Department of Clinical Sciences,
Danderyd Hospital, Karolinska Institutet,
Stockholm, Sweden

# Evaluation of retrospective patient record review as a method to identify patient safety and quality information in orthopaedic care

Maria Unbeck


Karolinska Institutet

Stockholm 2012

Front page illustration by Kamilla Andersson. www.kamillaandersson.se.
Layout Ringvor Hägglöf

ISBN 978-91-7457-708-2

Printed by


www.reproprint.se
Gårdsvägen 4, 169 70 Solna

To Matts, Christoffer, Malin and Rasmus

# Abstract

The great benefits of modern healthcare must be weighed against the risk of patient injury due to human intervention. Studies show that adverse events (AE) are identified in up to 16.6% of all hospitalisations. As a step toward preventing AEs, efforts are made to collect patient safety information at different levels in the healthcare systems. The information is neither effectively organised nor integrated within the healthcare systems, leading to difficulty achieving systematic analysis. This may be due to the use of different methods that yield qualitatively different information about AE.

The general aim of this thesis was to evaluate the capability of retrospective record review (RRR) methods to identify patient safety and quality information in orthopaedic care.

In papers I and II, 395 patient records were retrospectively examined for AEs using both traditional incident reporting methods and RRR for the same cohort. More AEs were identified using RRR than by using traditional incident reporting methods. Also, paper II showed that more AEs were due to deficiencies in care processes rather than to deficiencies in technical skills.

In paper III, the efficiency of an orthopaedic nursing improvement initiative, called "improvement theme months," was evaluated using case study methodology and a RRR of 2,281 patients. Results showed significant improvement over time in performance of risk assessment for pressure ulcers and lowered pressure ulcer prevalence. We found RRR easy to use and valuable as a method to assess improvement over time.

In paper IV, the RRR methods, Harvard Medical Practice Study (HMPS) and Global Trigger Tool (GTT) were compared for capability to identify AEs in a sample of 350 randomly selected orthopaedic admissions. Results showed that HMPS identified more AEs than GTT did. The overall positive predictive value (PPV) was 40% and 30% for HMPS and GTT methods, respectively.

Retrospective record review appears to achieve wider coverage when identifying orthopaedic AEs at a local level. Given that many current methods vary considerably in quality of data gathered and in coverage, which require multiple methods to be used concurrently, the wider coverage characteristic of RRR is an advantage. Consequently, RRR could play a vital role in quality and safety information systems in order to identify, categorise, and analyse quality and patient safety problems and to provide the basis for interventions. Increased awareness, consideration of risk factors, interventions focused on multidisciplinary and interdepartmental teamwork, and strategies that focus on healthcare processes may reduce the frequency of AEs in orthopaedic care. Also, RRR can incorporate a time series display of patient safetyintervention outcomes to drive change.

As a method, improvement theme months may serve to organise quality and lead to safety improvement in nursing. However, we found that it was associated with a lengthy period of time before new guidelines, quality indicators, and safety initiatives were noticed and became widely used in clinical practice. To achieve sustainable and significant improvement, interventions on many levels of the organisation were needed.

# List of publications

This thesis is based on the following papers, which will be referred to in the text by their roman numerals:

**I.** **Unbeck M**, Muren O, Lillkrona U.
Identification of adverse events at an orthopedics department in Sweden.
*Acta Orthopaedica*. 2008;79(3): 396-403.

**II.** **Unbeck M**, Dalen N, Muren O, Lillkrona U, Pukk Härenstam K.
Healthcare processes must be improved to reduce the occurrence of orthopaedic adverse events.
*Scandinavian Journal of Caring Sciences.* 2010;24(4):671-7.

**III.** **Unbeck M**, Sterner E, Mattias Elg, Fossum B, Thor J, Pukk Härenstam K.
The impact on evidence-based practice and patient safety of quality improvement theme months in orthopaedic nursing: a case study on pressure ulcer prevention.
*Submitted*.

**IV.** **Unbeck M**, Henriksson P, Jürgensen U, Muren O, Nilsson L, Schildmeijer K, Pukk Härenstam K.
Detection of adverse events by two methods for retrospective record review. A comparison of the Harvard Medical Practice Study method and the Global Trigger Tool.
*Manuscript.*

# CONTENTS

# List of abbreviations

| | |
|---|---|
| AE | Adverse event |
| ASA | American Society of Anaesthesiologists |
| CMIFS | California Medical Insurance Feasibility Study |
| GTT | Global Trigger Tool |
| HFMEA | Healthcare Failure Mode Effect Analysis |
| HMPS | Harvard Medical Practice Study |
| IHI | Institute for Healthcare Improvement |
| LOS | Length of stay |
| MNS | Modified Norton Scale |
| PCE | Potentially compensable event |
| PPM | Point prevalence measurement |
| PPV | Positive predictive value |
| QAHCS | Quality in Australian Healthcare Study |
| QI | Quality improvement |
| RCA | Root Cause Analysis |
| RF | Review form |
| RN | Registered nurse |
| RRR | Retrospective Record Review |
| UTCOS | Utah and Colorado Study |
| WHO | World Health Organization |

# BACKGROUND

"The problem of harm arising from healthcare management is not new nor is the interest in this. What has changed in the past 10-20 years is the general acceptance of the need for measurement and assessment of scale and causes of harm followed by interventions to improve the safety of healthcare" p. 28[1].

A pioneer to measure patient harm in surgery was Ernest Codman, a Boston surgeon in the early 20th century. He took a major step by starting to systematically detect and understand medical errors. He was dissatisfied with the observed quality of care and apparent unwillingness in healthcare to evaluate surgical outcomes. He proposed that hospitals should follow patients long enough after hospitalisation to determine whether treatment was effective or not. If an unsuccessful treatment was found, the hospital should attempt to determine why, in order to learn from these failures and prevent recurrence. During the years, 1911-1916, he noted 123 errors from 337 patients and analysed the causes in his own hospital, which he creatively named, "End Result Hospital". He openly published these results in annuals reports and distributed this information to major hospitals in order to challenge colleagues to, in turn, demonstrate their outcomes. Codman is credited for holding the first morbidity and mortality conferences for internal audit in hospitals. He assisted in the founding of the American College of Surgeons, and later, in collaboration with other organisations, created Joint Commission on Accreditation of Hospitals, the largest accrediting body in the US[2 3].

While healthcare-associated harm, or iatrogenic injury, had been noticed earlier, this area had seldom been systematically investigated using either retrospective or prospective approach. In the sixties, investigative studies were carried out[4 5]. In the early 1970's, the Californian Medical Insurance Feasibility Study (CMIFS) initiated the first major investigative use of RRR, a structured review approach[6]. Steel at al.[7] followed with a study using prospective structured record review that also included interviews with personnel and other information sources available at the hospital. Couch et al.[8] used a prospective study design to examine errors in surgery care.

The Harvard Medical Practice Study (HMPS) using retrospective record review found that AEs occur in 3.7% of the hospitalisations[9]. With the release of the Institute of Medicine's report, *To Err is Human*[10], such findings led to massive debate. The report articulated recommended actions to improve patient safety at all levels of the healthcare system. According to Wachter[11], the modern patient safety movement began with this report. Since then, many reports and official statements from that issue have been released.

*To Err is Human*[10] led many countries to conduct AE studies, and measure their AE rates. The reported outcomes highlight the need to improve patient safety. In the late 2007, the Swedish National Board of Health and Welfare performed a study based on the HMPS protocol[12 13]. A representative sample, consisting of 1967 admissions, was reviewed. In total, 12.3% of the 1967 admissions contained AEs, of which 70% were judged to be preventable.

The most common types of AEs were healthcare-associated infection and organ injury. Of the preventable AEs, 62% were related to surgical specialities. Fifteen percent of AEs identified in these study reports contributed to permanent disability or deaths.

The healthcare burden of AEs is considerable[14-18]. The healthcare burden of AEs includes both individually affected patients, as well as the whole healthcare system[19]. Reported in *To Err is Human* is that between 44,000 and 98,000 people die in US hospitals every year where the AE was a contributing cause of the death[10]. Extrapolating to the 1.2 million annual admissions during the sampled year, Swedish study results correspond to 105,000 preventable AEs and are calculated to have resulted in approximately 630,000 extra days of hospitalisation, contributing to 3,000 deaths and 50,000 unnecessary outpatient visits[13].

# Terms and definitions

In this section I will present and discuss different terms and definitions in the patient safety field.

Standardised definitions are adopted within the patient safety field, but standard usage is inconsistent[20-22]. The conceptual meaning of some terms is, at times, unclear, which confuses rather than clarifies[11]. Terminology has also been taken from different perspectives and is aimed toward various purposes. In some studies, consistent terms and definitions have been used but careful description of how the terms had been applied, for example, by reviewers, was lacking. Across studies, on the other hand, meanings behind terms were different. Consequently, one's ability to evaluate outcome within the same used methodology and between studies with different methodologies may be limited[23].

In 2009, the World Health Organization (WHO) presented a report on international classification for patient safety. The aim was to develop an international taxonomy in this field in order to promote standardisation of terminology, definitions and classifications. It also includes a glossary of patient safety terms taken from various references[20]. Patient safety is set as a subset of quality[11]. See a list of commonly used patient safety terms preferred by WHO[20] in Table 1.

To classify errors, a meaningful approach is to consider what the reasonable intention was in the first place[24]. According to Reason[24], errors can be classified as unintended slips and lapses. These occur when the plan is adequate but actions fail to go as planned. Slips are errors that typically occur at the task execution stage, in interrupted daily routine, which are observable. Lapses involve, generally, failure of memory. Reason notes, further, that errors that occur when actions go as planned, but the intended plan in not the correct, would be termed "mistakes".

Reason also uses the terms "active failures" and "latent conditions" when describing different types of errors. Active failures often occur in the "sharp end" of the system involving frontline personnel. These failures can, and often do, have immediate outcomes. Latent conditions related to the "blunt end" are present in all systems. Latent conditions arise from strategic and other top-level decisions, and tend to be removed from the direct control of the personnel. These decisions can, years later, in combination with local triggers, such as a high workload, create an adverse outcome[24].

Errors can result in near misses, no-harm incidents or AEs, but these are not all preceded by an error. Error can be found in the active delivery of care, as in, doing something wrong (acts of commission) or failure to do the right thing (acts of omission)[20 25-27]. Errors are always found retrospectively[18].

**Table 1.** Patient safety terms using WHO definitions[20]

| Terms | Definitions |
|---|---|
| Patient safety | The reduction of risk of unnecessary harm associated with healthcare to an acceptable minimum |
| Risk | The probability that an incident will occur |
| Event | Something that happens to or involves a patient |
| Near miss | An incident that did not reach the patient |
| No harm incident | An event reached a patient but no discernable harm resulted |
| Incident | An event or circumstance that could have resulted in, or did result, in unnecessary harm to a patient |
| Error | A failure to carry out a planned action as intended or application of an incorrect plan |
| Harmful incident (adverse event) | An incident that results in harm to a patient |
| Harm | Impairment of structure or function of the body and/ or any deleterious effect arising there from, including disease, injury, suffering, disability and death, and may be physical, social or physiological |
| Healthcare associated harm | Harm arising from or associated with plans and action taken during the provision of healthcare, rather than an underlying disease or injury |
| Preventable | Being accepted by the community as avoidable in the particular set of circumstances |
| Contributing factor | A circumstance, action or influence that is thought to have played a part in the origin or development, or to increase the risk, of an incident |

There are many definitions of AE[20]. The most common definition (or similar) in epidemiological retrospective record review studies has been that an AE is "an unintended injury or complication which results in disability, death or prolonged hospital stay and is caused by health care management rather than the patient's disease"[26]. It is important to distinguish between adverse outcome as a result of healthcare management from those arising from morbidity and mortality due to the patients underlying disease or condition[11]. AEs include complications, which are deemed as leading to harm but of low preventability[9 18].

An AE can be classified as preventable or non-preventable. Some preventable AEs can be related to negligence, such as, "care that fall below the standard expected of the average physician, other provider, or institution"[9 28 29]; or to violation, including, "a deliberate deviation from operation procedure, standard or rule"[20]. A penicillin reaction in a patient with known allergy is considered a preventable AE if the physician does not ask the patient

about allergy before prescription or does not look for or react to risk of hypersensitivity information in the patient record. On the other hand, if the patient takes an antibiotic for the first time, or has taken it before without any problem, it is considered non-preventable AE[28]. Judging what is preventable is often not clear-cut, and it may be difficult to distinguish between a preventable and non-preventable AE[11].

Kjellén[30] defines a safety information system as a "system that provides the information needed for decisions and signalling related to safety".

The patient record in which all documentation is made by healthcare personnel is referred to as the record in this thesis.

## VALIDITY AND RELIABILITY

Validity and reliability are important in the development of any new method. The validity and reliability also ought to be tested and evaluated when accepted methods are introduced in new contexts. The understanding of these terms is of value when evaluating different studies and data collecting methods used[1].

Accurate is a synonym for valid[31]. A method is valid if it measures what it is intended to measure[1]. High validity implies high reliability. Validity in terms of the extent to which the method approximates the true value of identified incidents is difficult to address as there is no "gold standard" to measure against[23 32]. Screening tools are thus not expected to be completely accurate. There may always be false positive and false negative outcomes[23].

Precise is a synonym for reliable[31]. The reliability of a method is its ability to consistently detect the same event in the same set of information[1]. High reliability does not guarantee high validity.

# METHODS TO IDENTIFY PATIENT SAFETY AND QUALITY INFORMATION IN HEALTHCARE

The focus of this thesis was to evaluate the capability of retrospective record review methods to identify patient safety and quality information in orthopaedic care.

Donabedian has developed a model for measuring quality that is widely adopted in healthcare. This model also provides a framework for measuring safety within the healthcare system. The model consists of measures of structure (how care is organised, resources, equipment), process (what we do) and outcome (the effect of what we do). To be able to measure structure, process and outcome indicators different data collecting methods must be used[33]. These data collecting methods can be used to both identify safety and quality issues. The main focus in this thesis, however, will be on patient safety information.

Quality, risks and incidents in the healthcare system can be identified and measured by using different methods. Some methods are used routinely in the clinic and some are more used in research[1]. Some methods are focused towards the incident rates, while others address risks, contributing factors and causes[18]. Collecting data is the first step of an organisational learning process; and data collection is necessary to assess risks and incident rates, and to provide the basis to prioritise where to deploy resources and how to implement change, as well as to later monitor progress in patient safety outcomes[34].

To measure patient safety and quality deficiencies, and to follow up on improvement initiatives is difficult. Much of our learning has come from individual clinical cases and not from aggregated data, the latter which may be needed to prioritise and evaluate safety efforts[35]. Healthcare systems in some cases do not have a quantitative safety information problem but, rather, the problem is of qualitative nature[34 35].

When different methods are used to collect information, the outcome is not easily comparable. Often, patient safety and quality information is collected at different levels in the healthcare system by different organisations. The resulting information can be difficult to organise and may not be collected in a manner that it can easily be integrated. Systematic analysis is difficult since different methods yield different information about incidents[18]. The measures are derived from different purposes, data collecting methods and dimension of care processes in patient care[36]. Some of the information is not easily accessible to healthcare leaders and clinicians, and the identification and dissemination of information is often delayed, which complicates matters. The knowledge of reliability and validity of different methods is often limited in relation to specific contexts[23]. In Sweden, we have few valid and reliable methods for identifying risks and incidents and measuring patient safety[37-40]. Källberg et al.[39] compared four incident reporting systems in Sweden and found several problems, including: differences in use of terms, variation of information details, problems with precision and presentation of data, lack of standardised system for classification and categorisations, and too many, complex and different reporting systems in use.

Much effort has been invested in methods to identify risk and incidents. Measuring risks and incidents is more problematic than measuring other outcomes or processes within the healthcare sector because these must be understood in the context of the system within which they occur. To choose the most appropriate method(s) to meet the measurable goals that are set, leaders, clinicians and researchers must understand the strengths and limitations of different methods[31].There has been much debate regarding which method is the most valid, reliable and feasible[1].

A wide range of data collecting methods can be used in patient safety and quality work, some of which are discussed below. Those used in this thesis will be presented first, in more detail. Some other methods available to identify patient safety and quality information will be briefly described in the end of this section.

## STRUCTURED RECORD REVIEW

Structured review of records is the most extensively used and studied method to identify AEs[41 42]. These may be conducted either prospectively or retrospectively. In this thesis the focus will be on the most used form, RRR.

Two separate approaches are used for RRR: by the use of screening criteria, and by the use of triggers. Both have traditionally focused on harm, while measuring AE rate retrospectively.

The screening criteria, or triggers, that are used in the first stage of the review process can be explicit or implicit. The first stage is often carried out by trained nurses. An explicit screening criterion/trigger is precise and does not require specific assessment (e.g., death or readmission within 30 days). An implicit screening criterion or trigger requires more assessment and judgement, and is, perhaps, more difficult to apply to the specific assessment requirements of the user. For example, to determine "inappropriate discharge to home" or a "procedure complication", additional assessment is necessary. In practice, explicit and implicit screening criteria respectively triggers are generally combined[23].

The second review stage, performed by physicians to judge what was found in stage 1, can, in general, also be either explicit or implicit. Here, explicit methods involve more objective application of some pre-determined criteria, events or specific types of AEs (e.g. wrong site surgery). The reviewer determines presence or absence of relevant criteria[23 42 43]. This method may be less dependent on expert judgement and may be more reliable; but, it may detect a less range of AEs due to narrow inclusion[43].

Reviewers make an expert judgement of the presence or absence of AEs when using implicit review. The review process is not constrained by pre-determined criteria. A series of structured questions guide the judgement as to the nature of the AE[1 23]. This method is less standardised, which may reduce reliability and require reviewer training, but its strength is that it may identify events and processes of diverse nature that are not generally identified by explicit methods[43]. Most epidemiological RRR AE studies have used this method[9 13 26 27 44-47].

### Retrospective record review using screening criteria

Epidemiological RRR studies that employ screening criteria have been reported for a number of countries, on regional or/and national levels. These studies report AE rates and evaluate how effectively RRR can be used to identify healthcare-associated harm to patients. Near miss or no-harm incidents have generally not been included in these studies[6 9 13 26 27 44 45]. This method is mostly used in research and, conversely, not used as a local routine data collection method by personnel or for evaluating initiated patient safety interventions[1].

*The California Medical Insurance Feasibility Study*

The first, large-scale study of healthcare-related malpractice or harm was the CMIFS. Its two-stage RRR was performed by using 20 screening criteria to screen 20,864 records of patients discharged during 1974 from 23 randomly selected acute care California hospitals[6].

This study was performed to evaluate if a system of no-fault compensation could be feasible and cost effective, in the context of rapidly increasing costs of malpractice insurance premiums, making malpractice insurance unaffordable for some medical profession specialities[6].

The incidence of potentially compensable events (PCE) was determined. A PCE was defined as "a temporary or permanent impairment of physical or mental function (including disfigurement) or economic loss in the absence of such impairment, which was caused by healthcare management"[6].

In this study, 970 patients were judged to be affected by a PCE (4.6 %). Ninety-four patients died as result of a PCE. Study authors concluded that 165 patients (0.79 % of all included patients) were affected by a PCE that the healthcare provider was legally liable for[6].

The methodology used in this study influenced the Harvard Medical Practice Study (HMPS).

*The Harvard Medical Practice Study*

The HMPS is a landmark epidemiological AE study and has become the model for subsequent, similar studies within the field.

An AE was defined as an "injury that was caused by medical management (rather than the underlying disease) and that prolonged the hospitalisation, produced a disability at the time of discharge, or both"[9]. The HMPS and the later Utah and Colorado Study (UTCOS)[48] incorporated the narrower definition of "negligence", rather than "preventability", which came to be used in subsequent studies[26]. A negligent AE was defined as "an injury caused by the failure to meet standards reasonably expected of the average physician, other provider, or institution"[28].

A two-stage record review was carried out on a random sample of 30,121 hospitalisations during 1984 in 51 acute care hospitals in the New York State. First, screening for one of 18 screening criteria (modified from the CMIFS) performed by trained nurses or medical-record administrators; and, second, implicit review by two physicians independently. An independent review by a physician-supervisor was performed if discrepancies in the identification of AEs were present between the reviewers[9].

The HMPS used a six-point causation and negligence scale. A score of ≥ 4 was regarded as being an AE resulting from healthcare management or being negligent. The HMPS used a different severity scale than was used for the CMIFS[6 28].

The state-wide AE incidence was estimated to 3.7 % of hospitalisations and 27.6 % of the AEs were due to negligence. AEs contributed to death in 13.6 % of the cases[9].

*The Utah and Colorado Study*
A second study was performed in the US, and served to investigate if the results from the HMPS were similar in other states, though using another time period. A random sample of 15,000 discharges during 1992 from 28 hospitals was included[48].

The AE and negligent definitions were the same as those used in the HMPS, as was the 2-stage review process. In the UTCOS, nurses carried out the screening stage and only one physician performed the review. The UTCOS also used 18 screening criteria but changed two criteria that had been used in the HMPS. Only those AEs with the highest disability were included[48].

The annual AE incident rate was estimated to 2.9 % of hospitalisations in both states. In Utah, 32.6 % of the AEs were deemed to be due to negligence, while the corresponding rate for Colorado was 27.4 %. Death occurred in, overall, 6.6 % of AEs[48].

*The Quality in Australian Health Care Study (QAHCS)*
The previously described studies were performed by assessing AEs related to negligent practice in a malpractice litigation context. The HMPS initiated a debate about AEs and quality of care in other countries. Australia was the first country to perform a national AE study. This was the first record review study of AEs undertaken from a quality improvement perspective with the objective to improve patient safety[1].

A random sample of 14,179 admissions from 28 hospitals in two states during 1992 was included and reviewed in two stages, as in previous studies, but with some modifications[26].

