoms and signs for their detection, which in turn often reside in the free text areas of electronic patient records.

The fact that the majority of identified computerized trigger systems has been developed and tested in large US hospitals, representing organisations which belong to the forefront of both informatics and patient safety development, raises concerns for the feasibility of automated trigger tools in smaller hospitals, and through commercially available, off-the-shelf applications. On the other hand, it must be noted that the systems presented in the reviewed publications reflect an earlier stage of development and maturity in hospital information systems, which is not necessarily representative of today's state of progress. Indicatively, in a recent interview, Dr. Classen communicated that IHI has already proof-tested the automation of the GTT in all 'leading EMR vendors at various health systems' (32).

A detailed comparison of the types of triggers used in the individual systems was beyond the scope of this work. An important observation, however, is that in many systems rules and triggers were continuously updated and modified in order to maintain their currency (vs. practice standards) and increase their accuracy. Although the benefits or even necessity of this approach may seem evident, at the same time it poses a challenge in assessing the performance of such systems over time (since the object of evaluation is not stable, and features which may have a critical impact on its success or failure are modified during the period of study). The need for fine-tuning and up-

dating of a computerized trigger system should also be accounted for when considering the maintenance demands of such an application.

Critical Success Factors

The performance of an automated trigger tool, in addition to the selection of appropriate triggers, relies on simple and reliable access to relevant clinical data. Ideally data would be timely, complete and in coded form (33). Data access is in direct dependence to the data storage and data integration solutions deployed by each organisation. The data repository may be distributed (i.e. a compilation/federation of separate databases of the same organisation), or centralized, where all relevant data are collected in one database (a data mart or data warehouse (34)). In the former case, data needs to be extracted at run-time from several other systems, such as eg. the hospital pharmacy database and the hospital laboratory system database, which in turn requires more time and does not allow for data quality control. Mixed approaches are also possible, where data residing in separate databases are first collected into a dedicated database for further processing supporting the trigger application. Having a single source of clinical data on which the trigger tool is applied provides a clear advantage, as demonstrated by the HELP system studies, as well as the experience of the MAG in the Karolinska University Hospital in Sweden.

In tandem to accessibility is data quality, which affects directly and significantly the effectiveness of a computerized trigger tool. The success of an automated trigger tool will largely rely on the completeness and accuracy of documentation in the patient record. What has not been documented in the patient's record cannot be found – particularly when searched for automatically.

Organisational and human factors also play a critical role. Leadership commitment to patient safety initiatives is essential for implementing a computerized trigger tool, given the significant amount of resources and sustained effort needed, both for training and introduction, as well as regular use and maintenance. As explained earlier, the establishment of a notification and reaction system is one of the key components of a computerized trigger tool system, if not the most critical one for its operational success. Research indicates that it is precisely in the area of human logistics and workflow that automated trigger tool solutions confront a too big barrier (35). The commitment of hospital staff to improve quality of care and correspondingly modify their behaviours, as well as clinician involvement in system development, are both central to the system's acceptance and integration with practices. Verifying accuracy of the system is also crucial in order to avoid too many false alarms and ensure relevance of system output for clinical decisions.

Reviewing the literature on computerized trigger tools has also confirmed the finding emerging from studies on the paper-based variant, i.e. that different methods of adverse event detection identify different types of events, with little overlap among

them. The most successful strategy in terms of resource demand and yield seems to be the combination of computer-based alerts and voluntary reporting.

The role of data/text mining

Data mining approaches have been explored in the context of patient safety with initially promising results. Although it was not an explicit focus area for this review, several publications were identified on the application of data mining techniques for patient safety purposes. The subject of data mining is anyhow briefly discussed here, because of the recent interest to utilize the method in the context of computerized trigger tools (as in the neurosurgery pilot reported in Chapter 4 and the Swedish MAG application). Most of the reported experiences with data mining concern limited parts of the patient record, most notably discharge summaries (36-40). Examples have also been reported with patient falls detection through incident reports analysis, as well as for purposes of pharmacovigilance, but these are not directly related to trigger tool methodology (41-42).

The process of utilizing data mining on large data sets (either coded or free text) for the identification of specific data or the discovery of previously unknown patterns and associations involves several steps and components, as documented by Hripcak et al (43):

1. Target events definition
2. Repository – Full clinical repository or a purposely defined subset.
3. Natural language processing (NLP) – parse the narrative data and create a fully coded repository.
4. Queries – Detect and classify errors. They may be generated manually or automatically.
5. Verification – Verify the accuracy of the detection and classification by manual review, thus calculating performance and adding to the database of known errors.
6. Error description – Use a systems approach or a cognitive approach to describe the newly detected errors.
7. Feedback and correction based on errors identified.

The aforementioned steps translate to a set of requirements, which are largely the reason for the slow uptake of data mining applications. Data mining requires specialised subject knowledge, skill and time to develop algorithms and NLP applications. In turn, algorithms require installation and instruction by experts, and their use most often relies on the integration of heterogeneous sources of data, an obstacle documented also in the efforts of Kaiser Permanente to automate the GTT (44). In addition to that, when the context is a trigger tool application the need of human review of results is not by-passed – rather it is focused to the second phase of the process. Hence it is still unknown whether the costs to implement such programs are offset by savings from eliminating the first phase of manual record review.

In spite of the aforementioned challenges, it is reasonable to expect that the use of data mining in healthcare and in the context of patient safety initiatives specifically will increase, on the one hand because of the sheer amount of valuable free text contained in the EHRs themselves, as well as in other data sources (such as patient safety reporting systems), and on the other hand because of the constant evolution and improvement of relevant applications. It should be kept in mind that any sort of data or text mining approach should not be seen as alternative, but rather as complementary to the use of structured data, since data mining also depends on the use of a reliable and suitable vocabulary/terminology set (which, in the Finnish context, could be in part or in whole based on the minimum data set).

Future target: Continuum of patient safety monitoring

In the course of further development of trigger tool systems, there is consensus that measurements achieved through real-time surveillance, which is the realm of computerized trigger tools, will provide an additional safety net allowing **intervention** to mitigate adverse events on the level of the individual patient.

Decision support systems in turn, focus on **prevention**, by providing advice at the time of decision making and can also be used (in connection to clinical guidelines) for the development and implementation of ‘negative’ triggers – i.e. warnings that an action consistent with clinical guidelines and high standard of care has not taken place, thus targeting errors of omission.

Retrospective surveillance and incident reporting focus on the **detection** of near misses and adverse events once they have occurred. The trend and requirement in this area is to move beyond manual methods to **both** electronic data analysis and automated tools for notification (45).

The development towards producing a unified view of patient safety through the combination of different methods and tools for its monitoring and analysis, must also be reflected in the development of suitable data sets to support it, as the authors of the report “Patient Safety: Achieving a new standard of patient care” had envisioned (46):

“...various approaches to adverse event detection ... demonstrate that it is not possible to simply identify a small set of clinical data elements specifically for adverse event detection, especially when addressing potential injuries due to errors of omission as well as injuries due to errors of commission. On the contrary, a broad range of data elements encompassing demographic information, signs and symptoms, medications, test results, diagnoses, therapies, and outcomes are required to: (1) detect adverse events through voluntary and mandatory reporting, chart review, and automated surveillance; (2) implement performance measures; and (3) use decision support tools (e.g., computerized physician order entry). Thus, comprehensive clinical and patient safety data are necessary for adverse event detection and monitoring.”