An AE was defined as "an unintended injury or complication which results in disability, death or prolongation of hospital stay, and is caused by health care management rather than the patient's diseases"[26].

Screening by trained nurses considered 18 screening criteria. These screening criteria were somewhat different to those used in the HMPS. Some screening criteria were rephrased (to a wider use) and some removed, but others were new. Positive cases were later reviewed independently by two medical officers. If the physicians disagreed they reviewed jointly again, presented their results for a third medical officer and reached consensus. Only the AE with the highest disability was included[26].

A 6-point scale, similar to the negligence scale used in previous studies, was used to assess the preventability of an AE. Preventability was assessed as "an error in management due to failure to follow accepted practice at an individual or system level", accepted practice being "the current level of expected performance for the average practitioner or system that

manages the condition in question". The causation scale was close to that in the HMPS but the QAHCS used a lower threshold for assessing causation (≥ 2). The severity scale was slightly dissimilar, using a 5-point scale instead of a 6-point scale as in the HMPS[26].

The QAHCS identified AEs in 16.6 % of admissions and, of these, 51 % were judged preventable. Most AEs resolved (77.1%) within 12 months but 4.9 % resulted in death. Orthopaedic AEs were common but judged to be less preventable than AEs identified in most other specialities[26].

### Retrospective record review using trigger tools

The concept of a "trigger" was used by Classen et al.[49 50] to describe an automated method using electronic triggers to detect potential adverse drug events. This method was found to be impractical in many settings due to costs and the requirement of customised software linkage to pharmacy databases[51]. Resar et al.[21] describe the evaluation of different trigger tools that do not require computerised technology.

The GTT methodology, developed in late 2003 and popularised by the Institute for Healthcare Improvement (IHI)[25], is a method now used internationally for retrospective reviews in patient safety work primarily in acute care settings. The GTT was primarily designed as a measurement tool in clinical practice to estimate and track AE rates over time, extending beyond traditional incident reports; and GTT aims to measure the effect of safety interventions[36 51-53]. The GTT has spread from collaborative projects to large-scale improvement initiatives, such as, IHI´s "5-Million Lives Campaign" in late 2006 to promote prevention of AEs in the US[25]. Specialised trigger tools have been developed, both before and after the initiation of GTT for various specific clinical settings[21 54-62]. The GTT is intended to be a practical tool to improve the detection of AEs, not an approach to measure the incidence of AEs[21].

The AE is defined as "unintended injury resulting from or contributed to by medical care that requires additional monitoring, treatment or hospitalisation, or that result in death"[25] and is a more inclusive definition than those used with HMPS methodology.

The GTT includes 54 triggers to identify potential AEs, for example, "post-operative increase in troponin levels" or "admission to intensive care post-operatively"[25]. Record review has been criticised for being too expensive and time consuming to be used in routine situations[21], therefore GTT restricts each primary reviewer to review the record for not more than 20 minutes. Primary reviewers, often registered nurses (RNs), are instructed to focus on their search for triggers, and not comprehensively read the record, as in the HMPS methodology. The review team discusses the findings together but the physician reviewer is the final arbitrator. The physician does not generally review the record but does authenticate consensus findings and the severity rating of the AE, as well as responds to questions from the primary reviewers[25].

The GTT neither includes omissions nor seeks to determine the nature of AEs. The GTT also use another severity scale compared to the HMPS methodology[25 63]. All identified AEs are included, not only the most severe as opposed to many of the studies using the HMPS methodology[25].

The Swedish GTT version[64] contains 53 triggers. The triggers "restraint use" in the care module and the use of the indefinite "other" as a choice in the medication module were excluded. A trigger, "occurrence of any postoperative complication", was added to the surgical module. The Swedish GTT version has included the same preventability scale as in the HMPS methodology[13], in contrast to the origin GTT version[25], which does not examine the preventability of an AE.

AE rates between 14.6-40 % have been detected by using GTT[51 52 54 65-68].

### Strengths and limitations

Retrospective record review is commonly used, uses available data[31] and has been found to identify more AEs compared with many other methods[16 22 69-78]. Record review can be repeated over time and specific AE types can be targeted, not only the overall AE rate[18]. The AE occurrence, severity and preventability can be overestimated due to, for example, hindsight bias; but it can also be underestimated, due to, for example, incomplete documentation. Quality of documentation affects results, including the detection of active failures and latent conditions[31]. A hindsight bias may be present when the situation actually faced by the personnel is inevitable grossly simplified by the reviewers in retrospective. Also, if an outcome is known to be adverse, a hindsight bias may result in more critical review, leading to the erroneous detection of an artificially high AE rate[18 79].

Studies have found that with RRR, inter-rater reliability is poor-to-moderate[31]. This could be due to several factors, including skills, clinical specialty, experiences and predetermined views held by the reviewers, as well as training or education in the methodology[1 23]. Criticism has been put on RRR as time consuming, expensive, weak on providing real-time information, and that it is not possible to collect additional information not already documented in the records[18 78].

In the discussion part of the thesis RRR strengths and limitations will be further addressed.

The validity of RRR is now widely accepted. Moreover, the methodology may be further developed by using information technology to search for patient safety information and the use of databases. Adaptations have been developed to counter limitations[1]. Olsen[1] describes in her thesis development of a method named "real-time record review" that uses local personnel to perform review close to the patient's discharge.

Table 2 displays differences between the HMPS methodology and the GTT.

**Table 2.** An overview of the origins of the Harvard Medical Practice Study method and the Global Trigger Tool

| Origin and perspective | Definition/inclusion and method of review | Review stage 1 | Review stage 2 | Criterion/trigger | Sample size and time frame for inclusion |
|---|---|---|---|---|---|
| **Harvard Medical Practice Study(HPMS)**[9 28 29] with subsequent modifications[13 26 27 44-47]<br><br>Medicolegal and focus on negligence the first studies and thereafter quality improvement and preventability perspective<br><br>Research method | An unintended injury or complication that results in disability at discharge, death or prolonged hospital stay and is caused by healthcare management rather than the patient's underlying disease<br><br>Includes both omission and commission<br><br>Adult, inpatients, often exclusion of e.g. psychiatric and rehabilitation patients<br><br>Two[9 26] - three[13 46] stage retrospective record review | Generally one reviewer per record<br><br>Screening for one of 18 criteria by trained nurses (can be other professionals)<br><br>Comprehensive reading<br><br>No assessment, generally only registration of found criteria, description of the potential AE, and a brief summary of the admission<br><br>No time limit | Mostly two physician reviewers per record<br><br>Detailed independent review<br><br>Assess the AE by using different scales according to e.g. causation, severity, preventability, timing, causes, and types<br><br>Generally includes only one AE per patient i.e. the most severe<br><br>No time limit | General for both methods:<br><br>An indication that patient harm may have occurred<br><br>Directs the medical reviewer to relevant parts of the records by the notes<br><br>Some screening criteria/triggers are AEs by definition e.g. healthcare-associated infections<br><br>Positive screening criteria/triggers may be without connection to patient harm i.e. false positive<br><br>HMPS method:<br><br>18 mostly implicit criteria | Random, big samples to measure the incidence and to generalise the result<br><br>An AE had to have occurred before and during and detected during and/or after index admission<br><br>Different inclusion periods before and after index admission |

| Origin and perspective | Definition/inclusion and method of review | Review stage 1 | Review stage 2 | Criterion/trigger | Sample size and time frame for inclusion |
|---|---|---|---|---|---|
| **Global Trigger Tool**[25 80]<br><br>Quality improvement tool for clinical practice<br><br>Track AE rate over time in a hospital or a clinic | Unintended injury resulting from or contributed to by medical care that requires additional monitoring, treatment or hospitalisation, or that results in death<br><br>Includes commission, excludes omission<br><br>Adult, inpatients, exclusion of psychiatric and rehabilitation patients<br><br>Two stage retrospective record review | Two reviewers per record<br><br>First screening independently for one of 54 triggers by trained nurses (can be other professionals), focus on triggers, no comprehensive reading, reads just relevant parts related to found triggers; second, consensus<br><br>Finds triggers, describes the potential AE, and categorise harm according to NCC MERP index<br><br>No assessment of preventability<br><br>Maximum 20 minutes per record | The team discuss the findings together<br><br>One physician, who does not generally review the record but does authenticate the consensus findings of the AEs, the severity rating, and answer questions from reviewers in review stage 1<br><br>The physician is the final arbitrator<br><br>All identified AEs are included<br><br>No time limit<br><br>The Swedish version includes the same preventability scale as used in the HMPS methodology[13 26 64] | 54 triggers, mostly explicit<br><br>The Swedish version[64] contains 53 triggers, the triggers “restraint use” in the care module and “other” in the medication module were excluded, and a trigger, ”occurrence of any postoperative complication“, was added in the surgical module | Random, small samples sufficient for the design of safety work over time<br><br>10 records every second week or 20 records every month per hospital<br><br>Length of stay at least 24 hours<br><br>An AE had to have occurred before and during and detected during and/or after index admission<br><br>30 days inclusion period before and after index admission |

NCC MERP, National Coordination Council for Medication Error Reporting and Prevention (NCC MERP) index[83]. Category A-D describes risk and no harm incident. Category E-I describes harm.

## INCIDENT REPORTING SYSTEMS

A reporting system refers to the processes and technology involved in the standardisation, formatting, communication, feedback, analysis, learning, response, and dissemination of lessons learned from the reported risks and incidents[42]. The scope of an incident reporting system may vary greatly. Some collect data about risks and incidents, while others are concerned with only severe incidents[41].

Leape[81] describes the characteristics of a successful reporting system, effectively part of a safety information system, as being non-punitive, maintains confidentiality, operates independently from of any authority with power to punish, contains a provision for expert analysis, is timely, is system-oriented and is responsive. Vincent[18] also argues that the system must be easy to use.

There are several incident reporting systems in use at local, regional and national levels in Sweden, used both by healthcare provider and personnel, and by patients and their relatives. These methods will be described in general and, then, overall strengths and limitations will be discussed.

### Incident reporting systems for healthcare providers and personnel

*Clinical incident reporting system*

Our healthcare system has learned from other high-risk industries, leading to implementation of clinical incident reporting systems, which are now widely used but with variable success[1 18].

Systems for self-reporting of incidents by personnel can be designed to be mandatory or voluntary[18]. In Sweden, since 1996, it is mandatory for healthcare personnel to report risks and incidents. In the current statute[82], healthcare personnel are required to report risk for preventable AEs, as well as events that have led to or might have led to a preventable AE. Clinical incident reporting systems vary between the numerous, different County Councils and regions.

*Lex Maria*

All Swedish healthcare provider organisations must urgently report all serious preventable AEs, or the risk thereof, also called "sentinel events", to the National Board of Health and Welfare, a government agency under the Ministry of Health and Social Affairs, according to the lex Maria statute[83 84]. This statute was established in 1937 after a medical error was blamed for four deaths at Maria Hospital in Stockholm. The statute has been revised over the years, and since 2006, includes suicide in relation to care. Of late, lex Maria has been noted to focus more on system failures[83 84]. The number of lex Maria reports has been increasing steadily, from 905 in 2000, to 2,124 in 2010[37 40].

*Medical Product Agency*

The Medical Product Agency statutes require that all healthcare provider organisations urgently report any serious as well as unintended side effects caused or associated with medical products. Also, reporting is required on side effects which are increasing in frequency[85 86]. Serious incidents concerning medical technical products require urgent report to both the Medical Product Agency and the manufacturer[87 88].

## Incident reporting systems for patients and relatives

*Patient ombudsman or direct contact with the healthcare provider*

In the first instance, the patient or relative is recommended to contact the healthcare provider in question by phone, e-mail or post if they have opinions about healthcare, noting that the provider has an obligation to investigate[89]. If problems arise, some hospitals have a patient ombudsman who serves as an impartial link between involved parties. Information from patients and relatives may provide valuable safety information[90].

*County Councils` Mutual Insurance Company*

Sweden like the other Nordic countries has a blame-free and non-tort national patient insurance system. According to the Patient Injury Act[91] a patient or a relative can file a malpractice claim to request compensation for injury in relation to healthcare. All Swedish County Councils and regions contribute economically to the insurance. Only injuries deemed preventable by an external medical specialist reviewer are compensated. The patients must report an injury within three years from when the patient became aware of the injury, and it must be within ten years of the occurrence.

Some 10,000 injuries are reported each year and about 45% of these are compensated[92]. Economical compensation is paid for disability, income loss, expenses, inconvenience as well as pain and suffering[93]. Claim rates vary significantly between specialities. Surgical specialities, especially orthopaedics, have the highest claim rates[94 95]. In the Nordic countries, lawsuits are very rare.

*National Board of Health and Welfare*

When the Patient Safety Act[82] was launched in January, 2011, the National Board of Health and Welfare took over the management of patient complaints from the Medical Responsibility Board (HSAN). Patients and their relatives can make complaints about an event, an individual personnel or a healthcare provider. The National Board of Health and Welfare has adopted a systematic approach for investigating complaints. They may initiate inspections in the same way as in lex Maria investigations. Healthcare providers or personnel may be criticised. Individual personnel may have licenses revoked. Supervision may be required. Patients must generally report the event within two years from the occurrence[89].

*Patients´ Advisory Committee*

A patient's advisory resource exists in every County Council and region. The Patients´ Advisory Committee is a central and independent authority to which patients or relatives can turn to when they have difficulties in contacting or addressing complaints concerning all public financed healthcare. The Patients´ Advisory Committee helps the patient to investigate what happened, propose solutions, inform and guide the patient further, if necessary, to the proper healthcare personnel, provider or authorities. However, the Patients´ Advisory Committee has no authority to award damages or to punish. The central aim is to improve healthcare services by reporting complaints and suggestions from the patients to different providers[96].

## Strengths and limitations with incident reporting systems

National and regional reporting systems may identify rare or serious events, and unforeseen consequences not previously identified by local incident reporting systems. Their value is

based on learning from other's experience[1 18 42 97]; but local systems can often collect more detailed data about local problems[1]. The experiences and observations made by patients and their relatives are important for healthcare in the patient safety work. Due to little overlap between the different reporting systems, data that is not available by other measures can be identified by using patient reported data[38 78 98-104]. Clinical incident reporting systems, besides providing data, may be essential to establish a positive patient safety culture and to raise awareness in these issues[18]. One advantage with incident reporting is that the method is always operational, accessible for all personnel within the healthcare organisation[1].

Some incident reporting systems, especially the national systems, have a time delay for reporting, processing and giving feedback[105]. Incident reporting systems are also biased by a high level of underreporting, selective reporting and hindsight bias[31]. The source of bias may vary over time, amongst personnel, specialities and hospitals; and by event type and severity[35]. Due to underreporting, some argue that data from incident reporting systems may be unreliable as a measure of patient safety and to evaluate the effect in patient safety initiatives[31 35 41 97 106]. These reports may increase learning and, in sufficient number, they can serve to identify improvement areas[23 41]. These reporting systems may need to be supplemented by other more systematic data collection methods[71].

## WHY ARE REPORTING RATES SO LOW AND WHAT AFFECTS THE REPORTING WILLINGNESS?

The underreporting, as mentioned earlier, applies to both mandatory and voluntary reporting systems and is well documented[16 22 54 66 70 71 75 77 95 100 103 107-109]. With incident reporting systems healthcare providers and personnel are obliged to report according to statute[82]. Patients and their relatives can use incident reporting systems to report deficiencies within the healthcare system. What are the reasons for such underreporting?

### Personnel

Completeness of clinical incident reports depends on the willingness to report[110]. Personnel must commit to report or not, and factors which influence this may vary between individuals and professionals[1 111]. Several barriers to clinical incident reporting are described in the literature, including: lack of time, interruptions in workflow, the safety culture or lack of, lack of knowledge and uncertainty as to what constitutes an incident, fear of litigation or disciplinary actions, difficulties to report, lack of feedback and a sense that reporting will not result in any improvement[11 18 97 103 112-115]. That nurses tend to report more incidents than physicians is well reported[112 116-120]. Curiously, physicians report more serious incidents but are less likely to report fatal events[117]. Noble and Pronovost[106] argue that such participation bias hinders the ability to identify and reduce specific risks mostly viewed by physicians and may misdirect interventions. This might also give an incorrect picture of incidents occurring in different care processes and, consequently, with such a bias, the pattern of incidents cannot be generalised.

### Patients and relatives

Patients and their relatives are an obvious source of AE reporting data, though subgroups, for example, by age, have variable reporting characteristics. Studies have shown that patients identify care-related AEs correctly[98 121 122]. Traditionally, patient reports of AEs have been collected, and exploited, by complaint- and claim-reporting systems, which, unfortunately,

underestimate the incident rate[78 103]. Bismark et al.[103] found that severe and preventable AEs have a greater likelihood to be lodged as a compliant. The elderly are a high risk group for AEs[27 75 123-125]; but, they are found to be the least likely to make complaints[103 126]. Possible reasons why patients do not report incidents can be several, including: lack of awareness for having sustained an AE, thinking that the incident is a known unpreventable complication or may be due to the underlying disease; advanced sickness; lack of knowledge that one can make a complaint or malpractice claim; and unwillingness or not sensing the need to report[78]. Having negative experiences were found to affect the patients´ perception of trust and safety in healthcare negatively[100 126], which may affect the reporting rate[100]. In the latter study, conducted in Sweden, they investigated reasons for not filing complaints and identified the main barriers to reporting. They found that patients did not find the strength to report, did not know where to turn and believed that a complaint would not make any difference.

### Healthcare providers

Öhrn et al.[95] found that in Sweden, severe AEs were widely underreported despite the mandatory reporting statute, lex Maria[83 84]. This may be due to many of these events being regarded as complications by physicians. Indeed, the senior chief medical officers' disagreements described in this study report, concerning what should be reported, indicates that the statute, itself, is not clear-cut.

## OTHER DATA COLLECTING METHODS

According to Swedish statutes, healthcare providers are obliged to analyse why incidents occur and how serious they are, rather than target blame. The aim is to achieve quality in systems concerning patient safety work in order to achieve a reduction of preventable AEs[82 127]. Several analytic methods are described in the literature. The methods most widely used in Sweden, as well as internationally, are the *Healthcare Failure Mode Effect Analysis* (HFMEA) and the *Root Cause Analysis* (RCA)[128]. These methods can lead to deep understanding of current weaknesses, contributing factors, future potential safety problems, maintain a system approach and lead to organisational and local learning in order to prevent reoccurrence. The RCA may be affected by recall bias, which may increase over time[18 23 129]. The HFMEA and the RCA tend to focus on severe outcome, require extensive resources and may be dependent on standardisation and education in methodology to provide valid and reliable results[18 41 129].

Since 1975, several *quality registers*, including across a wide range of specialist areas and diagnoses, have been established in Sweden (89 registers in 2011). A register contains individualised data concerning patient problems, medical interventions, and outcomes after treatment. The vision for the quality registries is to constitute an over-all knowledge system that is actively used on all levels for continuous learning, quality improvement and management of all healthcare services. Data is submitted from participated units and patients/relatives and is stored in central databases[130]. Aggregate data is available by these registers and epidemiological research can be performed[37]. Note, however, that validity and reliability within and between the registers and other methods may vary[99 131].

*Administrative data* sources, data often derived from the patient administrative system are used to identify patient safety information and are commonly used in the US. Administrative

data may contain relevant information but these are not established for the purpose of measure safety or quality[18]. Advantages are availability, low cost, digitally compiled, and broad population coverage. This method is also less susceptible to selection bias than some other methods e.g. incident reporting but has limited clinical data and may be incomplete due to inaccuracy and variability of coding practices which affects reliability, for example, primary and secondary medical diagnosis[31 132 133]. Michel[41] suggests these data be used for screening proposes and not to measure patient safety.

*Information technology* applications and methods are increasingly used to identify and provide feedback about incidents. Information technology enables digitised and automated data filtering, searching, counting, and so on. These functions can be used to search multiple, large data sets to focus on specific AEs, e.g., healthcare-associated infections[41 134]. The information technology approach can be costly and no standardised method exists to guide examination of data, and reliability is limited[31 41]. ”Drilling down” into data sets using generic information technology may also not enable one to detect latent conditions[31].

*Mortality and morbidity (M&M) conferences* are one of the oldest methods to examine AE issues. When used for internal evaluation, M&M conferences remain an important resource and are accepted, especially within surgical specialties[1]. Thomas and Petersen[31] report that M&M conferences can detect active failures and latent conditions but are limited by hindsight bias. The selection for conference cases is biased towards severe AEs, while cases that led to minor AEs, which constitute the majority, are often not presented. Olsen[1] cautions of the risk of poor analysis of events that misses complex contributing factors, and promotes individual blame. The reliability is likely to be variable due to several factors e.g. the objective, case selection, and the attendance of personnel[1 41].

*Interviews and questionnaires* can be valuable methods to detect incidents, active failures and contributing factors and can provide detailed data not otherwise available[23 42 135]. Interview as a method has pedagogical and communicative advantages, and could be more routinely used in continuous patient safety improvement[69]. Interview, however, does not provide data about trends[42]. Interviews are resource intensive[97]. Recall bias may also be of concern. Interviews and questionnaires used with different personnel categories may lead to different reporting of incidents[23 135]. For valid and reliable results, trained interviewers and validated questionnaires are needed.

*Observation or videotaping* may increase the understanding about the processes and dynamics that affect the outcome and identify solutions on safety problems[42]. Michel[41] describes some advantages with these methods, which are not as influenced by willingness to report, memory, ability to communicate or subject’s level of knowledge concerning errors. Other advantages are that these can be performed prospectively[97] and may be more sensitive to detect active failures than other methods[31]. However, observation may be costly, labour intensive and presents practical difficulties[1]. Thomas and Petersen[31] describe further some of the limitations with observation, such as, that it is time-intensive to train reliable observers, concern of confidentially, hindsight bias and the fact that personnel may alter their normal behaviour when knowing they are observed, known also as the Hawthorne effect.

*Clinical surveillance*, often used in epidemiological studies, could be effective to assess explicit incidents. The methodology may be costly and less useful in detecting contextual information on latent conditions due to the focus on specific outcomes or elements of care at a specific time and place rather than the full range of incidents[31]. Traditionally used outcome studies seldom separate events due to healthcare management from those resulting from the patient's underlying disease[136].

*WalkRounds* is a management tool, introduced in the US in 2000 by Allan Frankel, to sustain good relations with frontline personnel, promote conversations to identify incidents, and gather safety information to enhance decision making around patient safety in order to take appropriate action and to support the safety culture[137]. Since its introduction the method has been adopted and modified in a number of different countries[138]. The identified data can neither be used to estimate AE rates nor evaluate the AE reduction after patient safety initiatives; but, the effect on patient safety culture can be measured and this method is low-cost and easy to carry out[42].

An overview of the different data collecting methods is presented in Table 3.

**Table 3.** Overview of different patient safety information methods´ strengths and limitations

| Methods | Strengths | Limitations |
|---|---|---|
| Record review | Use already available data<br>Valid compared to most other methods<br>Can yield detailed information<br>Can detect active failures if adequate documentation<br>Can estimate prevalence and incidence<br>Can be used to assess the efficacy of interventions | Rely on documentation quality<br>Hindsight bias<br>Reliability concerns<br>Resource extensive for continuously use or on large scale |
| Incident reporting systems | Some can detect latent conditions<br>Provide multiple perspectives over time<br>Can be run as a routine<br>Can detect rare incidents (regional/ national reporting systems) | Unreliable as a measure of patient safety<br>Difficult to generalise<br>Unreliable to evaluate interventions<br>Reporting bias<br>Hindsight bias<br>Report propensity may vary between groups and over time<br>Lex Maria and patient claims are focused on severe events in accordance with the statutes´ purpose<br>Timeliness |
| HFMEA and RCA | Can yield detailed information<br>Can suggest latent conditions and future problems | Resource extensive<br>Reliability and validity of the conclusions<br>Recall bias (RCA)<br>Hindsight bias (RCA)<br>Tend to focus on severe outcome |

Table 3 continues on page 28.

| | | |
|---|---|---|
| National quality registers | Nation-wide comparison possible<br>Provide data for QI and healthcare planning<br>Epidemiological data | Timely feedback<br>Data quality and coverage may differ between registers and within registers<br>Focusing on complications |
| Administrative data including information technology (IT) | Utilises readily available data<br>Inexpensive (after initial investments, IT)<br>Routinely collected<br>Less susceptible to selection bias<br>Can screen big populations<br>May use real-time monitoring (IT)<br>Integrates multiple data sources (IT) | Affects by incomplete or inaccurate data<br>Data divorced from clinical context<br>Not good at detecting latent conditions<br>Expensive (initial IT investments)<br>Sensitive to programming, data entry errors and/or incomplete/not standardised data (IT) |
| Morbidity and Mortality conferences | Familiar and acceptable to especially surgeons<br>Can suggest latent conditions | Hindsight bias<br>Selection bias<br>Infrequently utilised<br>Focus on rare events<br>Often poor analysis<br>Risk of individual blame |
| Interview and questionnaire | May provide data otherwise unavailable Can yield detailed information<br>Can suggest active failures and latent conditions | Recall bias<br>Expensive (interviews)<br>Needs trained interviewers and validated questionnaires |
| Observation | Provide data otherwise unavailable<br>Detects active failures<br>Can yield detailed information<br>Prospective method | Expensive<br>Difficult to train reliable observers<br>Potential Hawthorne effect<br>Logistical problems<br>Poor at detecting latent conditions<br>Potential hindsight bias<br>Potential concerns about confidentially<br>Possible to be overwhelmed with information |
| Clinical surveillance | Potential valid and reliable<br>Prospective | Used to measure explicit defined events<br>Time consuming and expensive<br>Not good for detecting latent conditions |
| WalkRounds | Management commitment<br>Supports safety culture<br>Can detect latent conditions<br>Low-cost | Reporting bias<br>Non-standardised source of data |

Adapted and further developed from[18 31 41].

# LEARNING FROM AND ACTING UPON SAFETY INFORMATION

A better understanding of the frequency, nature and contributing factors of specific incidents is a first step to preventing reoccurrence. But measurement alone does not lead to safer healthcare; the incidents must be accurately analysed, lessons learned and appropriate action taken[23 42].

Modern safety research defines safety existing within a system, not as the absence of risks and incidents. Safety concerns the system's capability to deal with vulnerabilities and risks so that they do not harm, for example, personnel, equipment or patients. Other industries, such as aviation and nuclear power industry, have gone far beyond the healthcare services in efforts to develop systems to monitor and improve safety. These are called high reliability organisations. Experience has shown these industries that organisations need safety information systems that collect, analyse and feedback information about risks and incidents to the organisation, and that learning must take place to improve safety. Such systems combine information about risks and incidents with decades of analyses that generated knowledge about specific risks and incidents and actions taken[30]. Well-developed system to measure and monitor patient safety within the healthcare system is still limited[30 139]. One difference between incidents in the healthcare systems, and the high reliability organisations, is that severe incidents in healthcare usually affect one patient at a time and are less visible and harm happens to a third part, compared to an airplane crash with hundreds of deaths including the crew[10].

Safety information systems provide data that can facilitate decision making at different levels of the healthcare system, which aim to improve safety and quality. Ideally, safety information systems provide real-time information[78]. Kjellén[30] describes the requirements for an effective safety information system, including: that the data collection should be valid, reliable and provide adequate coverage; information should have relevance, be timely, be comprehensive but also easy to gather and available when needed; that methods used should be easily understood and accepted; that involvement be promoted and encouraged involvement; and that it should be cost-effective.

WHO[42] concludes that the healthcare system finds it difficult to learn from incidents and disseminate lessons learned to a wider audience, for example, making available data that may be used to formulate recommendations for system changes.

No-harm incidents and near miss are events that are often overseen, though there is potential for harm under other circumstances. Also, learning from these events could lead to more proactive change in patient safety initiatives[1]. Kaplan and Fastman[140] suggest that such events are a valuable source of patient safety information because of their similarity to, and greater frequency than, events causing harm to patients. No-harm incidents and near miss simply are an increased risk of harm but allow that some form of recovery took place, either by chance or by actions of an individual, team, or organisation before harm could occur.

Analysis of these events gives opportunity for better understanding of recovery mechanisms, in addition to what contributing factors are. Also, these constitute a low cost learning tool for safety compared with events where the patient is actually harmed[34 141 142].

## THE WIMMERA CLINICAL RISK MANAGEMENT MODEL

At the Wimmera Base Hospital, in rural Australia, the Wimmera Clinical Risk Management Model has been in use since 1989, one which incorporates a systematic, on-going approach for patient safety. It consists of five basic, sequential and coordinated components: multi-method AE identification tools; analysis; determination and prioritising risk of each AE; initiate interventions and finally monitoring the effects of these[77 143].

In the Wimmera Clinical Risk Management Model, an AE is defined as "an untoward patient event, which under optimal conditions is not a natural consequence of the patient's disease or treatment". All identified AEs are included in the review process[77].

Their review process differs from those used in other studies using screening criteria. Manual screening is performed on all medical records closely after patients have been discharged. Eight (later nine) explicit screening criteria were chosen in order to be easily and accurately used by non-clinical personnel as a part of their normal duties. A positive record in the screening process was then reviewed by a senior medical personnel using a standardised assessment. The AEs were, after the physician review, discussed by a surveillance committee comprised of personnel from a cross section of different clinical professions, and whose task was to identify situations that should be regarded as warranting implementation of measures for patient safety[77]. Some of the screening criteria are the same as in the HMPS methodology[28]. As opposed to the HMPS methodology, the Wimmera Clinical Risk Management Model uses the screening criteria obtained locally and continuously. They have also a different review process, using an AE definition that identifies a wider range of events and have some other assessments scales e.g. the severity scale[143 144].

In the Wimmera Clinical Risk Management Model, other systems of reporting AEs have also been implemented. For example, clinical incident reporting and general practitioner feedback on AEs are used. These reporting systems are intended to complement each other. The method is not built to identify all AEs, as is GTT, but to find enough to form a basis for intervention. The rate of AEs over time is measured and the effectiveness of implementation of interventions. This is a part of their routine, longitudinal safety measurements[77].

Evaluation of the Wimmera Clinical Risk Management Model found the annual AE rate decreased from 1.35 % for all discharged patients to 0.74 % after interventions between the first and eighth year when[77].

# What can this thesis add to the field?

When we started to plan our first study in the spring 2004, a limited number of terms and definitions were available in Sweden within the patient safety field. No RRR examining the occurrence of AEs had been performed in Sweden. Traditionally, prospective studies within surgery focused on specific patient groups or treatments and followed the complication rates without distinction whether a complication was related to healthcare management or to the patient's underlying condition or disease[136]. International AE studies using RRR reported few data about the incidence and nature of orthopaedic AEs because these data were often included with all other surgical specialities. This makes it difficult to learn in relation to orthopaedic patients and context. The results from these studies were often considered from a national or regional perspective not from a local, departmental one. We determined it would be valuable to locally collect patient safety information used to form a basis for interventions because the frequency and nature of AEs has been found to vary significantly between specialities[9 12 26]. In Sweden, incident data is available from several different reporting systems on different levels, as described earlier. Incidents can be registered in these systems but validity can be questioned regarding these data. Data on RRR and orthopaedic AEs was, at that time, limited. The Wimmera Clinical Risk Management Model was chosen due to its capability to use different safety information methods at the same time in a structured way.

While there is growing awareness of quality and safety problems in healthcare systems it remains uncertain how best to accomplish and sustain improvement over time. There is unfortunately no given "best way" to implement and improve quality and safety[145]. We wanted, therefore, to evaluate an on-going nursing project at the Orthopaedic Department by using a retrospective case study design in order to add knowledge concerning nursing quality and safety improvement. Another aim was to evaluate the feasibility of RRR for evaluating nursing outcome.

There has been theoretical debate on what is the "best" record review method. The methods using screening criteria, for example, HMPS, are mostly used in research and neither in local, routine data collection by, for example, clinical personnel, nor for evaluating initiated patient safety interventions[1]. GTT is a method designed to identify AEs in order to form a basis for improvement and to follow-up initiated interventions[25]. In the year 2006, the GTT was translated to Swedish, tested and introduced in small scale in some County Councils in their patient safety work. In 2008, a national handbook was written and distributed for the methodology. This is now in use in many County Councils in Sweden[64]. This method is widely used in hospitals in the US but the original GTT method is not scientifically well evaluated and published concerning, for example, PPV of the triggers. Most reported studies address the specialised trigger tools. We thought that it would be of interest to compare GTT with the more scientifically evaluated method, as was used in the Swedish national AE study (HMPS method). Moreover, our study would use the same cohort and definitions to allow comparison. To the best of our knowledge, such comparison has not been published before.

# AIMS

## GENERAL AIM

The general aim of this thesis was to evaluate the capability of retrospective record review methods to identify patient safety and quality information in orthopaedic care.

## SPECIFIC AIMS OF THE PAPERS

Paper I
To evaluate the capability of a modified Wimmera Clinical Risk Management Model to identify orthopaedic adverse events.

Paper II
To identify patient risk factors that may lead to adverse events in orthopaedic inpatients.

Paper III
To examine the design and outcome of quality improvement theme months in orthopaedic nursing and the feasibility of retrospective record review as a method for evaluation of impact on nursing practice.

Paper IV
To compare the capability of two retrospective record review methods to identify orthopaedic adverse events and patient outcome and to identify the screening criteria and triggers that predict orthopaedic adverse events.

# MATERIALS AND METHODS

## CONTEXT

All studies were performed at Danderyd Hospital, an acute care university hospital in the Stockholm metropolitan area with a catchments area of approximately 450,000 inhabitants. The hospital provides acute and elective specialist care focusing primarily on major disease groups. In 2010, the hospital had 435 beds and 101 technical beds; during that year had 43,190 inpatient admissions, of which 36,440 were acute. Also, that year, 12,048 and 4,920 operations were performed within inpatient and ambulatory care, respectively. The hospital had 3,400 employees in 2010.

The hospital has a four-ward, 50 (paper I and II) or 52 (paper III and IV) - bed, orthopaedic department, which treats both elective and acutely admitted inpatients. The Department admits approximately 3,500 inpatients each year, which are predominantly acute care. Patients with hip fracture and elective surgery for hip and knee replacements are the major patient groups.

## PATIENTS AND PROCEDURE

### Papers I and II

This study was based on retrospective record review.

The Wimmera Clinical Risk Management Model[77 143], described earlier, was translated into Swedish and adapted to our Swedish healthcare context by altering some of the safety information methods to those used in Sweden. Screening criteria relevant for orthopaedic care were added. The screening criteria used are presented in Table 1 in Appendix.

*Inclusion*

All digitalised records of the 395 orthopaedic inpatients during August and September 2004 were included. An AE had to have occurred during the index admission or within 28 days of discharge from the Orthopaedics Department to be included. Symptoms present when the patient was first admitted to hospital were excluded.

*Definitions*

An AE was defined as an untoward or unintended patient event caused by healthcare management. Healthcare management covers the actions of individual healthcare personnel and also the systems and care processes used in delivering healthcare. It includes both acts of omission and acts of commission[27].

A preventable AE was defined as an error in healthcare management due to failure in following accepted practice at the level of the individual or of the system[26 44].

*Data collection*
A three-stage retrospective review of digitalised records was used. The two-stage review process used at Wimmera Base Hospital was complemented with a third, consensus stage.

In stage 1, all records were screened by a senior orthopaedic RN for the presence of one or more of 12 predefined screening criteria (Table 1, Appendix). A check of the traditional incident reporting systems was also carried out on the same cohort. As a further complement general practitioners was contacted by post and requested to complete and return a standard form if they identified any potential AE within 28 days after the patient's discharge. Returned forms were included in the screening process. A county council-wide database in Stockholm, called Predo, was used for the detection of potential AEs that occurred within the first 28 days of discharge. If a potential AE had occurred, the medical record was requested from the appropriate healthcare provider and then screened.

Two senior orthopaedic surgeons received a copy of the records with potential AEs and the accompanying nurse screening form for independent reviews.

In stage 2, positively screened records were reviewed using a standardised protocol, review form (RF1), similar to the one used at Wimmera Base Hospital[143].

The degree of healthcare management causation[9 26] and of preventability[26] was assessed using a six-point scale. Events with a score of four or higher, requiring evidence that the causation and preventability is more likely than not, were included as AEs and preventable AEs, respectively (Table 2, Appendix). The severity of the AE was graded on a seven-point scale. A score of three or higher indicated a major AE (Table 3, Appendix)[143 144]. Additional assessments about the nature of the AE, including underlying causes, and which were based on existing information in the records, were also performed.

In stage 3, the physicians reviewed and analysed all the records from stage 2 together to obtain consensus. A RF2, identical to RF1, was completed for all events.

If a patient had more than one AE, each was reviewed and given its own RF1 and RF2.

To validate the screening process and to identify false negatives, the records of every tenth admission (n=28) that had not met any of the 12 screening criteria, or had been deemed not to contain a potential AE by the RN, were fully reviewed by one of the reviewers.

A flowchart of the review process and outcome is presented in the result part below (Figure 1).

### Paper III
We used a retrospective mixed-methods case study design[146].

In the first AE study I also systematically collected nursing data concerning, for example, risk assessment for pressure ulcer and malnutrition. The results showed that risk assessments were not being performed. These outcomes and poor results of annual county council-wide follow-ups led to the initiation of improvement theme months. The aim was to achieve

department-wide improvement of key quality and patient safety areas in a systematic, sustainable and timely manner.

*Improvement theme months*
The quality improvement interventions used nursing teams, consisting of RNs and nurses assistants. The teams focused on one improvement theme at a time in two-month cycles, hence the term, improvement theme months. Improvement theme months used a modified bottom-up approach that included defined objectives, easy-to-use follow-up measurement, education, changes to daily routines, "reminder months" and feedback on data. In each improvement theme months, different evidence-based concepts, e.g. risk assessments tools, were implemented at the same time in all orthopaedic wards.

*Data collection*
The orthopaedic nursing improvement efforts, for the years 2003-2010, were studied. Data collection included participant observation of the improvement theme months´ efforts and outcome; document studies and reviews of hospital performance indicator measurement. The relationship in time between different improvement initiatives was mapped and analysed using public local, hospital, county council, and national wide documentation.

This study incorporated retrospective data from records containing risk assessment and outcome of pressure ulcers. We studied the impact of the first of these improvement theme months (pressure ulcer prevention) because this enabled retrospective evaluation of process and outcome measures prior to, during and after improvement theme months using structured review. For the evaluation of the improvement theme months´ impact, monthly point prevalence measurement (PPM) data was collected for 46 months (January 2007-October 2010) using explicit RRR prior to, during, and after intervention. These 46 months included 2,281 admissions. The PPMs were scheduled on different days of the week so that all days of the week were included. Patients hospitalised for more than one month were only counted once in the result, i.e. the admission month. The 46 monthly point prevalence samples ranged from 28 to 66 admissions and the outcome were displayed in time series diagrams. These data were compared to annual PPMs mandated in the area of the County Council, and were collected from hospital administrative data, starting with the year 2003.

### Paper IV
In the last study (paper IV), we compared the HMPS and the GTT (Table 2).

*Inclusion*
Three hundred and fifty randomly selected orthopaedic admissions (one department), out of a total population of 3,701, were included, irrespective of length of stay. The orthopaedic admission in the random sample constituted the index admission.

To be included in the study, the AE had to be related to care given in the Orthopaedic Department and, additionally, one of the following criteria had to be met:
(i) The AE had to be caused within 30 days before index admission, leading to the index admission;
(ii) The AE had to occur and be detected during index admission; or
(iii) The AE had to be caused during index admission and detected within 30 days of index

discharge from the Orthopaedic Department. AEs in this criterion were not required to result in a new admission.

*Definitions*
An AE was defined, for both methods, as an unintended patient harm that was caused by healthcare, rather than the patient's underlying disease process. In contrast to other HMPS method studies[9 13 27 46], we did not require that the AE cause disability at the time of discharge or prolonged hospital stay. Both AEs due to acts of omission and acts of commission were included.

A preventable AE was defined as an error in healthcare management due to failure in following accepted practice at the level of either the individual or the system [26 44].

A no-harm incident is an event that reached a patient but no discernable harm resulted [20].

*Data collection*
Two teams comprised each of a RN and two physicians were assigned one to each method.

To ensure validation a thoroughly written manual and well-prepared education sessions were used as for training and familiarising team members with the method processes.

A two-stage RRR was performed, plus a crossover review.

In stage 1 of the review, all records were reviewed by the RNs, one for each team. They screened for the presence of one or more of 18 predefined screening criteria, and for 53 triggers, respectively. For every screening criterion or trigger detected, judgement was made by the RN as to whether the finding reflected the presence of a potential AE or not. The potential AE was described. The RN in the HMPS team also included potential no-harm incidents when using the screening criteria.

In stage 2 of the review, the physicians performed an independent review of the records containing a potential AE. Judgement was made regarding whether patient harm had occurred or not. A determination was made as to whether of healthcare causation and preventability was indicated using the same scale as in paper I and II (Table 2, Appendix)[9 26]. The severity of the AE was assessed. Importantly, however, the two methods comprised different severity scales (Table 4 and 5, Appendix)[9 63]. In addition to these scales, all physicians documented which screening criteria and triggers corresponded to each AE (Table 6 and 7, Appendix). If a record contained more than one AE, each was reviewed separately. The HMPS method team also investigated the nature of the AEs, based on the available information in the records, and according to HMPS methodology. The physicians in the HMPS team performed the same review procedure for each no-harm incident. The latter results will be presented in a separate paper not included in this thesis.

A random sample of records was selected to be reviewed by the physicians blinded to the other physician in respectively team in order to examine inter-rater reliability. After independent review of all potential AEs, the physicians in each team met together and discussed the duplicated, previously reviewed records containing potential AEs. If disagreement arose, they worked to reach consensus.

Aiming to compare the physicians' judgements across the full set of AEs, a crossover review was performed. All confirmed or rejected AEs were compared between the two methods and all discrepancies were analysed. This crossover review was the evaluation between the two methods to reach the "true documented and confirmed AE rate".

To validate the nurse review process in stage 1, we used internal validation for each method in a two-step procedure. A senior orthopaedic surgeon and a senior anaesthesiologist were available to provide advice on request. The RNs and physicians entered their findings from the different stages directly into a database.

A flowchart of the review process and outcome is presented in the result part below in Figure 2.

## STATISTICAL METHODS

An overview of the statistical methods used in papers I-IV is shown in Table 4.

**Table 4.** Overview of the statistical methods used in papers I-IV

| Method | Paper I | Paper II | Paper III | Paper IV |
|---|---|---|---|---|
| Descriptive | X | X | X | X |
| Cohen's kappa | X | | | |
| Spearman rank-order correlation coefficient | | X | | X |
| Chi-square test | | X | | X |
| Logistic regression | | X | | |
| Mann-Whitney *U* test | | X | | X |
| Fisher's exact test | | | X | |
| Wilcoxon matched pairs test | | | | X |

Overall, frequency, mean, percent, a 95 % confidence interval, median, range, inter-quartile range, time series diagram, and PPV were used for descriptive purposes.

Cohen's kappa was used to calculate inter-rater reliability between the two reviewers in paper I. The kappa values are scaled to correspond to different classifications that better describe agreement or consistency, as presented in Table 5[147].