Implications of the review findings for trigger tool activities in Finland

Finland is in a good starting point with the regard to the requirements for implementation of computer-based trigger tools, since several of the key components for such an endeavor are already in place:

- The country possesses extensive computerization and experience with use of IT systems in healthcare.
- A unique patient identifier exists, that allows bringing together disparate data on the same individual (provided certain security and data protection requirements are fulfilled).
- Coded data for electronic patient record systems are widely used and expanding. The means through which the minimum data set is made available and supported through EPR systems are still inadequately investigated. The key challenge will be ensuring the availability of high quality coded data in real-time, by efficiently supporting clinical documentation processes.
- On-going work on the Finnish minimum data set for electronic patient records, particularly in view of the latest efforts to incorporate processes and workflow in EPR structures, allows the opportunity to address also the needs of data standardisation for patient safety purposes.
- The EBMeDS research project on decision support systems (47) has produced positive experiences that would warrant investigating further the possibilities for coordinated activities between patient safety/trigger tool initiatives and decision support system implementation.

The investigation and further development of trigger tools should be considered also in the light of the first national patient safety strategy for Finland, and the latest updates to healthcare services legislation. The strategy encouraged organisations to develop and implement reporting systems at the local and regional level (48). The trend towards systematic patient safety monitoring was later on confirmed with the new Health Services Act (Terveysturvallisuuslaki) (49) and the associated Patient safety and Quality Decree (50) (both enacted in May 2011), that demand the following of patient safety levels, thus underlining the need for (at least) organisational level solutions.

During the last four years, voluntary incident reporting systems have established their presence in many Finnish healthcare service organisations, gradually also beyond the hospital environment. This has been a positive development, but given the fact that reporting systems capture only a fraction of actual adverse events, it will be necessary to think of additional methods to cover a largest part of the patient safety spectrum.

In the near term, trigger tools can function as a possible metric for following up the success rate of introduced patient safety promoting interventions. In the longer run, and in combination with the increased uptake and evolution of health information systems, the development and implementation of concurrent and subsequently real-time trigger systems should be the priority.

References

1. Classen DC, Pestotnik SL, Evans RS, Burke JP, Battles JB. Computerized surveillance of adverse drug events in hospital patients. *JAMA* 1991;266:2847-51.
2. Classen DC, Pestotnik SL, Scott Evans R, Burke JP. Description of a computerized adverse drug event monitor using a hospital information system. *Hospital Pharmacy* 1992;27:774-783.
3. Griffin FA, Resar RK. IHI Global Trigger Tool for Measuring Adverse Events. IHI Innovation Series white paper. Cambridge, MA: Institute for Healthcare Improvement; 2007.
4. Intensive Care Unit (ICU) Adverse Event Trigger Tool, VHA/IHI, January 2002. Available for download at: <http://www.ihl.org/IHI/Topics/CriticalCare/IntensiveCare/Tools/ICUAdverseEventTriggerTool.htm>
5. Surgical Trigger Tool for measuring Perioperative Adverse Events. IHI, August 2006. Available for download at: <http://www.ihl.org/IHI/Topics/PatientSafety/SafetyGeneral/Tools/SurgicalTriggerTool.htm>
6. Rozich JD, Haraden CR, Resar RK. Adverse drug event trigger tool: a practical methodology for measuring medication related harm. *Qual Saf Health Care* 2003;12:194-200.
7. Thor J, Lundberg J, Ask J, Olsson J, Carli C, Härenstam KP, Brommels M. Application of statistical process control in healthcare improvement: systematic review. *Qual Saf Health Care*. 2007 Oct;16 (5):387-99.
8. Triggers and Targeted Injury Detection Systems (TIDS) Expert Panel Meeting. Conference Summary. AHRQ, July 2008. Available at: <http://www.ahrq.gov/qual/triggers/triggers.pdf>
9. Modified Automated GTT. Information brochure (in Swedish). Available at; <http://www.uppdragsguiden.sll.se/upload/Utveckling/Kvalitetsarb/Patientsakerhet/GTT/MAG%20infobrochyr%20A4.pdf>
10. Hämläinen P, Reponen J, Winblad I. eHealth of Finland – Checkpoint 2008. National Institute for Health and Welfare. Report 1/2009, ISBN 978-952-245-002-9, ISSN 1798-0070, Helsinki 2009.
11. Peltomaa K, Savunen T, Kantola I, Gröndahl W. VSSHP:n potilasturvallisuusprojekti 2009 – 2011: kokonaisvaltainen kehittämishanke. In: Honkanen V, Härkönen M, Jylhä V, Saranto K (Eds). Potilasturvallisuusutkimuksen päivät 2011, Esitysten tiivistelmät, Helsinki 2011. p. 33.
12. Kinnunen M, Helovuola A. Kehittämissuunnitelma jatkuvaan ja suunnitelmalliseen potilasturvallisuuden edistämiseen. In: Honkanen V, Härkönen M, Jylhä V, Saranto K (Eds). Potilasturvallisuusutkimuksen päivät 2011, Esitysten tiivistelmät, Helsinki 2011. p.31.
13. <http://www.ihl.org/knowledge/Pages/Tools/IntrotoTriggerToolsforIdentifyingAEs.aspx>
14. AHRQ Patient Safety Tools and Resources. <http://www.ahrq.gov/qual/pstools.htm>
15. Health Foundation - <http://www.health.org.uk/areas-of-work/topics/patient-safety/>
16. Country brief Finland. eHealth Strategies project, Country report, 2010. Available at; http://ehealth-strategies.eu/database/documents/Finland_CountryBrief_eHStrategies.pdf
17. Kilbridge PM, Campbell UC, Cozart HB, Mojarrad MG. Automated surveillance for Adverse Drug Events at a community hospital and an academic medical center. *J Am Med Inform Assoc*. 2006;13:372-377.
18. Jha AK, Kuperman GJ, Teich JM, Leape L, Shea B et al. Identifying Adverse Drug Events: Development of a computer-based monitor and comparison with chart review and stimulated voluntary report. *J Am Med Inform Assoc*. 1998;5(3):305-314
19. Greenway L, Jeffs M, Turner K. Computerized management of respiratory care. *Respir.Care* 1993 Jan;38(1):42-53
20. Elliot CG. Quality assurance for respiratory care services: a computer-assisted program. *Respir Care* 1993; 38(1):54-9
21. Camacho LA, Rubin HR. Assessment of the validity and reliability of three systems of medical record screening for quality of care assessment. *Med.Care* 1998;36(5):748-751
22. Levy M, Azaz-Livshits T, Sadan B, Shalit M, Geisslinger G, Brune K. Computerized surveillance of adverse drug reactions in hospital: implementation. *Eur.J.Clin.Pharmacol*. 1999;54(11):887-892
23. Lederer J, Best D. Reduction in anticoagulation-related adverse drug events using