**Table 5**. The kappa values different classifications

| Classification | Kappa values |
|---|---|
| Poor agreement | Less than 0.20 |
| Fair agreement | 0.20 to 0.40 |
| Moderate agreement | 0.41 to 0.60 |
| Good agreement | 0.61 to 0.80 |
| Very good agreement | 0.81 to 1.00 |

In paper II, Spearman rank-order correlation was used for correlation analysis on variables that were not normally distributed. Chi-square was used to identify dichotomous variables

associated with AEs. The Mann-Whitney *U* -test was used to compare continuous variables not normally distributed. A multivariate analysis with AE as the dependent variable was performed using stepwise logistic regression. A significance level of ≥0.2 was required for inclusion in the regression model.

Descriptive statistics and time series diagrams were used to examine the patterns of improvement theme months' impact over time in paper III. A qualitative analysis of various initiatives at different levels in the healthcare organisation, and which was related to the improvement theme months, was also carried out. Fisher's exact test was performed for the hypothesis of the differences between two proportions.

In paper IV, Chi-square was used to compare HMPS and GTT methods regarding the respective proportions of verified screening criteria and triggers. The Mann-Whitney *U* test was used to analyse the numbers of screening criteria and triggers in relation to AE (yes/no). The difference in nurse review time between the HMPS and GTT methods was analysed by Wilcoxon matched pairs test. The Spearman rank order correlation coefficient ($r_s$) was used to assess the association between the methods regarding nurse review time and also to investigate "learning curves".

Differences between groups were considered to be statistically significant if the p-values were <0.05 in a two-tailed test.

Calculations were made using JMP version 5.1, SAS 9.1.3, Statistica 9.0, MINITAB 16.2.1 and Excel 2003.

# RESULTS

In this section main findings from the respectively papers are presented.

**PAPER I**

In 127 of the 395 records assessed (32%), one or more screening criteria was found to be positive and was judged to represent a possible AE at stage 1. These 127 records contained 136 potential AEs (Figure 1).

At stage 2, the two orthopaedic surgeons independently judged the number of AEs to be 59 and 52, respectively (Figure 1).

At stage 3, 60 patients of the 395 (15%) were deemed to have been be affected by 65 (16%) AEs (Figure 1). Four of these AEs were detected by searching the Predo database. Of all 65 AEs, 61 included harm, whereas, four were deemed as no-harm incidents.

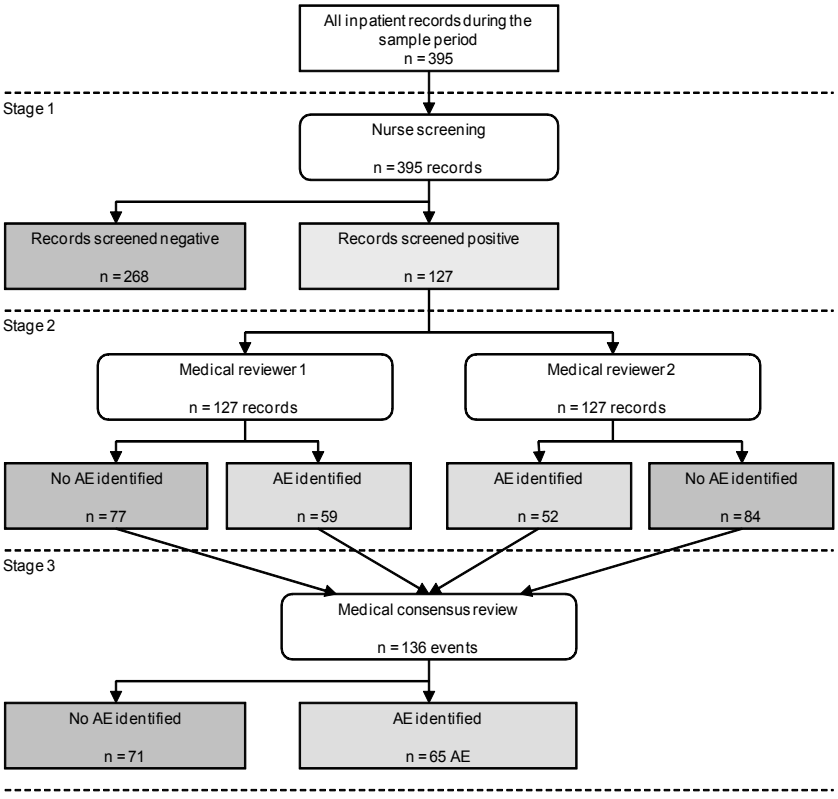


***Figure 1.*** *The three-stage review process for identifying adverse events using retrospective record review*

Seven potential AEs were identified by traditional incident reporting systems, four of which were deemed to constitute an AE and affected three patients. One of these patients had the same AE, a healthcare-associated infection, reported in two incident reporting systems, which was also identified by RRR. The overall number of identified different AEs was 67 (17%) (Table 6).

**Table 6.** Methods used for detecting adverse events

| | Medical record review | PSR | Lex Maria | PaN | HSAN | Clinical incident reporting | GP AE feedback |
|---|---|---|---|---|---|---|---|
| Number of AEs identified | 65 | 2* | 0 | 1* | 0 | 1 | 0 |

AE= adverse event, PSR=the Patient Injury Insurance, Lex Maria=the Swedish statute concerning reporting "sentinel events", PaN=the local Patients´ Advisory Committee, HSAN=the Medical Responsibility Board, GP=general practitioner
* One was also identified in the medical record review.

Of the 65 AEs found from the retrospective record review, 34 were estimated to be preventable. Thirty-four of 65 AEs were considered to constitute major AEs, ranging from three-to-six on the severity scale, and with the majority scaled as three.

The kappa values indicating reliability between the reviewers' judgements of the presence of an untoward or unintended patient event, healthcare causation, preventability and severity at stage 2 were fair-to-good, that is, 0.44, 0.35, 0.52 and 0.62, respectively. The overall PPV for the screening criteria was 48 %.

In the validation of the nurse screening process no unidentified AEs could be found.

## PAPER II

In paper II only AEs identified by RRR were included.

Patients with ASA classification three and above were found in paper II to be more predisposed to AEs ($p=0.03$) and high ASA scores were also strongly correlated with age ($p < 0.0001$). Stepwise logistic regression analysis showed that AEs were more likely in patients who underwent a surgical procedure, ($p=0.02$). Patients who experienced AEs had longer hospital stays than patients who did not experience AEs (median length of stay 6 vs. 3 days, respectively, $p< 0.0001$). Patients who experienced major AEs were significantly older than patients who experienced minor AEs ($p<0.01$).

Fifty-nine of 65 AEs occurred in patients who underwent surgical procedures. However, only nine of these 59 AEs were related specifically to surgical or anaesthesia technique. Overall, 56 of 65 AEs were judged to be due to deficiencies in the orthopaedic healthcare processes, and thus involved all categories of personnel.

Of all AEs, healthcare-associated infections (n=20) were most common.

## PAPER III

Substantial differences were found for risk assessment rates for pressure ulcers by using Modified Norton Scale (MNS)[148] both in the longitudinal follow-up (p=0.0001) and in the annual county council-wide measurements. Reduction in pressure ulcers rate was observed in the annual county council-wide measurements. In the longitudinal data, wider variation in the pressure ulcers rate was seen; but the baseline period of ten months, when compared with the corresponding period the last year showed a significant difference (p=0.004).

Measurement data over 46 months shows a difference compared to the advertised annual county council-wide measurements for MNS and pressure ulcers rates.

Explicit RRR was found to be a feasible method to collect data to evaluate improvement theme periods displayed in time series diagrams.

We found that it took a long time for new guidelines, quality indicators, and safety initiatives to be noticed and widely used in clinical practice. The results also showed that changes were moderate in the first years, and this suggested to us that for a sustainable improvement interventions on many levels were needed. The form and structure of applying improvement theme months may provide a possible way of organising quality and safety improvement initiatives in nursing.

## PAPER IV

In stage 1 of the review, 111 (HMPS method) and 132 (GTT) of the 350 records contained at least one positive screening criterion or trigger, and a potential AE. These were forwarded for physician review (Figure 2).

After the stage 2 review, the HMPS method team found AEs in 100 records (28.6%, 95% CI: 23.8-33.3), including 151 AEs that were due to healthcare, of which 131 (86.7%) were deemed preventable. Using the GTT, the physicians identified 99 AEs in 85 of the records (24.3%, 95% CI: 19.8-28.8) and 77 (77.8%) of these AEs were deemed preventable (Figure 2).

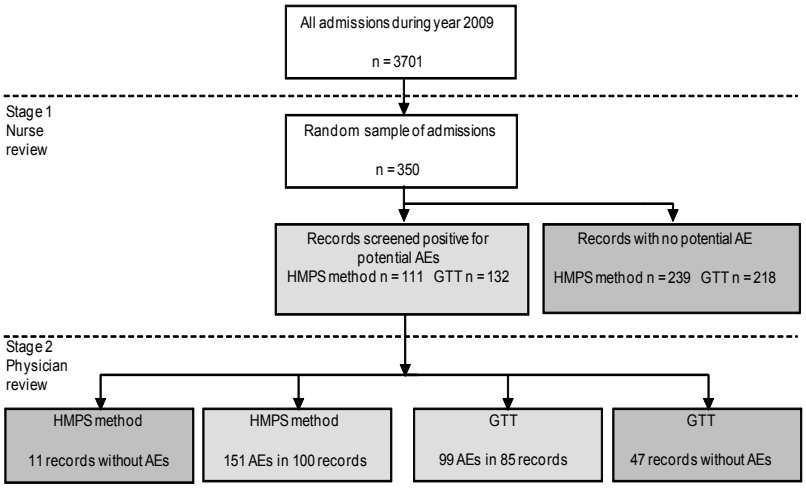


***Figure 2**. Two-stage review process for identifying adverse events*

In the two nurse-reviewer validation steps for GTT, seven respectively two previously unidentified AEs were included in the review process by the physicians. No previously unidentified AEs came to be later identified in one of either of the steps when using the HMPS method.

Four potential AEs not identified using the HMPS method, and 43 not identified by the GTT were identified in the crossover review process. After the stage 2 review and subsequent crossover review, a total of 155 and 137 AEs were identified in 104 (29.8%, 95% CI: 24.9-34.5) and 98 (28.0%, 95% CI: 23.3-32.7) of the 350 records by using the HMPS and GTT methods, respectively. Of these AEs, 135 (87.1%) and 110 (80.3%) were deemed preventable by the physicians. Overall, 160 different adverse events were identified in 105 records (30.0%, 95% CI: 25.2-34.8) using both methods.

To assess inter-rater reliability within the teams, 30 and 43 potential AEs were reviewed twice by physicians using the HMPS and GTT methods, respectively. The physicians' initial assessments in the twice-reviewed records were, before team discussions considering healthcare causation, coherent within the teams in 93% and 88% of the cases for the HMPS and GTT methods respectively and preventability in 100% and 95% of the cases, respectively.

The median nurse review time was significantly shorter when using the HMPS (3 minutes) compared to GTT (8 minutes) ($p<0.001$). The median review time for both physician reviewers using the HMPS method was six minutes. Median times for the team using GTT were four and eight minutes. No statistical difference was seen between the two methods´ physicians review times ($p=0.10$).

After crossover review, a total of 466 (range 1-11 per record) screening criteria and 737 (range 1-13 per record) triggers were identified in 195 and 233 records, respectively. The AE detection yield varied for each screening criterion and trigger. The percent of PPV was 40.3% (range 0.0-80.0) for the HMPS method and 30.4% (range 0.0-100.0) for the GTT ($p<0.001$). The PPV for no-harm incidents was, overall, 27.7% (0.0-82.8). When AEs and no-harm incidents were combined, the PPV for the HMPS method was 68.0% (range 0.0-92.0). The screening criteria with the highest PPV for AEs were "unplanned transfer from general care to intensive care", "unplanned removal, injury or repair of organ during surgery" and "hospital-incurred injury". The corresponding for the triggers were "post-operative increase in troponin level", "admission to intensive care post-operatively" and "clostridium difficile positive stool".

Most AEs resulted in minor, transient harm and the majority was judged to be preventable. The most common types of AEs were healthcare-associated infections and skin related.

The HMPS method found 118 no-harm incidents in 91 (26.0 %, 95% CI 21.4-30.6) of the 350 admissions, corresponding to on average 0.34 no-harm incidents per admission (range 0-3). Ninety-four of the 118 (79.7%) no-harm incidents were classified as preventable.

The most common types of no-harm incidents were found in patients not receiving critical drugs, for example, antibiotics or cortisone, or who had falls without harm or anaesthesia-related events.

# ETHICAL CONSIDERATIONS

Ethics approvals were obtained from the regional Ethics Committee of Stockholm for all studies. All studies were performed according to correct ethical practice.

The external reviewers signed confidentially agreements to maintain the security of the information.

All information about the patients was kept confidential during all stages of the study.

Permission was obtained to examine records from the Head of the Orthopaedic Department and the Chief Medical Officer at Danderyd Hospital.

Permission was obtained for the obtaining data from databases according to the Personal Data Act[149].

# DISCUSSION INCLUDING METHODOLOGICAL CONSIDERATIONS

The aim of this thesis was to evaluate the capability of retrospective record review methods for identification of AEs, and for follow-up review on local patient safety and quality information in orthopaedic care. In this section I will discuss my main findings including methodological considerations.

Papers I and II focused on comparing a RRR with the traditional incident reporting systems in Swedish healthcare, with respect to capability to identify AE occurrence and AE risk factors in orthopaedic care. Paper III evaluated a nursing improvement initiative. Paper IV compared the capability of two different RRR methods to detect AEs and their predictive value.

## COMPARISON OF SAMPLES

The studies included in this thesis are based on data from one of the largest orthopaedic departments in Sweden. However, the outcomes may differ from other orthopaedic departments and AE studies due to different case-mixes, review team composition with respect of medical specialities and reviewer performance ability based on skills and review experiences. Therefore, findings may not be generalised to other orthopaedic departments or AE studies. This will be further discussed in the section on validity and reliability. To the best of our knowledge, no other orthopaedic AE study is available for comparison with our results. Results illustrate the capability of RRR to identify and track orthopaedic quality and patient safety information; and RRR can be used to gain insight into the rates and nature of orthopaedic AEs.

Comparing the demographics of the study samples for papers I, II and IV, and the annual orthopaedic inpatient cohorts during 2004 and 2009, results show that the study samples are comparable to the whole inpatient cohort at the Orthopaedic Department (Table 7).

**Table 7.** Demographics of the study samples and the annual orthopaedic inpatients during 2004 and 2009

| Demographic variables | 2004 study sample n (%) | 2004 n (%) | 2009 study sample n (%) | 2009 n (%) |
|---|---|---|---|---|
| Total number of admissions | 395 | 2679 | 350 | 3701 |
| Men (%) | 153 (38.7) | 1035 (38.6) | 149 (42.6) | 1475 (39.9) |
| Age, median years (mean years) | 72 (66.1) | 72 (67) | 69 (66) | 70 (66.5) |
| Emergency admission (%) | 285 (72.2) | 1888 (70.5) | 250 (71.4) | 2810 (75.9) |
| Patients undergoing surgery (%) | 309 (78) | 2150 (80.3) | 270 (77.1) | 2795 (75.5) |
| Length of hospital stay, median days (mean days) | 5 (5.8) | 5 (5.5) | 4 (5.3) | 4 (5.4) |

In 2004, almost all of the AEs were found using record review rather than by using incident reporting systems that were in place. This may reflect the department's effectiveness in addressing safety at that time. Assuming that the findings from the two RRR accurately represent years 2004 and 2009, results can be extrapolated and compared with reported annual incidents by traditional systems. This extrapolation and comparison shows that considerable underreporting occurs five years later; and that RRR has better coverage (Table 8). An increase in reporting rates by personnel and by patients was seen when year 2004 was compared with 2009. Reporting rates may vary over time and an increase in the number of reported incidents may not necessarily reflect worse outcome. Instead, this may result from more structured efforts to highlight AE issues in departmental work since 2004, such as with increased awareness, openness and willingness by personnel, patients and their relatives to report. Between these years a 38% rise in inpatient admissions can be noticed.

**Table 8.** Comparison of papers I and IV of estimated annual AE rate using retrospective record review, outcomes from the traditionally incident reporting systems during the study period in paper I, and actual annual rates at the Orthopaedic Department

| Safety information methods | 2004 study sample n | 2004 total n | 2009 study sample n | 2009 total n |
|---|---|---|---|---|
| Retrospective record review | 65* | 441 | 160 (plus 118 no-harm incidents) | 1692 (2940 including both AEs and no-harm incidents) |
| Clinical incident reporting*‡ | 1 | 53 | na | 197 |
| County Councils´ Mutual Insurance Company‡ | 2 | 40§ | na | 80§ |
| The Medical Responsibility Board*‡ | 0 | 8§ | na | 9§ |
| Lex Maria*‡ | 0 | 4§ | na | 2§ |
| The Patient's Advisory Committee*‡ | 1 | 4§ | na | 21§ |

* May also include events without harm. ‡ The annual rates may include incidents related to outpatients. § Reporting year and this may not be the same year that the incident occurred. Na, not applicable due to no measurement of these variables in that study.

## METHODS FOR IDENTIFYING ADVERSE EVENTS

Papers I and II evaluated the capability of a modified Wimmera Clinical Risk Management Model to identify AEs at an orthopaedic department in Sweden. The model is well established at the rural Wimmera Base Hospital and uses multiple data collecting methods. The cost for the programme is estimated to be 0.1% of the total hospital budget[150]. Knowledge about the outcome and cost for this risk management model at larger hospitals with more specialist care is limited.

Paper I compared the AE rates between a RRR and the traditional incident reporting methods used in Sweden on the same cohort of 395 patients. The AE rate in the traditional reporting systems was found to be very low. Only two previously unrecognised AEs were found using these methods compared to 65 AEs found through the RRR. This finding is in line with other

studies where different AEs have been identified with different methods. Record review has been found to have better coverage compared to other methods[16 22 69 71-75 77 78 151].

Based on the results and the literature[16 22 23 54 66 71 75 77 95 100 103 107-109 151], it is likely that many incidents go unnoticed and/or unreported by healthcare providers, personnel, and patients. Lex Maria reports and patient malpractice claims tend to focus on more severe events as defined by law. This selective reporting, which can be seen in paper I, may lead to an incorrect and disparate picture of patient safety and the rate and nature of incidents.

It is essential to increase the reporting rates to achieve more valid and reliable data when using existing incident reporting systems. To be able to reduce the barriers of reporting requires a multilevel approach, including creating an open, safety culture. Some of the interventions may be made more useful if matched according to the different personnel's groups.

Most comparative studies, including those in this thesis, focus on comparing RRR with incident reporting systems, patient safety indicators and/or administrative/IT data within hospitals. Other methods, like observations, interviews, and surveys, may identify a higher number of AEs compared to RRR[98 108 152]. These methods may increase the understanding of contributing factors, both facilitating and hindering factors, Participation in both reporting and improvement efforts for patient safety by personnel and patients may bring better and more detailed data not otherwise available. These methods should be used more often to increase and deepen the patient safety knowledge.

## NATURE OF ORTHOPAEDIC ADVERSE EVENTS

The retrospective reviews presented in papers I and II found that 15% of patients had experienced an AE. In paper IV, the HMPS and GTT methods identified AEs in 30% and 28% of the records, respectively.

In the literature, the AE rate in orthopaedics is often included in total surgical AE rates, and is reported from as little as 2.7% to a high of 39%. Surgical AEs represent up to 66% of all AEs[27 45 54 66 67 108 136 153-156]. Studies report orthopaedics AE rates between 4.1% and 33%[9 26 45 123 124 154 157 158].

### Surgery as a risk factor for adverse events

While a surgical procedure increases AE risk in orthopaedic inpatients, we reported in paper II that most AEs did not relate to the skill of the surgeon nor the anaesthetist, but rather to the orthopaedic care processes outside the operating room. Griffen et al.[104] reviewed malpractice claims against general surgeons and found that deficiencies in care occurred more often pre- and post-operatively than during surgery itself. Olsen's[1] findings indicate that the majority of surgical patients' problems of care were related to their care on the wards. If the aim is to reduce the rate of AEs, then we must pay more attention to high risk patients during their entire care period. This calls for an emphasis on multidisciplinary and interdepartmental improvement that would focus on implementation of and adherence to evidence-based routines and procedures, and also on other aspects of care processes that can affect patient safety, such as, team training, standardised hand-offs, improved communication, and the development of a safety culture.

The nature and frequency of AEs appear to differ according to speciality and procedures utilised within different specialities[9 13 136 154 155 159 160]. Zegers et al.[161] found that AEs varied between hospitals and departments but that preventability only varied between departments. They argue that interventions to improve patient safety should, therefore, be tailored to specialities and local contexts. An error in one speciality or procedure may lead to a minor AE, whereas a similar error in another can lead to permanent disability[29]. This makes it difficult to compare some errors between specialties and their impact on patients and severity. Pukk-Härenstam et al.[94] found that surgical specialities accounted for 88% of patient claims to the County Councils´ Mutual Insurance Company, but for only 46% of hospital discharges. Orthopaedic and hand surgery care had the highest claims rate. The large number of claims in these specialities may be due to high expectations for positive outcomes among patient groups. Whether surgical AEs occur more frequently or simply more likely to be detected by patients and healthcare personnel is not well known[136]. Greater treatment complexity and invasiveness of care may increase the risk for an AE. In surgical care many such processes, whether pre-, peri- or post-operatively, can lead to AEs[29].

### Adverse events and length of hospital stay

In paper II we reported that patients experiencing an AE were found to have a longer hospital length of stay (LOS), which is consistent with others' findings[27 46 66 70 75 158] and may indicate that the orthopaedic AEs themselves prolonged LOS. It is possible that patient co-morbidity, case complexity and other reasons may have also prolonged LOS. Even admissions with mild AEs have been reported to affect the LOS[157 162]. Kable et al.[155] found that LOS was a significant predictor for the presence of an AE in surgical care. Prolonged LOS for patients with AEs not only reflects increased patient suffering but also carries economic consequences, including increased workload for healthcare personnel[44]. In an already stressed system this can lead to additional safety problems[1].

### Elderly and adverse events

We found in paper II that elderly patients were more predisposed to AEs during surgery and were affected by major AEs to a greater degree than younger patients. The latter finding is consistent with Zegers et al.[123]. Patients with ASA-classification three or above were also found to be more predisposed to AEs and high ASA score was also correlated to age. In contrast to many other studies[9 26 27 44 45 75 124 125 136 155 156 163], age was not a significant risk factor for AE in our study. This may be due to our smaller sample size or that the study population was much older in contrast to other AE studies that have presented age[26 44 45 75 136 160]. Elderly people are more likely to have more complicated diseases and underlying degenerative conditions that require more care activities. This increases the risk for an AE compared to younger patients. They are also more vulnerable to the effects of healthcare errors, for example, medication errors[29].

### Minor and major adverse events

In papers I and II, 31 of 65 AEs found by RRR were considered minor in contrast to those that were judged to be major according to the severity scale. This may have been due to focusing on medical AEs that might have led to unidentified minor AEs related to, for example, nursing care, although the severity scale used may have biased ratings. Most AEs we reported in paper IV were judged minor, which was possibly a consequence of including a more inclusive and broader range of AEs and the use of other severity scales. This is

consistent with other findings[12 26 27 45 159]. The severity scales used at Wimmera Base Hospital, HMPS methodology, and GTT are completely different which makes it difficult to compare severity outcomes from papers I, II and IV.

Applying a broader definition of an AE will shift some of the focus so as to prevent high-frequency, minor AE. The GTT AE definition may be better fitted to achieve this compared to the HMPS method's definition. Olsen[1] stated that even if an AE caused only minor physical hurt, it may still be to the detriment of patient psychological recovery, participation and the trust. Runciman et al.[164] argue that if interventions to improve patient safety are triggered only by the relatively more uncommon but severe outcomes, the more frequent, commonly occurring safety problems, which consume a great deal of resources, will go unaddressed. Interventions to reduce these minor AEs will positively affect a large number of patients and may be cost-effective.

## VALIDITY AND RELIABILITY

Comparing AE rates between RRR studies is associated with considerable difficulties. Data produced, regardless of methodology, may be biased by several factors, including: study setting, AE definitions, time frames for inclusion, case-mix, threshold for causation scale, documentation quality, representativeness of the sample and the fact that some studies only recorded one AE per patient[165]. Factors that may affect variations in judgements include reviewer skills, application of screening criteria/triggers and definitions, experiences and prevalent views of the reviewers, and training and education in the particular RRR methodology[1 23]. All these factors affect validity and reliability.

Having a digitalised record system enabled us to avoid missing records. However, in papers I and II, anaesthesia monitoring forms used during and after surgery, and the lists of medication in paper form, were not included, which may have led to some AEs being unidentified. We have had control of data quality and could refer to original records if something had been unclear and in order to validate the administrative data. We used review forms that had been used in other studies, with some modifications.

As with any retrospective record study, sufficient qualitative documentation is necessary to determine that an AE had occurred and to make judgements about preventability, severity and nature. This affects the outcome in each study in this thesis. Clinical incident reports, complaints and malpractice claims falling within the inclusion period in paper I were followed locally for three years. Because malpractice claims are reported up to ten years after occurrence a few claims may have been unnoticed.

Hindsight bias may, on the other hand, provide an overestimation of AEs in papers I, II and IV. Also, since elderly are more predisposed for AE and our study sampled adults of higher age compared to other studies would be expected to increase AE frequency[26 44 45 75 136 160].

Papers I and II review orthopaedic records from a summer month (August) during which time the proportion of trauma patients may have been increased while number of elective patients, especially arthroplasty cases, would be reduced. This may have influenced the pattern of observed AEs. Patients with symptoms related to other conditions acquired before index admissions were excluded. In our study for paper IV all admissions during

2009 were available for randomisation. In studies for papers I, II and IV we included only orthopaedic-related AEs. This may have decreased the observed overall AEs rate.

### Why do Wimmera Base Hospital and Danderyd Hospital outcome differ?

Our study at Danderyd Hospital for papers I and II reported a higher rate of AEs (17%) than was reported from the one of the Wimmera Base Hospital's studies (1.35%), in which nearly 50,000 patient records were screened over eight years[77]. This difference might perhaps be explained by the fact that Wimmera Base Hospital is located in a rural setting, in contrast to Danderyd Hospital which is a university hospital with probably more complex treatments and more risks for AEs. The Wimmera study also included several, different departments, not just surgical that often have higher AE rates, and used only eight explicit screening criteria, which did not require clinical judgement by first-stage review personnel. At Wimmera Base Hospital they later added the screening criterion, "any medical record recommended for review," but the outcome on the AE rate was not presented. In the study for papers I and II, all healthcare personnel notes in the records were examined for AEs. However, it was not so well described which notes were similarly examined in the Wimmera review process. We used a three-stage record review as compared to the two-stage process at Wimmera Base Hospital, which may have affected the outcomes since the reviewers identified a greater number of AEs after the consensus stage, which Wimmera does not have, as compared to stage 2. We obtained follow up data (county council-wide database) on patients but only four of a possible 65 AE were identified in this database, the rest having been identified either at Danderyd Hospital inpatient admissions or in outpatient visits. None of the 42 requested records from the general practitioner contained any AE, nor did the two from general practitioner feedbacks. The workload needed to identify these four additional AEs should be a topic for discussion.

### Rate of adverse event and preventability

One might wonder why the AE rate nearly doubled, as reported in paper IV compared to reported figures in papers I and II. Our papers I and II reported study aim to examine mainly medical AEs. This may have led to unidentified minor AEs, such as infiltrated intravenous infusions as we reported in paper IV. During the years between the studies, on the basis of what we learned from the RRR in organisations and among those engaged as reviewers, learning was achieved as to what constitutes an AE. Since the first study was performed, patient safety awareness has improved.

In papers 1, II and IV, we reported higher rates of AE in surgery than are reported in some surgical AE studies[136 153 154]. However, our AE rates are more in line with reports on other studies, for example, GTT studies[36 45 51 52 54 65-67 155 156]. The number of AEs that we found was high, and may be due to our broad AE definition that did not require that the patient should have experienced any disability or prolonged hospital stay as a result of the AE. This procedure was in contrast to other HMPS method studies that excluded minor AEs that did not fulfil the criteria for AE inclusion. These studies also often include only the most severe AE, while all AEs are included in ours, as do both Wimmera Clinical Risk Management Model and GTT. In contrast to GTT, AEs due to acts of omissions were included.

Our finding, presented in papers I and II, that about half of AEs are preventable, is consistent with prior surgical studies[136 154 155]. The preventability rates were, however, considerably higher in paper IV, and is more in line with the Swedish national AE study that used reviewers trained in the patient safety perspective. GTT studies seldom measure preventability in accordance to the origin method and, if preventability has been measured[51 52 68], other scales than the one used in the HMPS method have been used, making comparison difficult.

The high AE and preventability rates in paper IV may be due to having experienced reviewers. Other studies have recruited reviewers from the hospitals but their skills in reviewing records and safety perspective are seldom reported. Preventability may also change over time. The preventability rates we report in papers I, II and IV emphasise the need for more preventive initiatives in orthopaedic care to reduce the AEs rates. Sharek et al.[52] found that an experienced review team identified 30.2% AE; whereas, newly trained internal and external teams identified only 18.8 % respectively 15.3% AEs.

*Inter-rater reliability*

The kappa values indicating reliability of reviewer judgement reported in paper I were in line with other studies[26 27 44 46]. The levels of agreement between physician reviewers were better or equal in Paper IV to those found in other studies[9 13 26 46 52 53 70 123]. This may be an effect of using detailed manuals to support judgments, identifying expert physician reviewers, engaging reviewers in pre-study training sessions as well as later discussions; and that all reviewers were participated in study design and writing of study manuals. The number of reviewers may also affect the agreement rate. In papers I and II, the reviewers were not experts, did not have previous specific training in patient safety perspective, a Swedish language protocol was unavailable, and limited patient safety terms and definitions were available. This was also the first AE study conducted in Sweden and, consequently, no methodological support was available within the country. These factors may have influenced inter-rater reliability rating causation, preventability and severity; and that would affect the AE rate.

Lower inter-rater reliability has been reported for preventability than for AE-causation in some studies[26 46 70 166] suggesting preventability may be more difficult to judge, perhaps because it is more subjective[23]. Poor quality of documentation can contribute to poor agreement. Even if the studies are well-planned, definitions may not be adequately clear. A manual cannot describe all conceivable AEs because situational and individual factors must be applied in the review. Implicit review, such as is described in papers I, II and IV, may be more affected by subjectivity that explicit review, the latter which is addressed in paper III. In paper III, the first PPM months were reviewed together by the two senior orthopaedic RNs in order to encourage consistent identification of explicit variables.

Comparisons among five experienced GTT teams resulted in large differences in the number of identified AEs, as well as the assessments of preventability and severity. Only three out of 42 AEs were detected by all five teams[65]. Classen et al.[53] showed that inter-reviewer agreement could be improved by a single training session later followed by a two-hour formal training session.

### Comparison of the HMPS method and the GTT

Debate continues as to which RRR method is the most valid, reliable, cost efficient and feasible during the years. Comparison between HMPS and GTT methods is difficult due to differences in presentation of the rates, definitions and inclusion frames. To our knowledge, this has been the first study that compares the capability of HMPS and GTT methods in AE detection using the same cohort while using the same definitions. After the review stage 2, 52.5% more AEs were identified with the HMPS method than with the GTT. After crossover review 160 different AEs were found in 105 records with both methods together, and 155 and 137 of those were identified with HMPS method and GTT, respectively.

One could argue that the differences reported in paper IV for AE rate after review stage 2 may mostly be dependent review time in stage 1. The nurse review time is less using the HMPS method than for the GTT. The result of the GTT review time in stage 1 is similar to the findings in another Swedish study[65]. The GTT strategy to search for triggers, instead of reading the record comprehensively, may lower sensitivity. Searching for many triggers in different parts of the record may take longer than just comprehensively reading the text and searching for broad screening criteria in orthopaedic care comprising relatively short hospital stays and limited documentation. Reviewer familiarity with the specific record system and context may have affected the numbers of identified potential AEs in review stage 1. Two out of three members of the HMPS method team were familiar with the record system and local routines of the Department, as opposed to none of the GTT team.

After review and confirmation of AEs after stage 2, it was apparent that the HMPS team omitted four AEs and the GTT team had 43 unidentified AEs. We separately analysed the AEs identified by the HMPS method team but not those identified by the GTT team after stage 2. We excluded AEs that already had been rejected as AEs by the GTT team but were confirmed AEs by the HMPS method team, thus the 43 AEs noted above. Five of these 43 AEs identified by the HMPS team were subsequently rejected as AEs in the crossover review by the GTT team, five AEs were judged to result from lack of context knowledge; and four appeared to go unidentified as a result of the nurse reviewer lack of familiarity with the digitalised record system. Twenty-seven AEs were identified by the specific triggers, “procedure” and “care: other”; and 27 AEs were classified as temporary harm requiring intervention. There were no differences in type or rate of unidentified AEs, irrespective of whether assessment was performed early or late in the review process. The largest difference after review stage 2 was found among AEs causing minor harm. Our qualitative analyse indicated that knowledge of orthopaedic speciality or the digitalised record system may have been related to nine of 38 unidentified but later confirmed AEs by the GTT team after review stage 2.

The boundary between a no-harm incident and minor AE is not sharp and is subject to individual judgement, which may affect outcomes for any participant in a review process, irrespectively how thoroughly the study has been designed. The discrepancy in AEs identified between review stage 2 and the crossover review could be due to differences in the two review methods and the fact that this study included less severe AEs, which according to the GTT method, at least as usually interpreted before study start by our expert reviewers, were not considered AEs. Less severe AEs included infiltrated intravenous infusions and minor skin abrasions. Finding a difference in AE rate after the crossover review suggests that the

physicians in the GTT team were more likely to reject minor events as AEs than those in the HMPS method team. Classen et al.[53] have found that the greatest variability between the reviewers in data categorising severity of AEs related to the lowest harm level in the severity scale, that is, category E. Furthermore, the severity scale used in the GTT[25] required that an intervention had occurred to qualify as a minor AE (category E) which could have affected the AE rate and inter-rater reliability outcome. The severity scale used in the HMPS method is, by contrast, more inclusive of minor AEs[9 13]. The perception of minor AEs affects the review outcome but also, subsequently, the organisation's input regarding safety learning as discussed earlier.

### Positive predictive value

Paper IV reported that individual screening criteria (also paper I) and triggers varied in their yield of identified AEs: Some were always associated with AEs, while others never were. Some of the screening criteria and triggers are irrelevant for orthopaedic care (paper IV) and some were never identified in our samples. This is also our experience when using GTT in clinical patient safety work at different hospitals. The PPV may have been affected by the larger number of triggers as compared to numbers of screening criteria. None of the AEs found indicated a need for new screening criteria or triggers. However, the non-specific screening criterion "any other undesirable outcome not covered above" and triggers "procedure" and "care: other" were common and may necessitate more descriptions with examples in the manual to flag for AEs and facilitate the nurse review in order to create a more valid and reliable review process. A limitation with the screening criteria used in the HMPS method is that all screening criteria except three must occur during index admission, leading to that all readmissions, irrespectively of cause, are categorised into the two readmission screening criteria. Thus, screening criteria are indefinite about the type of AE that affected the patients outside index admissions. This may on the other hand have affected the total PPV positively. The total PPV reported in paper I was 48%, as compared to the HMPS and GTT method results reported in paper IV that showed 40% and 30%, respectively. This may be due to more explicit screening criteria. Naessens et al.[36] indicate that objective, explicit triggers may lead to higher agreement but it is not feasible to create explicit screening criteria or triggers for every conceivable AE[29]. For a skilled nurse reviewer the broad screening criteria used in HMPS method may be easier and quicker to use than searching for many triggers in surgical specialities.

## FEASIBILITY OF RETROSPECTIVE RECORD REVIEW

Record review has been criticised for being time consuming and expensive, does not provide real-time information, and that it is not possible to collect additional information not documented in the records[18 78]. In the development of GTT some of these limitations were considered. Olsen found[1] that by addressing some characteristics, including: experience, refinement and a structured education and training program, local reviewers can be used rather inexpensively to systematically produce detailed data close to the discharge of the patient. Prospective record review, combined with review close to discharge can enable personnel to gain additional information not documented in the records, to investigate active failures and latent conditions and control bias[1 41].

Local teams may collect the additional information mentioned above even when the review is conducted retrospectively, within 30 days of discharge. Real-time data is, of course, ideal

from a learning perspective. The results reported in paper IV show that the review time for the experienced RNs and physicians is acceptable in relation to outcome. There are few studies that have collected data about review time and our review times are shorter than most others. The length of hospital stays may affect the review time in stage 1, making comparisons difficult[1 59 65 150 167]. Classen et al.[53] discuss that if RRR can identify more AEs compared to traditional incident reporting methods the time used for review may be cost-effective. I will point out that one should not forget that the clinical incident report process also takes time, including interruptions in work flow and writing the reports. Those written reports should also be analysed and categorised in order to lead to meaningful learning. Using a RRR process at the departmental level will be further discussed in the section on record review below.

Retrospective record review has been proven to be just as effective as prospective collected data in identifying AEs, though more preventable AEs were identified with the latter method overall. However, in surgery the same rate of preventable AEs was found[69]. Vincent[129] states that there is no sharp division between retrospective and prospective techniques because the user uses previous knowledge when applying the prospective method.

## FEEDBACK OF PATIENT SAFETY AND QUALITY INFORMATION

Explicit record review is common when performing follow-ups and when monitoring progress in quality and patient safety. The performance data, as process and outcomes results, were identified using explicit RRR as reported in paper III. A strength in that paper was the use of longitudinal performance data displayed in time series diagrams over nearly four years, making it possible to analyse variation in performance data over time prior to, during and after interventions. The time series may provide a pictorial representation of change that may support real-time action based on what results are obtained on process and outcome[168 169]. Time series are also used in GTT to display and track the AE rate over time. To regularly collect and display performance data over time is an important part of QI initiatives to help create a learning-oriented culture conducive to systematic and continuous improvement. Healthcare personnel need to receive feedback on their performance in order to change practice, increase adherence, and improve quality and safety[30 170-173]. One specific lesson learned and then discussed in paper III was how to use more systematic, timely information feedback earlier in the improvement initiatives to support and catalyse change and to provide evidence over time of the impact of changes.

In follow-up sessions in which performance data were reported and discussed, nursing personnel reported they found that PPM was relatively easy to perform in daily practice. In my opinion, all types of measurements that personnel perform have educational elements. Therefore, it is important to understand the contributing factors behind the results. This information may help motivate and engage the nursing personnel toward further improvements.

Results of the mandatory annual PPM differed from longitudinal data on MNS outcome, and the pressure ulcer rate in paper III. This can give an inaccurate picture of the quality and safety of care both to local leaders and to decision makers. Reliance on one measurement per year could produce misleading. Such hasty conclusions about quality and safety in healthcare could misdirect quality and safety improvement efforts.

Decision-makers at different levels in the healthcare organisation need feedback data on performance to influence local, regional and national patient safety initiatives[41 173]. A different kind of tailored feedback data may be needed on different levels in the organisation. Feedback data may need to be more detailed to the personnel at a department as they have a different role in creating quality and safety compared to decision makers higher up in the organisation who often rely on aggregated data[173]. Benn et al.[174] describes 15 requirements for the design of effective feedback systems, including: creditability and content of information, capacity for rapid action, and role of leadership. Ideally, the feedback, learning and actions should occur on multiple levels. Specific quality problems, risks and incidents can be reduced at the local level by tailored interventions that appreciate the local context. Hospitals can reduce risks and the number of incidents while improving quality and safety; and this can involve several units. The broader healthcare system must be responsive to quality problems, risks and incidents, particularly, when multiple healthcare organisations are involved, and where a single organisations may have diffulty solving the problem[18 35].

## IMPROVEMENT OF PATIENT SAFETY AND QUALITY

Paper III described the locally developed improvement theme months, which were aimed to improve quality and safety within orthopaedic nursing care. The improvement theme months used a modified bottom-up approach that included defined objectives, easy-to-use follow-up measurement, education, changes to daily routines, "reminder months" and feedback on data. The improvement initiative and subject was approved by head nurses, but the content and ideas were generated from the bottom up, in the context of empowered QI teams, who were supported by a local facilitator. The initiative was in part inspired by findings from nursing data collected in the first study. Using the "Breakthrough model", healthcare organisations and their respective teams can close the gap between best and current practice, and contribute to continuous improvement work[175]. In Plan-Do-Study-Act cycles[176], teams used learning gained from prior improvement theme months to improve the next one. A wide range of barriers to change can affect the outcome of implemented interventions. They can be related to the individual personnel, the social and organisational context, the patient and the financial system. Assessment and understanding of barriers and strategies to promote change may contribute to a successful implementation of evidence into practice. Barriers and strengths were discussed in each improvement theme months. The use of a facilitator may also be an important strategy to provide support to personnel to implement change[177]. In our retrospective analyse we identified five general QI principles that have been used in the improvement theme months and that may have affected the result positively: creating urgency; leadership support; motivation, involvement and commitment; audit and feedback, and a plan for creating progress and sustainability.

Main findings in paper III included that implementation of improvement theme months was temporally and logically associated with significant improvement over time in both process compliance to MNS and patient outcome regarding pressure ulcer rates. Changes, overall, were more variable the first two years. Findings indicate that sustained improvement requires multilevel efforts with multiple components, including: education, contextual knowledge, motivated and engaged personnel, achievable goals, measurable outcome, structured and timely feedback, reinforcement, supportive organisational culture, and leadership support. Such things take time. We also found that it took long time for new guidelines, quality indicators, and safety initiatives to be noticed and become

widely used, not only among personnel as just mentioned but also among the healthcare leaders and the hospital management team. The reasons for this slowness are not clear and require further investigations to increase the implementation speed in the future. There is, unfortunately, no given "best way" to implement and improve quality and safety[145]. Synthesis across studies can help organisations to create an understanding of strategies to support improvement[178]. Our improvement themes months may be a learning example of how to work with quality and patient safety including empowering of the nursing personnel in these issues.

Paper III has some limitations that must be considered when interpreting the results. Some of the data was collected as part of the hospital's annual follow up measurements in relation to the County Council's quality indicators. Only simple graphical analyses of the annual PPM data have been possible to carry out due to poor data quality. This was also the reason that it was not possible to separate data concerning the Orthopaedic Department from the hospital's data, which would have allowed comparison between our department and all other departments.

Improvement theme months were a product of a local QI project. Research efforts came later, which has limited our ability to collect some data that could have been identified if the study has been performed prospectively. Interviews with head nurses and nursing personnel could have further increased our understanding of their perspectives about the QI project and implementation successes and barriers in relation to internal and external contextual factors.

However, attribution is difficult. Even if we can demonstrate a temporal connection between improvement theme months and desirable changes in performance, we cannot rule out the possibility that other circumstances caused those changes, at least in part, due to the internal and external contextual factors as displayed in the time series diagrams. Improvement theme months were also multifaceted and it difficult to distinguish the effect of any single intervention. After initiating improvement theme months, the performance of the Orthopaedic Department, as captured by the County Council's annual PPM, was better than the performance of the hospital overall. This suggests that improvement theme months had an impact on quality and safety.

# Conclusions

The overall findings and conclusions from the four papers in this thesis will briefly be summarised as follows:

- Retrospective record review may have wide coverage in capturing orthopaedic adverse events at a local level and may play a vital role in the quality and safety information system in order to identify, categorise and analyse quality and safety problems and to subsequently provide a basis for interventions.

- Due to variations in data quality and coverage, several different data collecting methods need to be used concurrently in order to obtain a comprehensive picture of deficiencies in healthcare.