- a trigger-based methodology. *Joint Commission Journal on Quality & Patient Safety* 2005;;31(6):313-318
24. Szekendi, M K; Sullivan, C; Bobb, A; Feinglass, J; Rooney, D; Barnard, C; Noskin, G A Active surveillance using electronic triggers to detect adverse events in hospitalized patients. *Quality & Safety in Health Care* 2006; 15(3):184-190
 25. Seger AC, Jha AK, Bates DW. Adverse drug event detection in a community hospital utilising computerised medication and laboratory data. *Drug Safety* 2007;30(9):817-824
 26. Kilbridge PM, Noirot LA, Reichley RM, Heard KM, Dunagan WC, Bailey TC. Computerized surveillance for adverse drug events in a pediatric hospital. *AMIA Annual Symposium Proceedings/AMIA Symposium.*:1004, 2008
 27. Ferranti J, Horvath M, Cozart H, Whitehurst J, Eckstrand J, Pietrobon R, Rajgor D, Ahmad A. A Multifaceted Approach to Safety: The Synergistic Detection of Adverse Drug Events in Adult Inpatients. *Journal of Patient Safety.* 2008; 4(3):184-190
 28. Herasevich V, Yilmaz M, Khan H, Hubmayr RD, Gajic O. Validation of an electronic surveillance system for acute lung injury. *Intensive Care Med.* 2009 Jun;35(6):1018-1023
 29. Kilbridge PM, Noirot LA, Reichley RM, Berchelmann KM, Schneider C, Heard KM, Nelson M, Bailey TC. Computerized surveillance for adverse drug events in a pediatric hospital. *JAMIA* 2009; 16(5):607-612
 30. Ramirez E, Carcas AJ, Borobia AM, Lei SH, Pinana E, Fudio S, et al. A pharmacovigilance program from laboratory signals for the detection and reporting of serious adverse drug reactions in hospitalized patients. *Clinical Pharmacology & Therapeutics* 2010;87(1):74-86
 31. Kohn LT, Corrigan JM, Donaldson MS. *To err is human: building a safer health care system.* Washington DC: National Academy Press; 1999
 32. AHRQ Morbidity and Mortality Rounds on the Web. In *Conversation with David C. Classen.* April 2012. Available online at: <http://webmm.ahrq.gov/printviewperspective.aspx?perspectiveID=119>
 33. Haug PJ, Rocha BHSC, Evans RS. Decision support in medicine: lessons from the HELP system. *Intern J Med Inform* 2003;69:273-284
 34. Protti D. How business intelligence is making healthcare smarter. *World View Report 2*, 2010. Available at: <http://webarchive.nationalarchives.gov.uk/20100408203612/http://www.connectingforhealth.nhs.uk/newsroom/world-view/protti10>
 35. A V Sorensen, M I Harrison, H L Kane, A E Roussel, M T Halpern, S L Bernard. From research to practice: factors affecting implementation of prospective targeted injury-detection systems. *BMJ Qual Saf* 2011;20:527e533. doi:10.1136/bmjqs.2010.045039
 36. Murff HJ, Forster AJ, Peterson JF, Fiskio JA et al. Electronically Screening Discharge Summaries for Adverse Medical Events. *J Am Med Inform Assoc.* 2003;10:339-350. DOI 10.1197/jamia.M1201
 37. Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. *J Am Med Inform Assoc.* 2005; 12:448-457. DOI 10.1197/jamia.M1794
 38. Forster AJ, Andrade J, van Walraven C. Validation of a Discharge Summary Term Search Method to Detect Adverse Events. *J Am Med Inform Assoc.* 2005;12:200-206. DOI 10.1197/jamia.M1653
 39. Ruud KI, Johnson MG, Liesinger JT, Grafft CA, Naessens JM. Automated detection of follow-up appointments using text mining of discharge records. *Int J Qual Heal Care* 2010; 22 (3): pp. 229-235
 40. Wright A, Chen ES, Maloney FL. An automated technique for identifying associations between medications, laboratory results and problems. *Journal of Biomedical Informatics* 43 (2010) 891-901. doi:10.1016/j.jbi.2010.09.009
 41. Lee T, Liua C, Kuob Y, Mills M et al. Application of data mining to the identification of critical factors in patient falls using a web-based reporting system. *Int J Med Informatics* 2011; 80: 141-150. doi:10.1016/j.ijmedinf.2010.10.009
 42. Trifiro G, Pariente A, Coloma PM, Kors JA et al. on behalf of the EU-ADR group. Data mining on electronic health record databases for signal detection in pharmacovigilance: which events to monitor? *Pharmacoepidemiology and drug safety* 2009; 18: 1176-1184. DOI: 10.1002/pds.1836

43. Hripcsak G, Bakken S, Stetson PD, Patel VL. Mining complex clinical data for patient safety research: a framework for event discovery. *J Biomed Inform* 2003; 36: 120–130. doi:10.1016/j.jbi.2003.08.001
44. Snow D. Kaiser Permanente Experience with Automating the IHI Global Trigger Tool. Presented at AHRQ 2010 Annual Meeting September 29, 2010
45. Classen DC, Metzger J. Improving medication safety: the measurement conundrum and where to start. *Int J Qual Saf Health Care* 2003;15:Suppl (1):i41-i47.
46. Adverse Event Analysis - Chapter 6. In: Patient safety : achieving a new standard for care / Committee on Data Standards for Patient Safety, Board on Health Care Services ; Philip Aspden ... [et al.], editors. National Academy Press, 2004, ISBN 0-309-09077-6 (hardcover)
47. Evidence-Based Medicine electronic Decision Support (EBMeDS) - Homepage: http://www.ebmeds.org/ebmeds/ebmeds_home.asp
48. Ministry of Social Affairs and Health, 2009. Finnish Patient Safety Strategy 2009-2013. (In Finnish. Edistämme potilasturvallisuutta yhdessä. Suomalainen potilasturvallisuusstrategia 2009–2013). Available at: http://www.stm.fi/julkaisut/nayta/_julkaisu/1383571 .
49. Healthcare Service Act, 30.12.2010/1326. (In Finnish. Terveysturvalaki 30.12.2010/1326) Available at: <http://www.finlex.fi/fi/laki/ajantasa/2010/20101326>
50. Decree of the Ministry of Social Affairs and health for quality management and the implementation of the patient safety plan. Sosiaali- ja terveysministeriön asetus laadunhallinnasta ja potilasturvallisuuden täytäntöönpanon laadittavasta suunnitelmasta

6 In Conclusion

Patient safety measurement and monitoring is an area where challenges abound. There is no silver bullet – not one all-encompassing, reliable measurement method. The problem concerns both the local (as in organisational), as well as the regional and national level of aggregation and analysis. Full review of patient records has been generally viewed as the golden standard, but due to its resource intensiveness it has remained primarily a research undertaking. In Finland, such an undertaking has even been deemed not necessary, at least at the national level. Until now, estimates of the number of patient safety incidents and adverse events have been based on extrapolation from results obtained in other countries (1).

And yet, healthcare organisations need to have a way to establish and follow their patient safety levels, and the effectiveness of their improvement efforts. In spite of the fact that relevant data is reported to and collected by a large number of authorities (2), the options are not that many when direct and fast access to useful data is the target. The wide popularity and steadily growing adoption of the HaiPro incident reporting system in Finnish settings bears testimony to that. Incident reporting systems, however, have certain well established limitations - among others, their ability to detect only a fraction of actual patient safety incidents, as well as the limited traceability of individual cases (3). On account of these limitations, (voluntary) reporting systems cannot function as the single patient safety monitoring tool. Combining an organisational incident reporting system with other data sources is necessary for a more reliable picture and a better understanding of patient safety status and changes (4).

When it comes to the hospital environment, the GTT method is a reasonable alternative and substitute for full medical record review, which can be applied as part of regular operations. Taking up the tool as part of an existing or beginning patient safety programme is not particularly resource-demanding, at least for large hospitals. Training materials are available, although formal validation of their Finnish translation is still missing (5).

The issue of the method's validity remains, and can be partly addressed through training and consistent use of the same reviewing team. Use of the tool as a benchmarking instrument however, is not to be recommended – both because of reliability, as well as variability in methods of implementation. The GTT is reported to detect a larger number, as well as different type of events compared to other methods. This feature is seen as a strength of the method, which it largely is. However, users of the GTT should be aware of this finding's dependence on at least two factors:

- the definition of adverse event that is used by the method;
- the level of uptake of other methods of measurement/monitoring.

Existing data on both these factors comes from the US, where the language, landscape and culture of patient safety are very different from Europe. There is very scarce evi-

dence comparing the GTT with other patient safety methods in a European context, where, for example, the rate of voluntary organizational incident reports seems to be much higher than that reported by US healthcare providers.

Another limitation of GTT closely connected to its definition of an adverse event, is the equal treatment of events that could have been prevented and events that were completely unavoidable. By not looking into preventability, the method limits its potential to guide patient safety improvement interventions. Often, however, organizations implementing the method have chosen to augment it with an assessment of preventability – a process that should be incorporated in the learning strategy of GTT reviews.

In addition to making a serious long term commitment when starting with GTT use, hospital management needs to make provisions for supporting the learning processes essential to patient safety improvement. The GTT methodology strictly implemented and in itself, will not provide any other information than trends of level of harm overtime. That type of information is a rather rough indicator, which albeit useful, cannot answer the questions that are central to patient safety learning: what went wrong? Why? Was the event preventable, and if yes, what was done to prevent, mitigate or alleviate harm, both in the specific instance and in the future?

This is the type of information usually contained in incident reports – only those cannot be connected to the care of any particular patient, nor do they often capture same type or equally serious events as the GTT review. The answer to the need for patient safety learning is emerging in the form of ‘hybrid’ systems, combining the GTT methodology with applications similar to incident reporting systems.

All of the above observations concern the paper-based method of the IHI GTT – even if the source of data would be the electronic patient record. But what about electronic trigger tools? Automated trigger tools, in contrast to their paper counterparts, can be used either retrospectively or prospectively. In the former case, the tool simply transfers the GTT methodology to an electronic environment, and it acts primarily as a monitoring tool for hospital management. In the latter example, the tool can function both as a real-time management tool, and as a preventive instrument in the case of individual patients – at least in theory. In practice, these tools are still in an experimental or at best pilot phase, with the exception of some notable pioneer organizations.

Automated trigger tools can be based either on the use of free text in patient records - which is subsequently processed through data mining applications- or on coded and structured data. In the Finnish context, the primary development emphasis has been on the latter approach, as far as the overall health IT infrastructure development is concerned. The forthcoming national archive for EHR data (6), and the accompanying Information Management Service will bring significant changes in the functionality of existing and new health information systems. In the resulting renewal process, the needs and requirements of applications supporting and promoting the achievement of patient safety targets should not be forgotten. Awareness of the data structures and

system functionalities needed for patient safety is required both from application developers, as well as of healthcare providers. The aforementioned emphasis on structured data should not be understood to mean that data mining approaches should be abandoned – on the contrary. Free text entries of patient records hold information of extreme, often irreplaceable value, and tools to harness it are highly needed.

Real-time trigger applications, in addition to their technology demands, also place work flow arrangements under a lot of pressure, and it is precisely the area of fitness with the working environment where such applications have often failed thus far. Still, there is no doubt that automated tools are the key to the future of patient safety improvement work. Healthcare provider organizations should be encouraged to participate actively in the research and development of such tools, as a way of ensuring that the next generation of health-IT enabled applications will truly serve their purposes. As a minimum, patient safety requirements and needs in terms of health information systems features and functionality should be incorporated in the organizations' information technology development strategy and planning.

In the case of both paper and electronic trigger tool systems, there are a number of requirements which are common to the tools' successful implementation. These are:

- Definitions: particularly those of adverse event and preventability
- Quality of documentation: trigger tools depend on complete and accurate data. What has not been documented is close to impossible to find – particularly when searching automatically
- Leadership commitment, in terms of supporting and promoting patient safety culture, ensuring the necessary resources and actively enlisting clinician involvement
- Need to emphasize preventability. Prevention of patient safety incidents should be the ultimate goal of patient safety initiatives. Our notion of preventability has to be based on consensus, so that it is operationalized consistently. Once an agreed definition and understanding of preventable events has been achieved, it should be reflected on multiple levels: in the health information systems used for documentation and for support of patient safety targets, in the utilization of evidence emerging from monitoring methods, and in organizational culture, in terms of acceptance and alertness.

National level guidance, support and coordination on the aforementioned areas are painfully needed. Both in the area of paper-based trigger tools, as well as that of computerized ones, the collaboration of several actors will be necessary in order to achieve progress – among others the Ministry of Social Affairs and Health, THL – through its dedicated Patient Safety Programme, Valvira (the National Supervisory Authority for Welfare and Health) and the hospital regions. THL's Patient Safety Programme, launched officially in the fall of 2011 (7), has taken the matter into consideration in the context of developing a framework for national patient safety monitoring activities. In the meantime, sustaining the collaboration with the other Nordic countries,

where work with trigger tools has progressed much further, will ensure the transfer of valuable know how and experience.

It should be noted that the material of this report focuses only on trigger tools in the hospital environment. Therefore it is not possible to say whether the GTT or other trigger tool methodology is suitable for other levels of healthcare services – such as primary care centres, nursing homes, etc. Evidence on this subject exists, and its analysis can be the focus of future research. Reconstructing the patient safety picture from the viewpoint of a patient's trajectory through various healthcare service chains and organisations can open up new paths for progress in patient safety.

References

1. Pasternak A. Hoitovirheet ja hoidon aiheuttamat haitat. *Duodecim*. 2006;122:2459–70.
2. Hämäläinen P, Rintanen H. Potilasturvallisuus ja sen mittaaminen. (In Finnish). In: Pekurinen P, Räikkönen O, Leinonen T. (Editors). *Tilannekatsaus sosiaali- ja terveydenhuollon laatuun vuonna 2008*. *Stakesin Raportteja 38/2008*, p. 103-118. ISBN 978-951-33-2290-8 (PDF). Stakes, Helsinki, 2008
3. AHRQ Evidence Report/Technology Assessment No. 43. Making Health Care Safer: A Critical Analysis of Patient Safety Practices. AHRQ Publication 01-E058, July 2001. Available at: <http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hstat1.chapter.59276>
4. Battles JB, Lilford RJ. Organizing patient safety research to identify risks and hazards. *Qual Saf Health Care* 2003;12(Suppl II):ii2–7.
5. GTT Manual Finnish translation. Peltomaa K, Gröndahl W, Ekblad U, Kantola I, Savunen T. Information available (in Finnish) at: http://www.thl.fi/fi_FI/web/potilasturvallisuus-fi/global-trigger-tool
6. KanTa – Homepage at: <http://www.kanta.fi/en/web/ammattilaisille/potilastiedonarkiston-maarittelyt>
7. THL Patient Safety with Skills Programme (Potilasturvallisuustaidola ohjelma). Homepage: http://www.thl.fi/en_US/web/potilasturvallisuus-en

Annex 1. Literature Review Search Strategies

Search strategy 1: EPR + Triggers + Safety

1. Medical Records Systems, Computerized/
2. electronic patient records.mp.
3. exp Electronic Health Records/
4. electronic health records.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
5. electronic medical records.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
6. computerized patient records.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
7. computerized medical records.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
8. electronic patient data.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
9. or/1-8
10. trigger tool*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
11. global trigger*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
12. trigger system*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
13. alert* system*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
14. alert* tool*.mp.
15. exp Adverse Drug Reaction Reporting Systems/
16. computerized monitor*.mp.
17. computerized surveil*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
18. computerized screen*.mp.
19. or/10-18
20. Medication Errors/ or exp Safety Management/ or exp Medical Errors/ or exp Accident Prevention/ or patient safety.mp. or exp Safety/
21. 9 and 19 and 20

Search strategy 2: EPR + Triggers + reviews

1. Medical Records Systems, Computerized/
 2. electronic patient records.mp.
 3. exp Electronic Health Records/
 4. electronic health records.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
 5. electronic medical records.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
 6. computerized patient records.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
 7. computerized medical records.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
 8. electronic patient data.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
 9. or/1-8
 10. trigger tool*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
 11. global trigger*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
 12. trigger system*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
 13. alert* system*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
 14. alert* tool*.mp.
 15. exp Adverse Drug Reaction Reporting Systems/
 16. computerized monitor*.mp.
 17. computerized surveil*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
 18. computerized screen*.mp.
 19. or/10-18
 20. review*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]
 21. 9 and 19 and 20
- Total References: 61

The search strategy was subsequently modified by taking into accounting indexing particularities of the target resources and applied to EBM databases:

Search Strategy for EBM Databases:

Database: All EBM Reviews - Cochrane DSR, ACP Journal Club, DARE, CCTR, CMR, HTA, and NHSEED

-
- 1 Medical Records Systems, Computerized/ (216)
 - 2 electronic patient records.mp. (11)
 - 3 exp Electronic Health Records/ (0)
 - 4 electronic health records.mp. [mp=ti, ab, tx, kw, ct, ot, sh, hw] (19)
 - 5 electronic medical records.mp. [mp=ti, ab, tx, kw, ct, ot, sh, hw] (63)
 - 6 computerized patient records.mp. [mp=ti, ab, tx, kw, ct, ot, sh, hw] (3)
 - 7 computerized medical records.mp. [mp=ti, ab, tx, kw, ct, ot, sh, hw] (19)
 - 8 electronic patient data.mp. [mp=ti, ab, tx, kw, ct, ot, sh, hw] (1)
 - 9 or/1-8 (309)
 - 10 trigger tool*.mp. [mp=ti, ab, tx, kw, ct, ot, sh, hw] (2)
 - 11 global trigger*.mp. [mp=ti, ab, tx, kw, ct, ot, sh, hw] (0)
 - 12 trigger system*.mp. [mp=ti, ab, tx, kw, ct, ot, sh, hw] (5)
 - 13 alert* system*.mp. [mp=ti, ab, tx, kw, ct, ot, sh, hw] (14)
 - 14 exp Adverse Drug Reaction Reporting Systems/ (83)
 - 15 Medical Errors/pc (28)
 - 16 or/10-15 (132)
 - 17 9 and 16 (4)
 - 18 computerized monitor*.mp. (4)
 - 19 computerized surveil*.mp. [mp=ti, ab, tx, kw, ct, ot, sh, hw] (1)
 - 20 16 or 18 or 19 (137)
 - 21 9 and 20 (4)
 - 22 reminder system\$.ab,kw,ot,ti. (128)
 - 23 20 or 22 (264)
 - 24 9 and 23 (11)
 - 25 computerized reminder\$.ab,kw,ot,ti. (37)
 - 26 23 or 25 (281)
 - 27 9 and 26 (14)
 - 28 monitoring alert\$.ab,kw,ot,ti. (5)
 - 29 26 or 28 (284)
 - 30 9 and 29 (14)
 - 31 from 27 keep 1-14 (14)

Annex 2. Reviewed studies on computerized trigger systems - additional details

Types of triggers, resources required and main findings of reviewed studies.

Study	Triggers	Resources	Main findings
Classen 1991	Series of simple rules examining chemistry, hematology, drug level and drug order data	2 person- hours per day for summary reports of 24h reviewed by a pharmacist.	Period of 18 months, 36653 patients, 557860 drug exposures, 731 ADEs detected and verified in 648 patients. Only 9 reports through traditional voluntary reporting system. ADE rate of 1,67% , average length of stay 13 days vs. 5 for non-ADE patients.
Classen 1992	Both voluntary and automated, incl. sudden medication stop orders, antidote ordering, certain abnormal lab values, specific laboratory tests	16 alerts per day on average, average time for follow up of each - 10'. Intensive educational efforts for the nursing and medical staff, to increase awareness of the ADE's and availability of computerized voluntary reporting	999 verified ADEs in a period of 2 years. 13% classified as severe, 83% moderate and 4% mild. Several of the signals used to activate the knowledge base have been changed during this period. Knowledge was also added to the knowledge base to reduce false positives.
Greenway 1993	Medical alerting system alerts the Medical Director and Respiratory Care staff to potentially harmful events that, if untreated, may result in increased morbidity or mortality	Respiratory care computer system developed by a Medical Informatics graduate student (in use since 1983). Medical director reviews alert reports once daily. Quality assurance activities (such as follow up of oxygen therapy and various respiratory care indicators) performed routinely and much more widely in a fraction of the time required for manual chart review.	Crucial element in the process is the availability of an integrated patient database that allows quality assurance and alerting functions to be performed without labor-intensive chart audits.
Camacho 1998	(1) the Uniform Clinical Data Set System (UCDSS), a software-based system developed by the Health Care Financing Administration (HCFA) version of 1993, (2) the HCFA generic screens, and (3) the Harvard Medical Practice Study (HMPS) screening criteria.	The UCDSS was much more labor-intensive with an average of 6.7 hours of nurse review per quality problem disclosed versus 2 hours required by HCFA and HMPS.	Despite standardization of procedures and multiple reviews by cardiologists trained in structured review, chance-corrected agreement (Kappa) between physicians in selected items was low to moderate (0.11-0.58) in ranges similar to those reported by others. Inter-rater agreement limited the validity of physician judgment as a reference standard to assess the performance of screening systems.
Jha AK 1998	ADE detection rules consisting of Boolean combinations of simple medical conditions (e.g. new medication orders, laboratory results above or below certain thresholds, medication orders associated with changes in lab values over time) Computer based rule editor to create the screening rules - Starting point: the published rules from the LDS Hospital study; improved on rules with known low PPV, to reduce false alerts 52 unique rules in database; several modifications during study period	11 person-hours/week vs. 55 person-hours/week for review and 5 person-hours/week for voluntary reporting List of alerts generation: 1 person-hour per week Review of associated charts: 10 person-hours per week Entering the rules for the ADE monitor: 2 programmer weeks (approx. \$2000) Ongoing maintenance: 1-2 programmer hours per month Entire project: 2-person years (but ADE monitor only part of it)	Crude adjusted rate of 21 ADEs per 1000 patient days – 9,6 for computer monitor (substantial number of mild ones). Computer-based monitor identified fewer ADEs than chart review, but much more than voluntary reporting. Overlap between different methods was small. Computer-based method identified more severe ADEs (more lab results & medication orders based), chart review identified more preventable ones, identification depending more on symptoms. PPV of alerts: 16 - 23%