- When using different methods to identify patient safety information the respective methods´ advantages, limitations and applications ought to be addressed.

- Increased awareness, consideration of risk factors, interventions focused on multidisciplinary and interdepartmental teamwork, and strategies that focus on healthcare processes may reduce the frequency of adverse events in orthopaedic care.

- Retrospective record review may be favourably used to monitor the outcomes of quality and patient safety interventions over time, preferably displayed in time series diagrams in order to drive change.

- Improvement theme months may provide a way to organise quality and safety improvement initiatives in nursing. However, we found that it took a long time for new guidelines, quality indicators, and safety initiatives to be noticed and widely used in clinical practice. Changes were moderate during the first years. To achieve sustainable improvement, interventions on many levels of the organisation were needed.

# PRACTICAL APPLICATIONS

## NO-HARM INCIDENTS

In the comparative study between HMPS and GTT methods, the HMPS method also addressed no-harm incidents (to be published separately). The HMPS method found 118 no-harm incidents in 91 (26.0 %) of the 350 admissions. Ninety-four (79.7%) of the 118 no-harm incidents were judged to be preventable.

We found that no-harm incidents were common, and fell mainly into three areas: important medication, such as prophylactic antibiotics, that were not being given; falls without harm; and anaesthesia-related events. All such cases point to systemic failures that put patients at risk of suffering an AE, and are an important signal to the personnel around the patient as well as local leadership that routines are failing to protect patients. If attention is given only to harm then important safety problems may be missed. Drug failures were also among findings presented in paper II, as were healthcare-associated infections in paper II and IV. These may illustrate the healthcare organisation's difficulty in learning from and acting upon patient safety information.

There are few studies that have used RRR, including both AEs and no-harm incidents, and then compared the outcomes in relation to other information methods[71 151]. Our no-harm incident rate comprised from 9.5% of the total annual inpatient population composed 60% of the annual clinical incident reporting rate as also included harm (Table 8). This result is consistent with that of others'[71 151] that show RRR identifies safety information about no-harm incidents in a structured way by using random samples and implicit review. This probably makes the outcome more valid and reliable than only using incident reporting systems, which are not systematic data collecting methods. The requirements of the Patient Safety Act[82] are that healthcare organisations and personnel must work more proactively on safety issues. If no-harm incidents are included when using RRR to detect AEs in the hospitals, as Swedish hospitals are obliged to do starting in 2012, the validity of no-harm incident data, learning, patient safety outcomes and the adherence to the Act may increase.

## RECORD REVIEW ON DEPARTMENTAL LEVEL

My viewpoint regarding the capability of RRR for patient safety work on a departmental level is based on the experience of others, as well as my own, along with findings from our various research studies, and practical work as a project leader for the implantation of the GTT at Danderyd Hospital.

To assess the AE rate at the department level and to develop the review process is important in increasing departmental safety learning. By analysing and categorising the nature of AEs during the review process according to e.g. scales used in the HMPS method, preferably with the help of an electronic database, structured enhanced learning can be achieved. GTT traditionally collects limited information about the nature of AEs. This safety learning can be used to guide local limited improvement resources into specific areas and/or processes

where tailored interventions and redesign are necessary to create resilience[51 161]. The use of an electronic database has several advantages: reduced risk of data loss, ensures data quality, guides reviewers in the review process and facilitates rapid statistical data and presentation[1 51]. Irrespective of method applied, it is important in clinical patient safety work to have stable internal review teams that can proceed to expertise and who produce consistent reviews, if trends are to be detected in less review time[52].

If every department has its own review team, the hospital can increase the number of personnel with both methodological and contextual knowledge and may help it achieve an increased local involvement in patient safety issues. Education and training, carefully designed user manuals, good methodological support and follow up may be success factors that increase the validity and reliability of the review process. Locally collected data using different methods for the identification of patient safety information should be used internally for improvement, though not for comparison between departments or organisations due to the validity and reliability factors mentioned earlier. However, it is valuable if different departments make the same judgement of comparable types of events to obtain valid data within the department and the hospital.

To learn from a wide range of deficiencies no-harm incidents may be included as discussed earlier. It is also important to have an inclusive definition and severity scale, to measure the nature of the incidents and have predictive screening criteria/triggers. To support local learning the steps in Figure 3, as will be described in next section, can be used. Data collected and analysed locally can be perceived as more credible and can therefore represent a better incentive for improvements. Local record reviews may lead to that the safety information data that forms the basis for fact-based and evidence-based interventions accumulates faster. The outcomes from all departments can also be aggregated and analysed at the hospital level in order to identify trends and patterns that may not be seen at the departmental level. By performing reviews on the departmental level information from the records can be supplemented with other information sources and may provide a more comprehensive picture of rate, nature and contributing factors of incidents which compensate for any lack of documentation in the records[1 69 71].

Safety discussions within the teams are important but my experiences is that the physicians in the team ought to do an independent review of the records that are forwarded from the reviewers in stage 1 in order to achieve a wider perspective in the judgements of the nature of incidents and increase the physician's safety learning. Incidents related to nursing may be reviewed in both stages of a RN who probably can make a better or equal judgement than a physician reviewer in these issues. Record review can also favourable be used in the follow ups of interventions and quality indicators. While having random samples every month or every second week these quality follow ups can be performed at the same time as the retrospective safety record reviews and therefore be more cost efficient than to perform separately reviews for these purposes.

## QUALITY AND SAFETY INFORMATION SYSTEMS

Multiple methods need to be used concurrent when collecting safety and quality information. Despite limitations various, single methods may have, combining them can serve to add qualitative and quantitative information needed to give a more accurate picture and a greater

understanding of incidents, their nature and contributing factors. Knowledge can be gained from a wide range of incidents[18 41 69 74 179]. The data collected must however have adequate validity, reliability and coverage. When choosing data collection methods consideration must be taken to the addressed problem, different methods´ strengths and limitations, the context in which the quality and safety problem occur and a balance of different constraints e.g. cost, work load, usefulness and expected outcome[1].

I have made an overview of the continuous improvement cycle in Figure 3 that illustrates how the quality and safety information system is linked at different steps. Some requirements for the respective steps are mentioned.

Information, per se, does not make any care safer. To achieve learning it also requires that the receiver is open to receive the information, to learn from it and to use the data in the continuous improvement work. The collected data must be reviewed, categorised and later analysed, often as aggregated data, to see patterns and trends and to learn from contributing factors. It would be of value if the quality and safety information systems in Sweden could develop and use the same terms and categorisations to facilitate aggregated analyses and learning in order to disseminate knowledge and inspire for actions to improve quality and safety. Healthcare organisations must be aware of that there is a risk if only rely on data with low quality and coverage as this may lead to decisions based on wrong interpretations and that the organisations cannot adequately measure if the care have been safer after implemented interventions.

More improvement areas will be found than can be managed at one time and by using standardised scales and tools prioritisations can be made. Events that are of low prioritisation must be further monitored[143]. The results from the former steps in the cycle form the basis to identify and select evidence-based interventions with the most impact on safety and quality outcome. Objectives and goals must be set in relation to the interventions. The goals ought to be specific, measurable, accepted, reasonable and timed (SMART)[180].

It is a challenge to improve quality and patient safety[33 181]. To increase quality and patient safety interventions outcome there is a need for bridging the gap between theory and practice in improvement efforts by combining professional and improvement knowledge[182 183]. Personnel and healthcare leaders need knowledge about how to apply and disseminate evidence in order to ensure that patients receive the intended evidence-based care[172 178].

In every step of the cycle reflections and evaluation also are needed. The outcome feedback must be timely and useful, and it should be delivered through convenient and effective channels to the appropriate levels in the organisation. The feedback loop on safety and quality issues is a continuously process rather than only temporary to effectively close the safety and quality feedback loop with help of timely corrective interventions[174] as described in Figure 3. To evaluate the outcome of the implemented inventions the cycle starts over again.

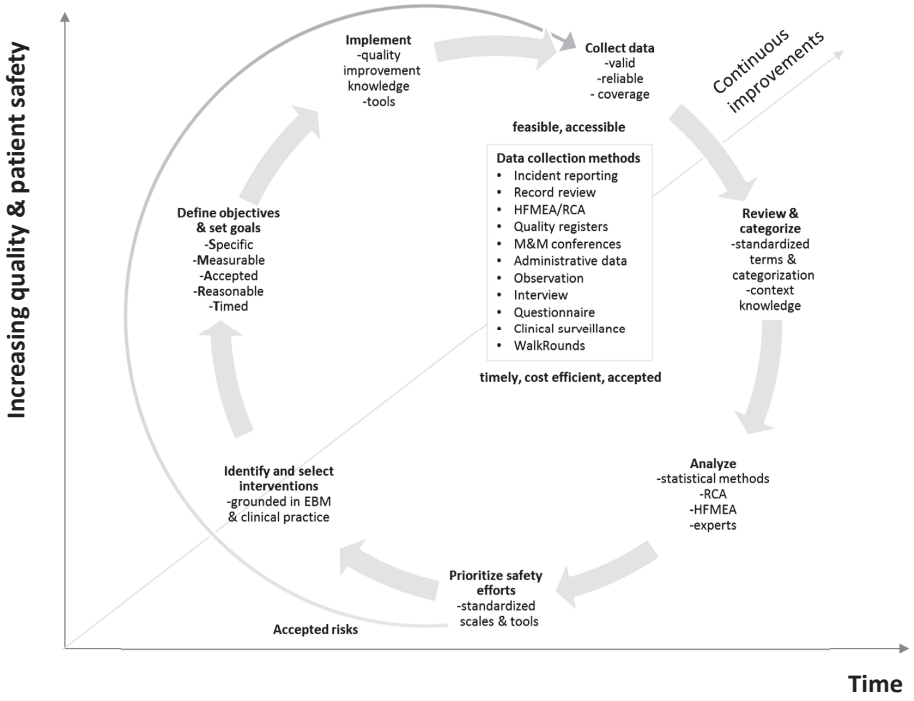


***Figure 3.*** *Quality and safety information system within the continuous quality and safety improvement cycle*

Being aware of the data quality problem and analysing and drawing conclusions carefully is important.

Data quality problems must not be an excuse for healthcare organisations to dismiss continuous quality and safety improvement work. Safety initiatives must be approached in parallel with developing better methods for measuring outcomes related to patient safety and quality. More research is necessary to develop and compare quality and safety information methods.

To radically reduce system failures and reduce the impact of risks and failures, multilevel interventions targeting both processes, work practice and safety culture are needed.

# FUTURE RESEARCH

Further development and modification of the record review methods based on both theoretical and practical experiences is important to determine the "optimal record review method", including combinations of screening criteria respectively triggers, systematic evaluation of manuals, education, implementation, method support, follow up and cost efficiency in relation to outcome. This may also include the validation of the use of information technology to identify AEs.

Further research in quality and safety deficiencies is also needed concerning validity, reliability and cost efficiency in relation to information richness outcome concerning the methods being considered. Research may also be performed to evaluate which combinations of methods are best suited to identify different types of quality and safety information, how this favourable should be analysed and categorised, being effectively fed back, and being used to prioritise, provide and inform the design and evaluation of improvement strategies and interventions.

Other important questions for future research are: how to further develop and implement a common terminology, definitions and classification in Swedish quality and incidents reporting systems to enable a systematic comparison and learning from different methods and their generated aggregated data and, how to involve patients and their relatives in a higher grade in order to identify quality problems, risks and incidents and use these data in the continuous QI.

# Populärvetenskaplig sammanfattning

Hälso- och sjukvård är i huvudsak till stor nytta, men är också en verksamhet fylld med risker som kan leda till att patienter kommer till skada. Kunskapen om att hälso- och sjukvården kan skada patienterna är inte ny inte heller intresset för dessa frågor. Det som är nytt är att det under det senaste decenniet pågått en allmän debatt om riskerna i vården. Detta har lett till en ökad insikt om behovet av att mäta och förstå sjukvårdsorsakade skador och av att förbättra säkerheten.

The Harvard Medical Practice Study (HMPS) som publicerades 1991 beskrivs ofta som den första avgörande retrospektiva journalgranskningsstudien för att identifiera skador orsakade av hälso- och sjukvården. Studien fann att skador inträffade vid 3.7% av alla vårdtillfällen. Det var dock först 1999 när the Institute of Medicines rapport *To Err is Human* kom som dessa resultat fick ett ordentligt genomslag följt av en massiv offentlig debatt. I rapporten angavs att skador orsakade av hälso- och sjukvården bidrog till att mellan 44 000 och 98 000 amerikanska patienter dog årligen, vilket motsvarade tre kraschade stora passagerarflygplan varannan dag! Rapporten föreslog även åtgärder för hur patientsäkerheten skulle kunna förbättras på olika nivåer inom hälso- och sjukvården. En del anser att den moderna patientsäkerhetsrörelsen började i och med denna rapport då den blev en varningsklocka även internationellt. Ett flertal länder började genomföra egna skadestudier. Dessa studier fann att skador uppstod vid upp till 16.6% av alla vårdtillfällen på sjukhus och att drygt hälften av dessa skador bedömdes vara undvikbara.

Hösten 2007 genomförde Socialstyrelsen en studie som baserades på samma protokoll som HMPS. Ett representativt urval omfattande 1967 vårdtillfällen granskades och man fann att skador inträffat vid 12.3% av alla granskade vårdtillfällen. Av dessa bedömdes 70% vara undvikbara dvs. vara vårdskador. Vid extrapolering av resulatet till alla vårdtillfällen under ett år motsvarar detta cirka 105 000 vårdskador vilket bl.a. resulterar i cirka 630 000 extra vårddygn.

Modern säkerhetsforskning definierar säkerhet i system inte som avsaknad av risker eller misstag utan i termer av systemens förmåga att kunna hantera sårbarheter eller risker så att dessa inte leder till skada för t.ex. personal, utrustning eller patienter. Andra högriskbranscher som t.ex. flyget och kärnkraftsindustrin har kommit längre i att utveckla system för att följa upp och förbättra säkerheten. Erfarenheter därifrån pekar på att organisationer behöver informationssystem för att samla in, analysera och återföra information om såväl risker som avvikelser i verksamheten för att lära av dessa för att förbättra säkerheten. I hälso- och sjukvården finns det inte något väl utvecklat system för mätning och uppföljning av patientsäkerheten på ett tillförlitligt sätt. Även förmågan att lära av inträffade avvikelser och att sprida säkerhetsinformation har brister.

En avsevärd mängd patientsäkerhetsinformation samlas in på olika nivåer inom hälso- och sjukvården men informationen är spridd i olika icke integrerade system och är kategoriserad

på olika sätt vilket försvårar en systematisk analys. Avvikelser kan identifieras på många olika sätt. Vissa metoder används rutinmässigt i hälso- och sjukvården medan andra mestadels har använts i forskningssammanhang. En del av dessa metoder kan användas för att räkna ut incidens och andra är mer användbara för att identifiera bakomliggande dolda fel i systemen. Mätningar är viktiga för att identifiera de vanligaste riskerna och avvikelserna som kan utgöra underlag till förbättringar och för att följa upp effekterna.

Med proaktiva informationskällor, t.ex. riskanalyser,kan man identifiera risker innan en skada sker. De befintliga rapporteringssystemen, där både vårdgivare och personal samt patienter och anhöriga kan rapportera avvikelser är dock ofta reaktiva. Rapporteringsystem kan vara lokala, regionala eller nationella. Alla dessa rapporteringssystem kan ge kunskap om brister i patientsäkerheten. Flera av dessa informationskällor har dock långa rapporteringstider vilket påverkar möjligheterna att snabbt återföra aktuella erfarenheter till verksamheten. En del av dessa rapporteringssystem fokuserar mer på allvarliga skador vilket kan ge en skev bild av avvikelsernas natur och frekvens. Dessutom saknas ofta ändamålsenliga system för återkoppling av analyser till beslutfattare på olika nivåer i verksamheterna samt till personalen. Det krävs även att mottagaren är öppen för att ta emot informationen, lära av den och använda underlaget i det systematiska förbättringsarbetet.

Studier har visat att det finns en betydande underrapportering vad det gäller de olika befintliga rapporteringssystemen. Långt ifrån alla avvikelser upptäcks och av dem som upptäcks rapporteras enbart ett fåtal varav de flesta inte är skador. Rädsla för rättsliga påföljder, att personalen inte vet vad som ska rapporteras eller hur, bristande återkoppling till rapportörer och verksamheten påverkar rapporteringsfrekvensen. Detta i kombination med en kultur inom sjukvården där uppfattningen att professionella inte gör fel kan bidra till att avvikelser inte rapporteras eller diskuteras öppet. Detta kan leda till att möjligheter till lärande missas och leder även till svårigheter i få en korrekt bild av hur ofta avvikelser inträffar och hur säker vården är.

Prospektiva uppföljningar, direkta kliniska observationer, enkäter och struktuerade journalgranskningar har mestadels använts i forskningsamamanhang och är mer anpassade för att ange prevalens och i vissa fall incidens gällande avvikelser samt för att utvärdera effekten av patientsäkerhetarbetet. De flesta av dessa är resusintensiva och kan vara svåra att använda kontinuerligt i det kliniska arbetet.

Det är viktigt att vara medveten om de olika metodernas styrkor och svagheter vid analys av data.

Det övergripande syftet för denna avhandling var att utvärdera retrospektiva journalgranskningsmetoders förmåga att identifiera patientssäkerhets- och kvalitetsinformation i ortopedisk vård.

I avhandlingen ingick fyra artiklar baserade på tre studier.

Kliniskt patientsäkerhetarbete handlar om att identifiera och förebygga händelser i patientvården som kan resultera i skada. Studier visar att frekvenser och typer av avvikelser varierar mellan olika medicinska specialiteter. Det är av värde att ha kunskap om de olika specia-

liteternas risksituationer och avvikelser för att kunna bedriva ett effektivt patientsäkerhetsarbete. I de tidigare gjorda internationella journalgranskningsstudierna finns få ortopediska data avseende skador tillgängliga eftersom de oftast redovisas ihop med övriga kirurgiska skador. Hur journalgranskningsdata relaterade till andra svenska avvikelserapporteringssystem gällande ortopedi var okänt.

I den första studien som presenterades i artikel I och II granskades 395 patienters digitaliserade journaler i tre steg med hjälp av 12 kriterier och standardiserade bedömningsformulär. Metoden utgick ifrån the Wimmera Clinical Risk Management Model. Det genomfördes också en genomlysning av de traditionella rapporteringssystemen för avvikelser gällande samma urval. Resultaten visar att betydligt fler avvikelser identifierades med hjälp av retrospektiv journalgranskning. Med denna metod fann vi att 60 (15%) patienter drabbades av 65 avvikelser. I de traditionella rapporteringssystemen fanns endast fyra avvikelser registrerade av vilka två även identiferades i journalgranskningen. Totalt idenfierades 67 olika avvikelser. Drygt hälften av avvikelserna bedömdes som undvikbara. Av de patienter som ingick i journalgranskningsdelen av studien hade de som drabbades av avvikelser dubbelt så lång vårdtid jämfört med de som inte drabbats av en avvikelse. Trots att 59 av avvikelserna drabbade opererade patienter berodde endast nio av dessa på bristande kirurgisk eller anestesiologisk teknik, de resterande bedömdes bero på brister i vårdprocesserna. Patienter som drabbades av allvarligare avvikelser var signifikant äldre än de som drabbades av lindrigare avvikelser. De vanligaste avvikelserna var olika typer av vårdrelaterade infektioner och sen upptäckt av urinretention.

I den andra studien som redovisas i artikel III utvärderades effekten av ortopedklinikens förbättringsinitiativ benämnd temamånader. Den vetenskapliga utvärderingen av utfallet bestod av en fallstudieansats med utvärdering av process- och utfallsmått genom retrospektiv journalgranskning (46 punktprevalensmätningar) före, under och efter förbättringsinitiativen. Totalt ingick 2281 patienter. Resultatet visade signifikanta skillnader över tid vad gäller riskbedömning för trycksår och trycksårfrekvens. Efter förbättringsinitiativens start uppvisade ortopedkliniken bättre resultat i de årliga landstingsövergripande mätningarna än sjukhuset som helhet. Skillnader i resultaten mellan förannonserade årliga mätningar och oannonserade longitudinella retrospektiva mätningar kunde tydligt iakttas vilket leder till slutsatsen att förannonserade mätningar kan ge en felaktig bild av vårdens kvalitet. Journalgranskning kan vara en värdefull metod för att utvärdera förbättringsarbetens utfall över tid vars resultat med fördel kan återkopplas till verksamheten med hjälp av tydliga tidsseriediagram.

Artikel IV rapporterar resultaten från den tredje och sista studien. HMPS metoden har ofta beskrivits som oprecis och resursintensiv jämfört med GTT. GTT används nu av många olika landsting i Sverige men dess validitet vad gäller identifiering av skador i relation till en mer vetenskapligt utvärderad journalgranskningsmetod har enligt vår kännedom inte tidigare undersökts. I studien granskades 350 slumpvisa ortopedvårdtillfällen under 2009. Två team som vardera bestod av en sjuksköterska och två läkare granskade samma journaler med varsin journalgranskningsmetod i en två-stegs granskningsprocess plus ett utvärderingssteg. I utvärderingssteget jämfördes alla bekräftade skador och avslagna händelser från respektive metod och alla skador som missats av en av metoderna inkluderades i en ny granskningsrunda. Resultaten visade att HMPS metoden identifierade drygt 50% fler

skador på kortare tid än GTT efter steg två i granskningsprocessen. Totalt identifierades 160 olika skador vid 105 vårdtillfällen (30% av vårdtillfällena) med båda metoderna. HMPS metoden identifierade sammanlagt 155 skador vid 104 vårdtillfällen (30% av vårdtillfällena) efter utvärderingssteget. Av dessa bedömdes 135 (87.1%) som möjliga att undvika. Motsvarande siffror för GTT var 137 skador vid 99 vårdtillfällen (28% av vårdtillfällena), varav 110 (80.3%) bedömdes vara undvikbara. Skillnaderna mellan metoderna bedömdes bero på metodologi och att GTT i högre grad exkluderade mindre allvarliga skador. Övergripande positivt prediktivt värde för att identifiera skada var 40% för HMPS metoden och 30% för GTT.