Study	Triggers	Resources	Main findings
Levy 1999	Signals consist of absolute values of, or changes in routine laboratory tests (blood count, liver and kidney function tests, serum electrolytes and glucose and drug plasma levels), which may indicate potential ADRs..	The online laboratory data for all admissions were screened daily for such ALS by a specially developed computer program. Patient files, including round notes and discharge summaries were reviewed for mention of ADRs. In addition, all charts were reviewed by a team of clinical pharmacologists for clinical evidence of potential ADRs, for their severity and the likelihood of them being ADRs and to see whether or not they were recognized as such during hospitalization by the staff physicians. For grading the probability of an event being an ADR the Naranjo algorithm score was used, while for verification of ADRs, the Iowa Drug Information Service, Micromedex, Medline 1966±1997 and own adverse-drug-reaction-oriented home database were used	Seventy-one ADRs were detected in 64 of the 199 (32%) admissions. 27% of the ADRs were serious, 9% of the admissions were due to ADRs. Two hundred and ninety-five ALS were generated involving 69% of the admissions. 61% of the ADRs were identified by ALS. ALS were present in 58% of the ADR negative admissions. 85% of the ADRs were recognized as such and 19% of the ALS-positive ADRs were not recognized by the staff physicians. In 19% of the admissions three or more ALS were found.
Lederer 2005	Lab-based patient specific INR triggers and pharmacy-based patient specific Vit. K triggers. Brief clinical course to describe the factors involved with the trigger, special emphasis on appropriate use and dosage of warfarin. Also, concurrent medications, diet, physical state and adequacy of lab monitoring and medication adjustments based on lab reports.	ADE Reduction Team: three vice presidents of medical affairs, four pharmacists, two pharmacy directors, three nursing leaders, two outpatient medical directors, senior VP of clinical improvement, medical director of clinical improvement, clinical improvement department support personnel. Initial individual training with the clinical pharmacists by the physician consultant to ensure accuracy of data extraction and ADE harm scoring. Medical director of clinical improvement read through all extracts - ongoing feedback to clinical pharmacists. Senior nursing leadership conducted nursing staff education - repeated at intervals to ensure coverage of new employees. Education and updates to staff on the whole. ADE Reduction Team members needed monthly data feeds to fine-tune the interventions. Administrative leadership - quarterly reports, organization on the whole updates on 4 to 6 months frequency.	Reductions in ADEs related to warfarin administration by 45% in inpatient and 52% in outpatient management. During data collection warfarin management and alert processes were continually improved to achieve better performance. Identification of outpatient events was notified to outpatient medical directors for chart review and further improvement cycle action and re-education. ADE rates standardized to 10,000 patient days, allowing large to small facility comparisons on event-per-patient days rather than raw numbers. Initiative laid the groundwork for real-time alert-based interventions - automation of 3-day INR check up order, alerting of pharmacists when INR >3 or increase >0.8. Project developed and endorsed on senior leadership level, which was responsible for the initiative's widespread diffusion.
Kilbridge 2006	Rules include orders for anti-dotes, toxic drug levels, combinations of laboratory values or trends and medication orders (eg. rapidly dropping platelet count in patient on heparin). Patient medication data extracted from different pharmacy systems each day, lab and demographic data are queried in the common clinical data repository	Lists of potential ADEs are reviewed by specially trained pharmacists - causality and severity. Review of patients record (80% resolved through EPR - 20% requires paper chart review); comparison of findings amongst themselves and then with a physician.	8-month study period (March-October 2005) University hospital: 25177 patients, 1116 ADEs in 900 patients – 4,4 ADEs per 100 admissions 3.6 times amount of ADEs identified compared to voluntary reporting Community hospital: 8029 patients 501 ADEs in 399 patients 6,2 ADEs per 100 admissions 12,3 times amount of ADEs identified compared to voluntary reporting Certain types of ADEs higher in community hospital (antibiotic-associated colitis, drug-induced hypoglycemia, anticoagulation-related ADEs)

Study	Triggers	Resources	Main findings
Szekendi, 2006	Records selected on the basis of 21 electronically obtained triggers (incl. abnormal lab values, high risk and antidote medications) from laboratory and pharmacy databases Rated for preventability (four-point scale, dichotomous for analysis) and severity (NCC MERP index) - categorized by event type based on local incident reporting system taxonomy.	Assignment for review by nurse case study manager. Review by patient safety team - registered nurse and pharmacist, plus two physicians for conflict resolution. Minimum time to complete an entire record review was 35 minutes for charts not needing physician review, and 45 minutes for those that did. Discussion with clinicians directly providing patient care (additional info or necessary changes to patient care)	3 month period, 327 records reviewed, at least one AE in 243 (17%). 163 in-patient preventable events and 138 that did not lead to harm. Interventions to prevent or ameliorate harm (according to record entry) in 47 patients Interventions most frequently related to incorrect prescribing of medications and to inconsistent or incorrect documentation of clinical information (particularly the latter likely to escape). Of the 163 preventable events, 101 directly related to triggers, 62 not-triggered. Non-triggered events were more likely to be severe (surgical or procedural complications) - Errors in diagnosis and clinical management may also go undetected by lab and pharmacy triggers alone
Seeger AC 2007	Rules: Derived from a previously developed set; combinations of drugs (based on pharmacy order and not medication administration), laboratory data and patient demographics that met certain conditions - linking. Designation of priority level: low-medium-high-critical. Possibility to adapt according to local practices. Activated if all criteria were met, and changed to alert if amount of time to react (Good Medical Practice interval - rule specific) was exceeded. Out of 787 rules in library, 358 were used bc of inability to reliably capture patient weight, and bedside glucose values not recorded electronically.	Average cost for preventable ADE \$US4685 (based on 1997 values - Bates study JAMA) - with 37 preventable ADE's per year: annual cost = \$US173 345. Plus 94 non-preventable ADE's detected. Personnel utilization: estimated that about 1,5 hours of pharmacist time per day needed to triage, follow up and report results. With a labour cost of \$US40 per hour the annual labour costs = \$US15600. Labour costs for IT implementation and maintenance = about \$US10000 annually. Set up, administration and operation of the system in the first year \$US99 000 and \$US56 350 for following years. Total first-year costs = \$US124 600, Subsequent year costs = \$US81 900, compared with annual benefits of \$US173 345 (NOTE: assuming that the level of ADE's remains the same over the years and that drastic content maintenance of rules database is not needed). Costs calculations are based only on high and critical alerts - not known how they would be affected when also less significant alerts would be taken into account.	In 6 months: over 90000 pharmacy medication orders and 841028 lab results, among 3428 patient admissions to medical or surgical services. 8829 activations of the rule set: 528 high or critical, 664 medium and 2355 low priority. Chart review of 56 high or critical charts found 5 non-preventable and two preventable ADEs - extrapolation: about 94 non-preventable and 37 preventable ADEs annually that could be detected through this method.
Kilbridge 2008	Combinations of demographic, laboratory and pharmacy data from SLCH clinical systems.	Potential events evaluated by two physicians and a pharmacist for causality by using the Naranjo algorithm, and for severity using NCC MERP. System surveys all SLCH inpatients on a daily basis.	922 admissions, 95 unique alerts generated in 57 patients with 59 visits. Most common alerts: impairment of renal function (28), hypokalemia (23), seizures (11), hepatotoxicity (7), hypomagnesaemia (6). 12 alerts had score of 5 or greater (probable or certain causality) and MERP scores of E or greater (temporary harm), one permanent harm (G). 8 of the 12 ADE's represented drug-inducing hypokalemia (supplementation). One ADE for each: hepatotoxicity, naloxone, C difficile toxin, hypomagnesaemia. PPV of the whole set (13%)

Study	Triggers	Resources	Main findings
Ferranti 2008	The system evaluates medication, laboratory, and patient demographic information against a set of 14 clinical rules (triggers) that span 3 main categories: abnormal laboratory results, use of antidotes, and drug-lab combinations. Although nearly 130 rules have been deployed since the system's inception, in operationalizing surveillance only high-risk rules with high true positive rates are considered (Table 1).	The daily trigger list was evaluated by 3 clinical pharmacists who performed chart review to determine whether an ADE occurred. Pharmacists identified all possible medications involved in the event and assigned a causality score using the Naranjo algorithm ¹⁸ and a severity score using the DUH 7-point scale. All events scored with causality Q5 and a severity Q3 were considered ADEs. Pair wise inter-rater reliability scores (k statistic) exceeded 0.88 for each rater pair. Surveillance chart review requires clinical resources, hence significant institutional support is needed to implement and sustain such a program. Focus should be on specific high-volume, high-risk drugs that fit the hospital's profile and clinical service range. Can also be utilized for longitudinal trending of ADE rates	Surveillance detected 710 ADE's out of 7365 triggers (6,93/1000 patient days) - voluntary reporting 205 from 1693 reports. Out of 875 unique ADE's only 40 (5,6%) identified by both systems. Most surveillance events related to hypoglycemia, while reported events belonged to the miscellaneous category
Herasevich 2009	Acute Lung Injury sniffer was triggered when both criteria below were met within a single 24-h period: (1) Qualifying arterial blood gas analysis: the ratio of partial pressure of oxygen over inspired oxygen concentration (PaO ₂ /FIO ₂) ³⁰⁰ and (2) Qualifying chest radiograph report: free text Boolean query containing trigger words: ("bilateral" AND "infiltrate") OR "edema."	The radiologist is available 24 h a day, and the reports of portable chest radiographs of ICU patients are available in electronic form on an average of 2 h after the chest X-ray is obtained. Daily e-mail messages alerted study investigators about potential ALI cases. Two intensivist researchers (MY and HK) blinded to clinicians' diagnoses reviewed the electronic medical records within 48 h of the alert and assigned the diagnosis of ALI according to the American-European Consensus Conference criteria (AECC)	ALI developed in 325 patients and was recognized by bedside clinicians in only 86 (26.5%). ALI "sniffer" demonstrated excellent sensitivity of 96% (95% CI 94–98) and moderate specificity of 89% (95% CI 88–90) with a positive predictive value ranging from 24% (95% CI 13–40) in the heart–lung transplant ICU to 64% (95% CI 55–71) in the medical ICU.
Kilbridge 2009	Combinations of demographic, encounter, laboratory and pharmacy data from SLCH clinical systems. Rule set expanded for ADEs more common in children, eg, medication induced seizures, electrolyte abnormalities, hypoglycemia, medication induced GI dysfunction. Total of 32 rules (criteria) Most common true positive alerts were: hypokalemia (66), hypomagnesemia (19), nephrotoxicity (18) and naloxone administration (9). Most frequently implicated medications were diuretics, antibiotics, immunosuppressants, narcotics and anticonvulsants.	Two independent assessments by pharmacists, approximately three times a week, and a final assessment by third reviewer - physician expert, whose evaluations were viewed as the gold standard. Material: alert info, plus critical patient data (incl. current medication lists, relevant lab values, patient weight and demographic data). Pharmacists had access to other online systems, incl. Hospital pharmacy system and enterprise clinical data repository to assist in evaluation. Events with Naranjo (probability or causality) score of 5 or higher were then scored for severity using the NCC-MERP scoring system. Also recorded: responsible medications, narrative of the event. Naranjo 5 or higher and MERP E or higher: considered ADEs. Pharmacists spent an average of 7 hours per week each evaluating the alerts.	6,889 non oncology patient admissions, generating 40,250 patient days. 1226 alerts generated, 160 true ADEs detected, representing 4 ADEs per 1000 patient days or 2,3 ADEs per 100 admissions. 135 events caused temporary harm to the patient (E score), 20 required prolonged hospitalization (F), 4 suffered permanent harm (G) and 1 died of multisystem disease complicated by drug induced nephrotoxicity from gentamicin and vancomycin. Average patient age suffering ADE was 6,3 years, while average of all admitted non-oncology patients was 6,8. Greatest numbers of ADEs occurred in the critical care units with 56 (35%) in cardiac ICU, 43 (27%) in general pediatric ICU and 12 (7,5%) in NICU. Composite PPV was 13% (total # of ADEs/total # of alerts), ranging from 100% to 0. Rates consistent with chart based studies in pediatric population, but 50% than those detected by limited rule set.

Study	Triggers	Resources	Main findings
Ramirez 2010	Signals from LIS: grouped under 6 specific conditions: Agranulocytosis, Pancytopenia, Thrombocytopenia, Liver Injury, Hyponatremia, Rhabdomyolysis ADRs caused by accidental or intentional overdose excluded, medical errors of prescription, dispensation or administration excluded, but included if error in decision making (contraindicated medication or drug interactions). ADRs by chemo drugs excluded (expected side effects of treatment - explained by pharmacodynamic properties).	First phase: on file laboratory data of all admissions screened 7 days a week, 24h a day for predefined ALSs. Second phase: after ALS identification team of clinical pharmacologists identified the patients as to avoid duplication and reviewed their EMRs. Non ADRs excluded. Third phase: individual review of remaining patient records, patient visits and/or interview with attending physician. After confirmed ADR - extensive data collection and reporting - complete adverse event report sent to SPhSystem. Generation of ALS automatic (resources needed for database program not mentioned). Review of EMRs (resources not stated) + individual assessment and medical record review for almost 22% of ALS with plausible ADR. In order to identify a single SADR EMRs of about 16 patients had to be reviewed and 5 patients visited. Staff needed to start and maintain the program during its first year can be calculated as one person/year, with significant part of the work directed towards program set up. Hiring of one person to generate just two identified SADRs per week does not appear to be justified. However, the program introduces many advantages in the hospital and can be easily expanded to other ALSs with the addition of fewer resources than those needed to initiate it. Alternative systems for pharmacovigilance: 1. Spanish Pharmacovigilance System database and 2. minimum basic data set - "adverse event in therapeutic use" ICD9-CM as principal diagnosis. No ADRs identified from ALSs were directly communicated by healthcare professionals to System 1, PPLSH identified all ADRs identified through ICD9-CM codes, plus 107 others - Detection of ADRs not detected through other means. Additional advantage: tool for identifying patients experiencing ADRs and subsequently collecting detailed information necessary for studying the risk factors associated with SADRs (on the genetic level as well).	54,525 hospital admissions. 839,668 routine laboratory tests (blood count, liver function, serum electrolytes) on inpatient and emergency patients. 7,198 ALS, corresponding to 1,732 inpatients with at least one ALS each. Review of EMRs showed that in 1,212 ALSs patients (70%) a cause other than an SADR was present. Therefore, it remained plausible that in 520 patients (30%) an SADR had occurred. After personal visit and review of their medical records, 413 of those patients were excluded from the ADR category and 110 suspected ADRs were found in 107 patients. All were confirmed as SADRs according to the norms of the Spanish Pharmacovigilance system as well. More than 50% of SADRs in adults (45) concerned patients >60y. and in children (11) patients <3y. All SADRs classified as serious because they required hospitalization (77 cases) or prolonged hospital stay (33 cases). 31 were life threatening. The Incidence of drug-induced reactions detected by PPLSH was 20,17% per 10,000 inpatients-year. Mean length of stay with an ADR was 12 days (4,4 more than without an ADR) and the mean excess stay due to ADR was 3,5 days. 77 cases of SADRs had developed in the community and led to hospitalization, while 33 occurred in the hospital. PPLSH produced an increase in number of SADR notifications not included in the program from 6 in the previous year to 53 during the year of the study.

Annex 3. Triggers and Minimum Data Set (Ydintiedot) for Finnish EPR systems

The feasibility of implementing computerized trigger tools through the minimum (core) data set defined for structured electronic patient record (EPR) systems in Finland was also explored in the context of the pilot project in the field of neurosurgery, a novel clinical domain for trigger utilization.

Minimum (or core) data elements include identification data, clinical data and other data (e.g. patient consent, treatment plan – See Figure 1 and Table 1 for further details). The set has been piloted in several organisations and its implementation in various EPR systems is presently on-going.

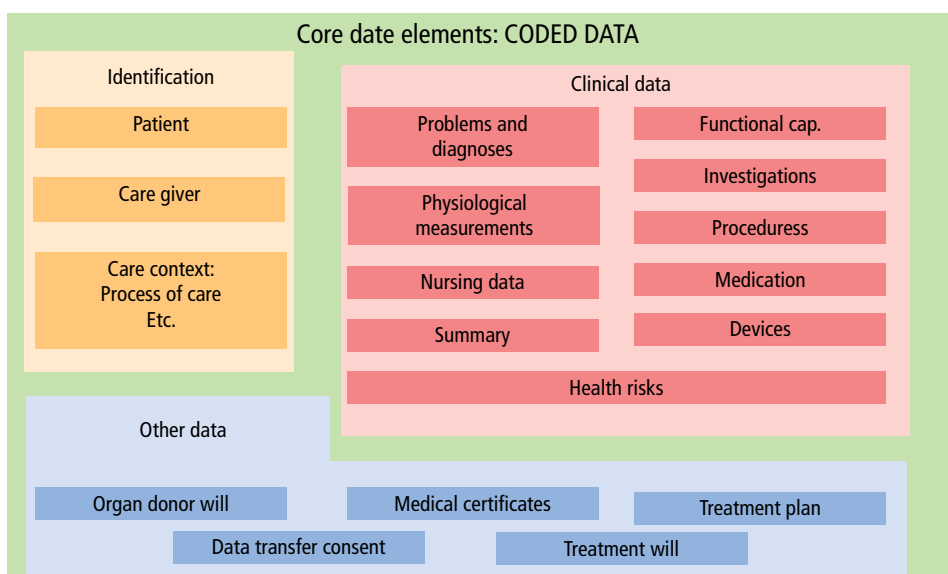


Figure 1. Components of the Finnish core data set (nationally defined since 2004, updates & further specifications ongoing) – Graph from presentation given by P.Hämäläinen and J. Komulainen