Sammanfattningsvis så finns det ingen enskild datainsamlingsmetod som kan identifiera alla risker och avvikelser. Trots de begränsningar som en del av metoderna har så kompletterar de varandra genom att olika kvalitativa och kvantitativa data angående frekvens och natur identifieras. Flera studier, inklusive vår, har visat att retrospektiv journalgranskning jämfört med traditionella datainsamlingsmetoder för patientsäkerhetsinformation är den metod som enskilt fångar flest skador. När olika metoder har jämförts så finns det en väldigt liten överlappning i utfallen. Med en ökad medvetenhet om riskfaktorer och interventioner baserade på tvärprofessionella och klinikövergripande strategier som fokuserar på hälso- och sjukvårdens processer kan förhoppningsvis avvikelserna inom ortopedisk vård minska. Retrospektiv journalgranskning kan vara en viktig del i ett säkerhetsinformationssystem och i patientsäkerhetsarbetet samt kan även användas för att identifiera tillbud. Vid användning av journalgranskning bör olika metoders svagheter och styrkor tas i beaktande. Retrospektiv journalgranskning kan även användas för att följa upp kvalitets- och patientsäkerhetsinitiativ över tid och data kan redovisas i överskådliga tidsseriediagram för att stimulera utvecklingen.

Modellen som tillämpades i samband med temamånaderna underlättade att organisera förbättringsinitiativ inom ortopedisk omvårdnad. Det tog dock lång tid innan nya vårdprogram, kvalitetsindikatorer och patientsäkerhetsinitiativ började användas regelbundet i lokal klinisk praxis. Förbättringarna var moderata de första åren och det behövdes interventioner på flera nivåer för att uppnå stadigt ökande resultat.

Utbildning, noggrant utformade manualer och bra metodstöd och uppföljning kan vara framgångsfaktorer för att öka validitet och reliabilitet i journalgranskningarna. Genom att varje klinik kan ha egna granskningsteam kan sjukhusen öka antalet personer med metodkompetens och även uppnå lokalt engagemang i dessa frågor. Lokalt insamlade och analyserade data kan upplevas som mer trovärdiga och kan därmed utgöra ett bättre incitament för förbättringsarbete. Data från lokala journalgranskningar kan snabbare omsättas till faktabaserade patientsäkerhetsinitiativ. Genom lokala granskningar kan information från journalen kompletteras med andra informationskällor och ge en mer heltäckande bild av avvikelsernas natur och bakomliggande orsaker.

# ACKNOWLEDGEMENTS

This thesis has emerged from my own learning and experiences in patient safety and quality both when I was working as a RN and as an administrator within this field as well as a researcher in these issues. The thesis is now the end of my doctoral studies. A number of people have generously contributed to my work during these doctoral student years. This thesis has not been possible to carry out without the support and encouragement from my supervisors, colleagues both at the hospital and in research, friends and family. Thanks to you all!

I would especially like to express my gratitude to:

**Karin "Kiku" Pukk Härenstam,** my main supervisor, for your constant support, engagement and encouragement. Thank you for generously sharing your scientific and clinical knowledge about patient safety and quality improvement. You always inspired me to think one step further. I have enjoyed our discussions during this time and I hope we will continue them after the dissertation.

**Olle Muren,** my co-supervisor, for valuable advices and being with me on my research journey from the beginning. Thank you for all helps to plan and carry out the adverse event studies.

**Nils Dalén,** my co-supervisor, for being such a supportive coach and for your constructive comments during the years.

**Peter Henriksson,** my co-supervisor, for excellent scientific inputs and valuable advices.

**Ulf Lillkrona,** my former supervisor and former Head of the Orthopaedic Department, for believing in me and introducing me into the scientific world.

**Michael Soop,** my mentor, for stimulating discussions and for sharing your expertise in patient safety and especially in methodological issues in record review. Thank you for reading my drafts with outstanding feedback and for your patience with all my never ending detailed questions.

**Mats Brommels,** the director of Medical Management Centre (MMC), for "adopting" me to MMC and generously giving me access to all activities and doctoral student colleagues.

**Eila Sterner,** for true friendship, endless encouragement and optimism, co-authorship, and collaboration during the years. I hope we both will have more time this summer to sitting and doing nothing at Lill-Stallis.

**Lena Nilsson, Kristina Schildmeijer** and **Urban Jürgensen,** for fun and creative collaboration, co-authorship and for sharing your experiences in working with GTT.

**Drew Gaffney** and **Duncan Neuhauser,** for valuable patient safety and quality discussions, advices and help with drafts.

**The Vinnvård research group**: **Kiku Pukk Härenstam, Magna Andreen-Sachs, Drew Gaffney, Johan Thor, Italo Masiello, Carl Savage, Mats Hedsköld, Karolina Peltomaa** and **Eva Wesslén-Eriksson,** for your devotion in our common research projects and support in my research.

**Pamela Mazzocato,** for encouragement, collaboration and learning discussions about healthcare processes.

**Gustaf Neander,** current Head of the Orthopaedic Department, for your encouragement and support in resources.

**Colleagues at the Orthopaedic Department at Danderyd Hospital,** with special thanks to **Paula Kelly-Pettersson, Heléne Sjöö** and **Bodil Samuelsson,** for your support, pep talks and all good laughter.

**Jakob Ask, Mattias Elg** and **Elisabeth Berg,** for excellent statistical help.

**Anders Forsberg, Ann Fornander, Gunilla Berglund, Mi Ta, Krister Eriksson and Liivi Rimling,** for valuable help with different kind of administration in relation to the studies and for help with data.

**Nina Ringart,** at the Department of Clinical Sciences, Danderyd Hospital, for kind support with various things during my doctoral student years.

All personnel at the **Medical Library** at Danderyd Hospital for kind help to find references and especially, **Erika Wiberg** and **Petra Wallgren Björk** who have supported me with fast and valuable help with EndNote.

Thanks all my lovely and supporting friends for having a lot of good times. Especially **Elisabeth Atte**, **Sara Aynehband** and **Elisabet Tengqvist** for being there for me and for listen to my monologues about all ups and downs during the research trip.

**Alf** and **Gullan**, my dear parents, and my brothers **Nicklas** and **Håkan with family** for love and encouragements.

**Christoffer** with partner **Frida**, **Malin** and **Rasmus** my wonderful children who often have wondered what I was doing behind the computer and all papers and books. I am back!

And finally **Matts**, my beloved husband without whose support and endless patience these research years had been difficult to carry out. There are not words enough …….

# References


1. Olsen S. Detection of adverse events in surgery by record review [Thesis for doctoral degree (Ph.D.)]. London: University of London; 2006.
2. Neuhauser D. Ernest Amory Codman, M.D., and end results of medical care. *Int J Technol Assess Health Care* 1990;6:307-25.
3. Neuhauser D. Ernest Amory Codman MD. *Qual Saf Health Care* 2002;11:104-5.
4. Schimmel EM. The hazards of hospitalization. *Ann Intern Med* 1964;60:100-10.
5. McLamb JT, Huntley RR. The hazards of hospitalization. *South Med J* 1967;60:469-72.
6. Mills DH. Medical insurance feasibility study. A technical summary. *West J Med* 1978;128:360-5.
7. Steel K, Gertman PM, Crescenzi C, et al. Iatrogenic illness on a general medical service at a university hospital. *N Engl J Med* 1981;304:638-42.
8. Couch NP, Tilney NL, Rayner AA, et al. The high cost of low-frequency events: the anatomy and economics of surgical mishaps. *N Engl J Med* 1981;304:634-7.
9. Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *N Engl J Med* 1991;324:370-6.
10. Kohn LT, Corrigan JM, Donaldson MS. To err is human: building a safer health system. Washington, D.C.: National Academy Press 2000.
11. Wachter RM. Understanding patient safety. New York: McGraw-Hill Medical 2008.
12. Soop M, Fryksmark U, Koster M, et al. [Adverse events in hospitals are common. The majority can be avoided according to a study of medical records]. *Lakartidningen* 2008;105:1748-52.
13. Soop M, Fryksmark U, Koster M, et al. The incidence of adverse events in Swedish hospitals: a retrospective medical record review study. *Int J Qual Health Care* 2009;21:285-91.
14. Brown P, McArthur C, Newby L, et al. Cost of medical injury in New Zealand: a retrospective cohort study. *J Health Serv Res Policy* 2002;7 Suppl 1:S29-34.
15. Ehsani JP, Jackson T, Duckett SJ. The incidence and cost of adverse events in Victorian hospitals 2003-04. *Med J Aust* 2006;184:551-5.
16. Williams DJ, Olsen S, Crichton W, et al. Detection of adverse events in a Scottish hospital using a consensus-based methodology. *Scott Med J* 2008;53:26-30.
17. Hoonhout LH, de Bruijne MC, Wagner C, et al. Direct medical costs of adverse events in Dutch hospitals. *BMC Health Serv Res* 2009;9:27.
18. Vincent C. Patient safety. Chichester: Wiley-Blackwell 2010.
19. Wilson RM, Harrison BT, Gibberd RW, et al. An analysis of the causes of adverse events from the Quality in Australian Health Care Study. *Med J Aust* 1999;170:411-5.
20. World Health Organization. Conceptual Framework for the International Classification for Patient Safety. Final Technical Report: World Health Organization; 2009 January 2009.

21. Resar RK, Rozich JD, Classen D. Methodology and rationale for the measurement of harm with trigger tools. *Qual Saf Health Care* 2003;12 Suppl 2:ii39-45.is
22. Naessens JM, Campbell CR, Huddleston JM, et al. A comparison of hospital adverse events identified by three widely used detection methods. *International Journal for Quality in Health Care* 2009;21:301-7.
23. Sari B. Study of the scale, nature and causes of adverse events and methods to identify them [Thesis for doctoral degree (Ph.D.)]. York: University of York; 2009.
24. Reason JT. Managing the risks of organizational accidents. Aldershot, Hants: Ashgate 1997.
25. Griffin F, Resar R. IHI Global Trigger Tool for Measuring Adverse Events (Second Edition). Second ed. Cambridge, Massachusetts: Institute for Healthcare Improvement 2009.
26. Wilson RM, Runciman WB, Gibberd RW, et al. The Quality in Australian Health Care Study. *Med J Aust* 1995;163:458-71.
27. Baker GR, Norton PG, Flintoft V, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ* 2004;170:1678-86.
28. Hiatt HH, Barnes BA, Brennan TA, et al. A study of medical injury and medical malpractice. *N Engl J Med* 1989;321:480-4.
29. 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.
30. Kjellén U. Prevention of accidents through experience feedback. London: Taylor & Francis 2000.
31. Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. *Journal of General Internal Medicine* 2003;18:61-7.
32. Vincent C. Understanding and responding to adverse events. *New England Journal of Medicine* 2003;348:1051-6.
33. Pronovost P, Holzmueller CG, Needham DM, et al. How will we know patients are safer? An organization-wide approach to measuring and improving safety. *Crit Care Med* 2006;34:1988-95.
34. Thomas MJ, Schultz TJ, Hannaford N, et al. Mapping the limits of safety reporting systems in health care--what lessons can we actually learn? *Med J Aust* 2011;194:635-9.
35. Pronovost PJ, Morlock LL, Sexton JB, et al. Improving the Value of Patient Safety Reporting Systems Assessment. In: Henriksen K, Battles J, Keyes M, et al., eds. *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 1: Assessment)*. Rockville: Agency for Healthcare Research and Quality 2008.
36. Naessens JM, O'Byrne TJ, Johnson MG, et al. Measuring hospital adverse events: assessing inter-rater reliability and trigger performance of the Global Trigger Tool. *Int J Qual Health Care* 2010;22:266-74.
37. Ödegård S. Säker vård: patientskador, rapportering och prevention. Göteborg: Nordiska högskolan för folkhälsovetenskap (NHV) 2006.
38. Pukk Härenstam K. Learning from patient injury claims. Stockholm: Karolinska Institutet; 2007.

39. Kallberg AS, Goransson KE, Ostergren J, et al. Medical errors and complaints in emergency department care in Sweden as reported by care providers, healthcare staff, and patients - a national review. *Eur J Emerg Med* 2011. [Epub ahead of print].
40. Öhrn A. Measures of patient safety: studies of Swedish reporting systems and evaluation of an intervention aimed at improved patient safety culture. Linköping: Department of Medical and Health Sciences, Linköping University 2012.
41. Michel P. Strengths and weaknesses of available methods for assessing the nature and scale of harm caused by the health system: literature review. 2004 [cited 2012 March 24]; Available from: http://www.who.int/patientsafety/research/P_Michel_Report_Final_version.pdf.
42. World Health Organization. WHO draft guidelines for adverse event reporting and learning systems. From information to action. 2005 [cited 2012 March 24]; Available from: http://www.who.int/patientsafety/events/05/Reporting_Guidelines.pdf.
43. Lilford RJ, Mohammed MA, Braunholtz D, et al. The measurement of active errors: methodological issues. *Qual Saf Health Care* 2003;12 Suppl 2:ii8-12.
44. Davis P, Lay-Yee R, Briant R, et al. Adverse events in New Zealand public hospitals I: occurrence and impact. *N Z Med J* 2002;115:U271.
45. Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001;322:517-9.
46. Schioler T, Lipczak H, Pedersen BL, et al. [Incidence of adverse events in hospitals. A retrospective study of medical records]. *Ugeskr Laeger* 2001;163:5370-8.
47. Forster AJ, Asmis TR, Clark HD, et al. Ottawa Hospital Patient Safety Study: incidence and timing of adverse events in patients admitted to a Canadian teaching hospital. *CMAJ* 2004;170:1235-40.
48. Thomas EJ, Studdert DM, Burstin HR, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care* 2000;38:261-71.
49. Classen DC, Pestotnik SL, Evans RS, et al. Computerized surveillance of adverse drug events in hospital patients. *JAMA* 1991;266:2847-51.
50. Classen DC, Pestotnik SL, Evans RS, et al. Description of a computerized adverse drug event monitor using a hospital information system. *Hosp Pharm* 1992;27:774, 6-9, 83.
51. Good VS, Saldana 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. *Qual Saf Health Care* 2011;20:25-30;46:654-78
52. 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;46:654-78.
53. Classen DC, Lloyd RC, Provost L, et al. Development and Evaluation of the Institute for Healthcare Improvement Global Trigger Tool. *J Patient Saf* 2008;4:169-177.
54. Griffin FA, Classen DC. Detection of adverse events in surgical patients using the Trigger Tool approach. *Qual Saf Health Care* 2008;17:253-8.
55. Kaafarani HM, Rosen AK, Nebeker JR, et al. Development of trigger tools for surveillance of adverse events in ambulatory surgery. *Qual Saf Health Care* 2010;19:425-9.

56. Resar RK, Rozich JD, Simmonds T, et al. A trigger tool to identify adverse events in the intensive care unit. *Jt Comm J Qual Patient Saf* 2006;32:585-90.
57. 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.
58. Takata GS, Mason W, Taketomo C, et al. Development, testing, and findings of a pediatric-focused trigger tool to identify medication-related harm in US children's hospitals. *Pediatrics* 2008;121:e927-35.
59. Sharek PJ, Horbar JD, Mason W, et al. Adverse events in the neonatal intensive care unit: development, testing, and findings of an NICU-focused trigger tool to identify harm in North American NICUs. *Pediatrics* 2006;118:1332-40.
60. Matlow AG, Cronin CM, Flintoft V, et al. Description of the development and validation of the Canadian Paediatric Trigger Tool. *BMJ Qual Saf* 2011;20:416-23.
61. Larsen GY, Donaldson AE, Parker HB, et al. Preventable harm occurring to critically ill children. *Pediatr Crit Care Med* 2007;8:331-6.
62. Rosen AK, Mull HJ, Kaafarani H, et al. Applying trigger tools to detect adverse events associated with outpatient surgery. *J Patient Saf* 2011;7:45-59.
63. National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP). NCC MERP Index for Categorizing Medication Errors. [cited 2012 March 11]; Available from: http://www.nccmerp.org/medErrorCatIndex.html.
64. Strukturerad journalgranskning för att identifiera och mäta förekomst av skador i vården enligt metoden Global Trigger Tool: handbok för patientsäkerhetsarbete. Stockholm: Kommentus 2008.
65. Schildmeijer K, Nilsson L, Arestedt K, et al. Assessment of adverse events in medical care: lack of consistency between experienced teams using the global trigger tool. *BMJ Qual Saf* 2012;21:307-14.
66. Nilsson L, Juhlin C, Krook H, et al. [Structured scrutiny of medical records can increase patient safety]. *Lakartidningen* 2009;106:2125-8.
67. Sjodahl R, Hultkrantz P, Melander H, et al. [High frequency of postoperative complications. Among patients with hospital stay of at least 5 days almost every third is affected]. *Lakartidningen* 2010;107:2636-9.
68. Landrigan CP, Parry GJ, Bones CB, et al. Temporal trends in rates of patient harm resulting from medical care. *N Engl J Med* 2010;363:2124-34.
69. Michel P, Quenon JL, de Sarasqueta AM, et al. Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. *BMJ* 2004;328:199.
70. Sari AB, Sheldon TA, Cracknell A, et al. Extent, nature and consequences of adverse events: results of a retrospective casenote review in a large NHS hospital. *Qual Saf Health Care* 2007;16:434-9.
71. Olsen S, Neale G, Schwab K, et al. Hospital staff should use more than one method to detect adverse events and potential adverse events: incident reporting, pharmacist surveillance and local real-time record review may all have a place. *Qual Saf Health Care* 2007;16:40-4.
72. Marang-van de Mheen PJ, van Hanegem N, Kievit J. Effectiveness of routine reporting to identify minor and serious adverse outcomes in surgical patients. *Qual Saf Health Care* 2005;14:378-82.

73. Samore MH, Evans RS, Lassen A, et al. Surveillance of medical device-related hazards and adverse events in hospitalized patients. *JAMA* 2004;291:325-34.
74. Hogan H, Olsen S, Scobie S, et al. What can we learn about patient safety from information sources within an acute hospital: a step on the ladder of integrated risk management? *Qual Saf Health Care* 2008;17:209-15.
75. Classen DC, Resar R, Griffin F, et al. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff (Millwood)* 2011;30:581-9.
76. Wolff AM, Bourke J. Detecting and reducing adverse events in an Australian rural base hospital emergency department using medical record screening and review. *Emergency Medicine Journal* 2002;19:35-40.
77. Wolff AM, Bourke J, Campbell IA, et al. Detecting and reducing hospital adverse events: outcomes of the Wimmera clinical risk management program. *Medical Journal of Australia* 2001;174:621-5.
78. Christiaans-Dingelhoff I, Smits M, Zwaan L, et al. To what extent are adverse events found in patient records reported by patients and healthcare professionals via complaints, claims and incident reports? *BMC Health Serv Res* 2011;11:49.
79. Henriksen K, Kaplan H. Hindsight bias, outcome knowledge and adaptive learning. *Qual Saf Health Care* 2003;12 Suppl 2:ii46-50.
80. Griffin F, Resar R. IHI Global Trigger Tool for Measuring Adverse Events. Cambridge, Massachusetts: Institute for Healthcare Improvement 2007.
81. Leape LL. Reporting of adverse events. *New England Journal of Medicine* 2002;347:1633-8.
82. SFS 2010:659. Patientsäkerhetslag. Stockholm: Riksdagen 2010.
83. SOSFS 2005:28. Socialstyrelsens föreskrifter och allmänna råd om anmälningsskyldighet enligt Lex Maria. Stockholm: Socialstyrelsen 2005.
84. SOSFS 2010:4. Ändring i föreskrifterna och allmänna råden (SOSFS 2005:28) om anmälningsskyldighet enligt lex Maria. Stockholm: Socialstyrelsen 2010.
85. LVFS 2001:12. Läkemedelsverkets föreskrifter om säkerhetsövervakning av läkemedel. Uppsala: Läkemedelsverket 2001.
86. LVFS 2006:4. Föreskrifter om ändring i Läkemedelsverkets föreskrifter (LVFS 2001:12) om säkerhetsövervakning av läkemedel. Uppsala: Läkemedelsverket 2006.
87. Läkemedelsverket. [cited 2012 March 24]; Available from: http://www.lakemedelsverket.se/.
88. SOSFS 2008:1. Socialstyrelsens föreskrifter om användning av medicintekniska produkter i hälso- och sjukvården. Stockholm: Socialstyrelsen 2008.
89. Socialstyrelsen. Om du vill klaga på vården. Stockholm: Socialstyrelsen; 2011.
90. Wolff AM, Bourke J. Reducing medical errors: a practical guide. *Medical Journal of Australia* 2000;173:247-51.
91. SFS 1996:799. Patientskadelag. Stockholm: Riksdagen 1996.
92. Landstingens Ömsesidiga Försäkringsbolag. Om du skadas i vården – kan du ha rätt till ersättning enligt patientskadelagen. [cited 2012 March 24]; Available from: http://www.patientforsakring.se/resurser/dokument/informationsmaterial/Om-du-skadas-i-varden.pdf.