Table 1. Structured data entry levels of the core data set

Minimum data set domain	Structured and coded data entry or, in the absence of codes, structuring on heading level
Problems and diagnosis: risk factors, diagnosis and cause of treatment	Coded
Factors affecting health	Status and amount of smoking, mapping of alcohol use and information on physical exercise in structured format. (CDA definition missing)
Physiological measurements	Structured and coded
Nursing	Structured and coded
Investigations	Structured and coded
Procedures	Structured and coded
Medication treatment	Structured and coded
Prevention: vaccinations	Structured and coded (CDA definition is missing)
Statements	General CDA R2 certificate/statement/form structure, use of classifications in the knowledge fields
Ability to function	Structured at heading level
Assistive devices	Structured and coded
Summary of episode of care	Structured at heading level (The summary may also be structured and coded. In the case of utilizing earlier entries in the patient record their structuring needs to be maintained).
Data concerning the planning of further care	Structured and coded
Consent – patient consent regarding the release of data in the context of a specific episode of care – patient consent regarding release of data on overall healthcare services provided	General CDA R2 certificate/statement/form structure, use of classifications in the knowledge fields

The list of thirteen triggers developed by the Neurosurgery- Neurology department at the University Hospital of Tampere was juxtaposed to the national core data set, and to the respective GTT set. The purpose was to assess the possibility of trigger representation through data available in Finnish EPR systems in coded form, hence enabling their automated detection.

The cross-tabulation results show that the necessary data are to be found in the following core element categories: problems and diagnoses, physiological measurements, investigations, procedures, medications, laboratory results and nursing care (Tables 2 and 3).

Table 2. Neurosurgery pilot triggers vs. Core Data Set elements

Clinical deterioration (neuroworsening)	Problems and diagnoses; Physiological measurements
Radiological procedure performed because of changes in clinical status	Investigations
Procedure undertaken because of changes in clinical status	Procedures
Complication caused by treatment procedure	Problems and diagnosis; Procedures
Wound problem	Problems and diagnosis; nursing
Disruption in the circulation of cerebrospinal fluid	Problems and diagnosis; Investigations
Infections Identified through the use of antibiotics and/ or lab values (e.g. high CRP)	Medication treatment; Investigations.
'Rebound' - patient returns to the ICU	Care unit
Prolonged ICU stay (for non-neurological reasons) E.g. pulmonary, renal, hepatic etc complications, in a patient who is considered "stable" in terms of brain status	Care unit, length of stay
Re-intubation	Procedures
Electrolyte disturbance requiring treatment Usually hyponatremia	Laboratory
Fluid disturbance requiring treatment [Hypo- hypervolaemia , hormonal disturbances (SIADH, diabetes insipidus)]	Laboratory, physiological measurements
Blood pressure disturbance requiring treatment (Low blood pressure, hypertension that requires medication)	Physiological measurements, medication treatment

The respective coding and classification systems indicated for representing these elements are: ICD-10, FinLOINC, FinCC (for nursing care), and national classifications for laboratory and radiological investigations.

Table 3. Sample of GTT triggers vs. Finnish minimum data set elements

GTT TRIGGERS (surgery, ICU, care modules)	CORE DATA ELEMENTS
Intra- or Post-Operative Death	Problems and diagnoses
X-Ray Intra-Operatively or in Post Anesthesia Care Unit	Investigations
Change in Procedure	Procedures
Occurrence of Any Post-Operative Complication	Problems and diagnoses; Procedures
Return to surgery	Problems and diagnoses; Nursing care
Post-Operative Increase in Troponin Levels Greater than 1.5 Nanogram/ml (possible cardiac event)	Problems and diagnoses; Investigations
Clostridium difficile Positive Stool	Investigations
Readmission to the Intensive Care Unit	Care unit
Admission to Intensive Care Post-Operatively	Care unit, length of stay
Intubation or Rentubation or Use of BiPap in Post Anesthesia Care Unit (PACU)	Procedures
Glucose Less than 50 mg/dl	Laboratory
Rising BUN or Serum Creatinine Two Times (2x) over Baseline	Laboratory, physiological measurements
Over-Sedation/Hypotension (Care)	Physiological measurements, medication treatment

Following this initial analysis, further exploration is required with regard to: representation of the reason(s) leading to certain procedures and/or investigations, the impact of the time axis or phase in an episode of care on the interpretation of a trigger, the consequences of multiple coding alternatives for the same term (depending on who is the user entering data), and the possibility for (coded) data storage in various parts of the record vs. quality and efficiency of information retrieval.

This study was based on the generic guidance documentation provided for the utilization of the core data set. Hence, the impact of the choices made by each EPR system vendor when implementing the core data set within a certain system is still unknown. Moreover, since utilization of the core data in any 'added-value' application is dependent on an agreement between the client(s) and the respective EPR system vendor, the cost of implementing automated trigger tools may at present be deterring for healthcare provider organisations.

Annex 4. IHI GTT Worksheet used in the Southwest Hospital District of Finland patient safety programme

	HOITOMODUULI	+	Tapahtuman kuvaus sekä haitan asteen arviointi	LÄÄKEHOITO	+	Tapahtuman kuvaus sekä haitan asteen arviointi
H1	Verituoitteet			L1		
H2	Elvytys			L2	Clostridium diff.	
H3	Dialyysi			L3	P-APTT >100 s.	
H4	Positiivinen verivijely			L4	INR > 6	
H5	RTG tai UÄ (trombi)			L5	P- gluk < 3	
H6	Hkr tai Hb lasku 25%			L6	S-Urea /S-Krea	
H7	Kaatuminen, putoaminen			L7	K-vitamiini	
H8	Painehaavauma			L8	Antihistamiini	
H9	Paluu 30 pv. sisällä			L9	Lanexat	
H10	Leposteiden käyttö			L10	Narcanti	
H11	Sairaala-infektio			L11	Pahoinvointilääkkeet	
H12	Aivohalvaus			L12	Ylisedaatio/hypotensio	
H13	Siirto tehost. hoitoon			L13	Äkil. lääkityksen lopetus	
H14	Toimenpide				Muu	
H15	Muu				TEHOHOITO	
	KIRURGINEN			T1	Pneumonia	
K1	Reoperaatio			T2	Palaaminen teho-osastolle	
K2	Toimenpiteen muutos			T3	Toimenpide teholla	
K3	Postop. tehohoito			T4	Intub. / reintubaatio	
K4	Intub/ reintub/ BiPap				PERINATAALI	
K5	Intra- / postop. RTG			P1	3. tai 4. asteen repeämät	
K6	Intra- /postop. kuolema			P2	Verihutalearvo < 50 000	
K7	Respiraattorihoito yli 24h			P3	Vuoto yli 500ml /1500ml	
K8	Intraoperat. adren/noradr.			P4	Postpartum lääkkeet	
K9	Postop. troponiini yli 1,5			P5	Imukuppi- /pitiisyntyty	
K10	Elinvaurio, poisto, korj.			P6	Yleisanestesia	
K11	Postop. komplikaatio				PÄIYYSTYS	
				E1	Paluu 48 tunnin sisällä	
				E2	Hoidon kesto yli 6 h	