93. Landstingens Ömsesidiga Försäkringsbolag. Vilken ersättning kan du få? [cited 2012 March 24]; Available from: http://www.patientforsakring.se/Vilken-ersattning-kan-du-fa.html.
94. Pukk-Harenstam K, Ask J, Brommels M, et al. Analysis of 23 364 patient-generated, physician-reviewed malpractice claims from a non-tort, blame-free, national patient insurance system: lessons learned from Sweden. *Qual Saf Health Care* 2008;17:259-63.
95. Ohrn A, Elfstrom J, Liedgren C, et al. Reporting of sentinel events in Swedish hospitals: a comparison of severe adverse events reported by patients and providers. *Jt Comm J Qual Patient Saf* 2011;37:495-501.
96. Patientnämnden. [cited 2012 March 24]; Available from: http://www.patientnamndenstockholm.se/index.html.
97. Murff HJ, Patel VL, Hripcsak G, et al. Detecting adverse events for patient safety research: a review of current methodologies. *J Biomed Inform* 2003;36:131-43.
98. Weingart SN, Pagovich O, Sands DZ, et al. What can hospitalized patients tell us about adverse events? Learning from patient-reported incidents. *J Gen Intern Med* 2005;20:830-6.
99. Ohrn A, Olai A, Rutberg H, et al. Adverse events in spine surgery in Sweden: a comparison of patient claims data and national quality register (Swespine) data. *Acta Orthop* 2011;82:727-31.
100. Wessel M, Lynoe N, Juth N, et al. The tip of an iceberg? A cross-sectional study of the general public's experiences of reporting healthcare complaints in Stockholm, Sweden. *BMJ Open* 2012;2(1):e000489.
101. Skar L, Soderberg S. Complaints with encounters in healthcare - men's experiences. *Scand J Caring Sci* 2011. [Epub ahead of print].
102. Soderberg S, Olsson M, Skar L. A hidden kind of suffering: female patient's complaints to Patient's Advisory Committee. *Scand J Caring Sci* 2011;26:144-50.
103. Bismark MM, Brennan TA, Paterson RJ, et al. Relationship between complaints and quality of care in New Zealand: a descriptive analysis of complainants and non-complainants following adverse events. *Qual Saf Health Care* 2006;15:17-22.
104. Griffen FD, Stephens LS, Alexander JB, et al. The American College of Surgeons' closed claims study: new insights for improving care. *J Am Coll Surg* 2007;204:561-9.
105. Vincent C, Davy C, Esmail A, et al. Learning from litigation. The role of claims analysis in patient safety. J Eval Clin Pract 2006;12:665-74.
106. Noble DJ, Pronovost PJ. Underreporting of patient safety incidents reduces health care's ability to quantify and accurately measure harm reduction. *J Patient Saf* 2010;6:247-50.
107. Suresh G, Horbar JD, Plsek P, et al. Voluntary anonymous reporting of medical errors for neonatal intensive care. *Pediatrics* 2004;113:1609-18.
108. Wanzel KR, Jamieson CG, Bohnen JM. Complications on a general surgery service: incidence and reporting. *Can J Surg* 2000;43:113-7.
109. Stanhope N, Crowley-Murphy M, Vincent C, et al. An evaluation of adverse incident reporting. *Journal of Evaluation in Clinical Practice* 1999;5:5-12.
110. Zegers M, de Bruijne MC, Wagner C, et al. Design of a retrospective patient record study on the occurrence of adverse events among patients in Dutch hospitals. *BMC Health Serv Res* 2007;7:27.

111. Walshe K. Adverse events in health care: issues in measurement. *Qual Health Care* 2000;9:47-52.
112. Vincent C, Stanhope N, Crowley-Murphy M. Reasons for not reporting adverse incidents: an empirical study. *Journal of Evaluation in Clinical Practice* 1999;5:13-21.
113. Jeffe DB, Dunagan WC, Garbutt J, et al. Using focus groups to understand physicians' and nurses' perspectives on error reporting in hospitals. *Jt Comm J Qual Saf* 2004;30:471-9.
114. Kingston MJ, Evans SM, Smith BJ, et al. Attitudes of doctors and nurses towards incident reporting: a qualitative analysis. *Med J Aust* 2004;181:36-9.
115. Evans SM, Berry JG, Smith BJ, et al. Attitudes and barriers to incident reporting: a collaborative hospital study. *Qual Saf Health Care* 2006;15:39-43.
116. Milch CE, Salem DN, Pauker SG, et al. Voluntary electronic reporting of medical errors and adverse events. An analysis of 92,547 reports from 26 acute care hospitals. *J Gen Intern Med* 2006;21:165-70.
117. Rowin EJ, Lucier D, Pauker SG, et al. Does error and adverse event reporting by physicians and nurses differ? *Jt Comm J Qual Patient Saf* 2008;34:537-45.
118. Taylor JA, Brownstein D, Christakis DA, et al. Use of incident reports by physicians and nurses to document medical errors in pediatric patients. *Pediatrics* 2004;114:729-35.
119. Hirose M, Regenbogen SE, Lipsitz S, et al. Lag time in an incident reporting system at a university hospital in Japan. *Qual Saf Health Care* 2007;16:101-4.
120. Nuckols TK, Bell DS, Liu H, et al. Rates and types of events reported to established incident reporting systems in two US hospitals. *Qual Saf Health Care* 2007;16:164-8.
121. Zhu J, Stuver SO, Epstein AM, et al. Can we rely on patients' reports of adverse events? *Med Care* 2011;49:948-55.
122. Friedman SM, Provan D, Moore S, et al. Errors, near misses and adverse events in the emergency department: what can patients tell us? *CJEM* 2008;10:421-7.
123. Zegers M, de Bruijne MC, Wagner C, et al. Adverse events and potentially preventable deaths in Dutch hospitals: results of a retrospective patient record review study. *Qual Saf Health Care* 2009;18:297-302.
124. Sari AB, Cracknell A, Sheldon TA. Incidence, preventability and consequences of adverse events in older people: results of a retrospective case-note review. *Age Ageing* 2008;37:265-9.
125. Thomas EJ, Brennan TA. Incidence and types of preventable adverse events in elderly patients: population based review of medical records. *British Medical Journal* 2000;320:741-4.
126. Evans SM, Berry JG, Smith BJ, et al. Consumer perceptions of safety in hospitals. *BMC Public Health* 2006;6:41.
127. SOSFS 2011:9. Socialstyrelsens föreskrifter och allmänna råd om ledningssystem för systematiskt kvalitetsarbete. Stockholm: Socialstyrelsen 2011.
128. Riskanalys & händelseanalys: handbok för patientsäkerhetsarbete. Stockholm: Socialstyrelsen 2009.
129. Vincent CA. Analysis of clinical incidents: a window on the system not a search for root causes. *Qual Saf Health Care* 2004;13:242-3.

130. Sveriges Kommuner och Landsting. Kvalitetsregister. [cited 2012 March 24]; Available from: http://www.kvalitetsregister.se/.
131. Franneby U, Sandblom G, Nyren O, et al. Self-reported adverse events after groin hernia repair, a study based on a national register. *Value Health* 2008;11:927-32.
132. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. *Qual Saf Health Care* 2003;12 Suppl 2:ii58-63.
133. Rivard PE, Rosen AK, Carroll JS. Enhancing patient safety through organizational learning: Are patient safety indicators a step in the right direction? *Health Serv Res* 2006;41:1633-53.
134. Bates DW, Evans RS, Murff H, et al. Detecting adverse events using information technology. *Journal of the American Medical Informatics Association* 2003;10:115-28.
135. Gawande AA, Zinner MJ, Studdert DM, et al. Analysis of errors reported by surgeons at three teaching hospitals. *Surgery* 2003;133:614-21.
136. Gawande AA, Thomas EJ, Zinner MJ, et al. The incidence and nature of surgical adverse events in Colorado and Utah in 1992. *Surgery* 1999;126:66-75.
137. Frankel A, Grillo SP, Pittman M, et al. Revealing and resolving patient safety defects: the impact of leadership WalkRounds on frontline caregiver assessments of patient safety. *Health Serv Res* 2008;43:2050-66.
138. Sveriges Kommuner och Landsting. Att mäta patientsäkerhetskultur: tipsguide - från mätning till åtgärder. Stockholm: Sveriges Kommuner och Landsting 2011.
139. Runciman WB, Williamson JA, Deakin A, et al. An integrated framework for safety, quality and risk management: an information and incident management system based on a universal patient safety classification. *Qual Saf Health Care* 2006;15 Suppl 1:i82-90.
140. Kaplan HS, Fastman BR. Organization of event reporting data for sense making and system improvement. *Qual Saf Health Care* 2003;12 Suppl 2:ii68-72.
141. Carthey J, de Leval MR, Reason JT. The human factor in cardiac surgery: errors and near misses in a high technology medical domain. *Annals of Thoracic Surgery* 2001;72:300-5.
142. Habraken MM, van der Schaaf TW. If only....: failed, missed and absent error recovery opportunities in medication errors. *Qual Saf Health Care* 2010;19:37-41.
143. Wolff AM, Taylor S. Wimmera Clinical Risk Management Model. A systematic approach to reducing medical errors. A framework for implementation and signposts for success. A manual. Horsham: Artisan Design 2001.
144. Craddick JW, Bader B. Medical management analysis: a systematic approach to quality assurance and risk management. Auburn (CA): Joyce W. Craddick 1983.
145. Bate P, Mendel P, Robert GB. Organizing for quality: the improvement journeys of leading hospitals in Europe and the United States. Oxford: Radcliffe 2008.
146. Yin RK. Case study research : design and methods. London: SAGE 2009.
147. Altman DG. Practical statistics for medical research. London: Chapman and Hall 1991.
148. Ek AC, Unosson M, Bjurulf P. The modified Norton scale and the nutritional state. *Scand J Caring Sci* 1989;3:183-7.
149. SFS 1998:204. Personuppgiftslag. Stockholm: Riksdagen 1998.
150. Wolff AM. Limited adverse occurrence screening: an effective and efficient method of medical quality control. *J Qual Clin Pract* 1995;15:221-33.

151. Sari AB, Sheldon TA, Cracknell A, et al. Sensitivity of routine system for reporting patient safety incidents in an NHS hospital: retrospective patient case note review. *BMJ* 2007;334:79.
152. Weissman JS, Schneider EC, Weingart SN, et al. Comparing patient-reported hospital adverse events with medical record review: do patients know something that hospitals do not? *Ann Intern Med* 2008;149:100-8.
153. Briant R, Morton J, Lay-Yee R, et al. Representative case series from public hospital admissions 1998 II: surgical adverse events. *N Z Med J* 2005;118:U1591.
154. Zegers M, de Bruijne MC, de Keizer B, et al. The incidence, root-causes, and outcomes of adverse events in surgical units: implication for potential prevention strategies. *Patient Saf Surg* 2011;5:13.
155. Kable AK, Gibberd RW, Spigelman AD. Adverse events in surgical patients in Australia. *International Journal for Quality in Health Care* 2002;14:269-76.
156. Kable A, Gibberd R, Spigelman A. Predictors of adverse events in surgical admissions in Australia. *Int J Qual Health Care* 2008;20:406-11.
157. Arvidsson S, Ouchterlony J, Nilsson S, et al. The Gothenburg study of perioperative risk. I. Preoperative findings, postoperative complications. *Acta Anaesthesiol Scand* 1994;38:679-90.
158. Neale G, Woloshynowych M, Vincent C. Exploring the causes of adverse events in NHS hospital practice. *J R Soc Med* 2001;94:322-30.
159. Davis P, Lay-Yee R, Briant R, et al. Preventable in-hospital medical injury under the ”no fault” system in New Zealand. *Quality & Safety in Health Care* 2003;12:251-6.
160. Marang-van de Mheen PJ, Mertens BJ, van Houwelingen HC, et al. Surgery groups differed in adverse outcome probabilities and can be used to adjust hospital comparisons. *J Clin Epidemiol* 2005;58:56-62.
161. Zegers M, De Bruijne MC, Spreeuwenberg P, et al. Variation in the rates of adverse events between hospitals and hospital departments. *Int J Qual Health Care* 2011;23:126-33.
162. Feldman L, Barkun J, Barkun A, et al. Measuring postoperative complications in general surgery patients using an outcomes-based strategy: comparison with complications presented at morbidity and mortality rounds. *Surgery* 1997;122:711-9; discussion 9-20.
163. Aranaz-Andres JM, Aibar-Remon C, Vitaller-Murillo J, et al. Incidence of adverse events related to health care in Spain: results of the Spanish National Study of Adverse Events. *J Epidemiol Community Health* 2008;62:1022-9.
164. Runciman WB, Edmonds MJ, Pradhan M. Setting priorities for patient safety. *Quality & Safety in Health Care* 2002;11:224-9.
165. de Vries EN, Ramrattan MA, Smorenburg SM, et al. The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care* 2008;17:216-23.
166. Bates DW, Oneil AC, Petersen LA, et al. Evaluation of Screening Criteria for Adverse Events in Medical Patients. *Medical Care* 1995;33:452-62.
167. Silver MP, Hougland P, Elder S, et al. Statewide identification of adverse events using retrospective nurse review: methods and outcomes. *J Nurs Meas* 2007;15:220-32.
168. Perla RJ, Provost LP, Murray SK. The run chart: a simple analytical tool for learning from variation in healthcare processes. *BMJ Qual Saf* 2011;20:46-51.

169. Thor J, Lundberg J, Ask J, et al. Application of statistical process control in healthcare improvement: systematic review. *Qual Saf Health Care* 2007;16:387-99.
170. de Vos M, Graafmans W, Kooistra M, et al. Using quality indicators to improve hospital care: a review of the literature. *Int J Qual Health Care* 2009;21:119-29.
171. Grimshaw J, Eccles M, Tetroe J. Implementing clinical guidelines: current evidence and future implications. *J Contin Educ Health Prof* 2004;24 Suppl 1:S31-7.
172. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet* 2003;362:1225-30.
173. Rasmussen JS, I. Proactive risk management in dynamic society. Karlstad, Sweden: Swedish Rescue Services Agency 2000.
174. Benn J, Koutantji M, Wallace L, et al. Feedback from incident reporting: information and action to improve patient safety. *Qual Saf Health Care* 2009;18:11-21.
175. Institute for Healthcare Improvement. The Breakthrough Series: IHI's Collaborative Model for Achieving Breakthrough Improvement. Cambridge, Massachusetts: Institute for Healthcare Improvement 2003.
176. Deming WE. Out of the crisis: quality, productivity and competitive position. Cambridge: Cambridge University Press. 1986.
177. Forsner T. Turning guidelines into clinical practice: findings from an implementary study. Stockholm: Department of Public Health Science, Karolinska Institutet 2010.
178. Baker GR. The contribution of case study research to knowledge of how to improve quality of care. *BMJ Qual Saf* 2011;20 Suppl 1:i30-5.
179. Battles JB, Lilford RJ. Organizing patient safety research to identify risks and hazards. *Qual Saf Health Care* 2003;12 Suppl 2:ii2-7.
180. Bergman B, Klefsjö B. Kvalitet från behov till användning. Lund: Studentlitteratur 2007.
181. Shekelle PG, Pronovost PJ, Wachter RM, et al. Advancing the science of patient safety. *Ann Intern Med* 2011;154:693-6.
182. Brandrud AS, Schreiner A, Hjortdahl P, et al. Three success factors for continual improvement in healthcare: an analysis of the reports of improvement team members. *BMJ Qual Saf* 2011;20:251-9.
183. Batalden PB, Stoltz PK. A framework for the continual improvement of health care: building and applying professional and improvement knowledge to test changes in daily work. *Jt Comm J Qual Improv* 1993;19:424-47; discussion 48-52.

# APPENDIX.

## USED SCALES, SCREENING CRITERIA AND TRIGGERS

Scales, screening criteria and triggers used in paper I, II and IV are listed below.

The screening criteria used in paper I and II and at Wimmera Base Hospital[77] are displayed in Table 1.

**Table 1.**

| Screening criteria | Wimmera Clinical Risk Management Model's screening criteria |
|---|---|
| 1. Unplanned readmission as a result of healthcare management within 28 days | 1. Unplanned readmission within 28 days of discharge |
| 2. Transfer from general care to intensive care | 2. Transfer from general care to intensive care |
| 3. Unplanned transfer to another department | 3. Transfer to another acute care facility |
| 4. Unplanned return to operating theatre within seven days | 4. Return to operating theatre within seven days |
| 5. Death | 5. Death |
| 6. Cardiorespiratory arrest | 6. Cardiac arrest |
| 7. Length of stay greater than 7 days | 7. Length of stay greater than 21 days |
| 8. External cause of injury codes | 8. Booked for theatre and cancelled |
| 9. Hospital-acquired infection or sepsis | 9. Any medical record recommended for review (added later) |
| 10. Complications during surgery | |
| 11. Hospital complications that developed during admission or within 28 days of discharge (e.g. myocardial infarction, stroke, or pulmonary embolism) | |
| 12. Other undesirable outcome not covered by other criteria | |

Causation and preventability scale used in paper I, II and IV are shown in Table 2[26].

**Table 2.**

| Causation and preventability scale |
|---|
| 1. Little or no evidence |
| 2. Slight evidence |
| 3. Not likely (less than 50:50 odds) |
| 4. More likely than not (greater than 50:50 odds) |
| 5. Strong evidence |
| 6. Virtually certain evidence |

Severity scale used in paper I and II is presented in Table 3[144].

**Table 3.**

| Severity scale |
|---|
| 1. Minor severity |
| 2. Minor temporary |
| 3. Minor permanent |
| 4. Major temporary |
| 5. Major permanent |
| 6. Potential major or major contributing |
| 7. Death |

Severity scale used by HMPS method team in paper IV is shown in Table 4[9 13].

**Table 4.**

| Severity scale HMPS |
|---|
| 1. Minimal impairment, recovery within 1 month |
| 2. Moderate impairment, recovery within 1 to 6 months |
| 3. Moderate impairment, recovery within 6 to 12 months |
| 4. Permanent impairment, degree of disability <50% |
| 5. Permanent impairment, degree of disability >50% |
| 6. Contributed to patient death |
| 7. Unable to determine |

Severity scale used by GTT team in paper IV, National Coordination Council for Medication Error Reporting and Prevention (NCC MERP) index, is displayed in Table 5[63].

**Table 5.**

| Severity scale GTT | |
|---|---|
| E. | Contributed to or resulted in temporary harm to the patient and required intervention |
| F. | Contributed to or resulted in temporary harm to the patient and required initial or prolonged hospitalization |
| G. | Contributed to or caused permanent patient harm |
| H. | Intervention required to sustain life |
| I. | Contributed to patient death |

Screening criteria used by the HMPS method team in paper IV can be seen in Table 6[13].

**Table 6.**

| Screening criteria HMPS method | |
|---|---|
| 1. | The index admission was an unplanned admission related to previous healthcare management within 30 days |
| 2. | Unplanned re-admission after discharge from index admission within 30 days including outpatient visits |
| 3. | Hospital-incurred patient injury or no-harm incident |
| 4. | Adverse drug reaction |
| 5. | Unplanned transfer from general care to intensive care |
| 6. | Unplanned transfer to another acute care hospital |
| 7. | Unplanned return to the operating room |
| 8. | Unplanned removal, injury or repair of an organ during surgery |
| 9. | Other patient complication |
| 10. | Development of neurological deficit not present on admission |
| 11. | Unexpected death |
| 12. | Inappropriate discharge to home |
| 13. | Cardiac or respiratory arrest |
| 14. | Injury related to abortion or delivery |
| 15. | Healthcare-associated infection or sepsis |
| 16. | Dissatisfaction with care documented in the patient's medical record |
| 17. | Documentation or correspondence indicating litigation |
| 18. | Any other undesirable outcome not covered above |

All screening criteria except 1, 2 or 4 must occur during index orthopaedic admission to be included as positive screening criteria.

Triggers used by the GTT team in paper IV are listed in Table 7[64].

**Table 7.**

| Triggers GTT |
|---|
| *Care Module triggers* |
| 1. Transfusion of blood or use of blood products |
| 2. Abrupt drop in haemoglobin |
| 3. In-hospital stroke |
| 4. Codes or arrest |
| 5. Dialysis |
| 6. Positive blood culture |
| 7. X-ray or Doppler studies for emboli or deep vein thrombosis |
| 8. Falls |
| 9. Pressure ulcers |
| 10. Re-admission within 30 days |
| 11. Healthcare-associated infections |
| 12. Transfer to higher level of care |
| 13. Procedure |
| 14. Care: other |
| *Medication Module Triggers* |
| 1. Clostridium difficile positive stool |
| 2. Partial Thromboplastin Time (PTT) greater than 100 seconds |
| 3. International Normalised Ratio (INR) greater than 6 |
| 4. Glucose less than 3 mmol/litre |
| 5. Rising BUN or serum creatinine two times (2x) over baseline |
| 6. Vitamin K administration |
| 7. Diphenhydramine administration |
| 8. Flumazenil administration |
| 9. Naloxone administration |
| 10. Anti-emetic administration |
| 11. Over-sedation/hypotension |
| 12. Abrupt medication stop |
| *Surgical Module triggers* |
| 1. Return to surgery |
| 2. Change in procedure |
| 3. Admission to intensive care post-operatively |
| 4. Intubation/reintubation / BiPaP in post anaesthesia care unit |
| 5. X-ray intra-operative or in post anaesthesia care unit |
| 6. Intra- or postoperative death |
| 7. Mechanical ventilation greater than 24 h post-operatively |
| 8. Intra-operative administration of epinephrine, norepinephrine, naloxone or flumazenil |
| 9. Post-operative increase in troponin levels |
| 10. Change of anaesthetic during surgery |
| 11. Consult requested in post anaesthesia care unit |
| 12. Occurrence of any postoperative complication |
| 13. Pathology reports normal or identifying specimen unrelated to initial surgical diagnosis |
| 14. Insertion of arterial or central venous line during surgery |
| 15. Operative time greater than 6 h |
| 16. Removal/injury or repair of organ during operative procedure |
| *Intensive Care Module Triggers* |
| 1. Pneumonia onset |
| 2. Readmission to intensive care unit |
| 3. In-unit procedure |
| 4. Intubation/reintubation |
| *Emergency Department (ED) Module Triggers* |
| 1. Re-admission to the ED within 48 hours |
| 2. Time in ED greater than 6 hours |

The five GTT perinatal triggers were not applicable in this study.