# Detectie van adverse events in administratieve databanken

*KCE reports 93A*

Federaal Kenniscentrum voor de Gezondheidszorg
Centre fédéral d'expertise des soins de santé
2008

## Het Federaal Kenniscentrum voor de Gezondheidszorg

Voorstelling : Het Federaal Kenniscentrum voor de Gezondheidszorg is een parastatale, opgericht door de programma-wet van 24 december 2002 (artikelen 262 tot 266) die onder de bevoegdheid valt van de Minister van Volksgezondheid en Sociale Zaken. Het Centrum is belast met het realiseren van beleidsondersteunende studies binnen de sector van de gezondheidszorg en de ziekteverzekering.

## Raad van Bestuur

Effectieve leden : Gillet Pierre (Voorzitter), Cuypers Dirk (Ondervoorzitter), Avontroodt Yolande, De Cock Jo (Ondervoorzitter), De Meyere Frank, De Ridder Henri, Gillet Jean-Bernard, Godin Jean-Noël, Goyens Floris, Kesteloot Katrien, Maes Jef, Mertens Pascal, Mertens Raf, Moens Marc, Perl François, Smiets Pierre, Van Massenhove Frank, Vandermeeren Philippe, Verertbruggen Patrick, Vermeyen Karel.

Plaatsvervangers : Annemans Lieven, Bertels Jan, Collin Benoît, Cuypers Rita, Decoster Christiaan, Dercq Jean-Paul, Désir Daniel, Laasman Jean-Marc, Lemye Roland, Morel Amanda, Palsterman Paul, Ponce Annick, Remacle Anne, Schrooten Renaat, Vanderstappen Anne.

Regeringscommissaris : Roger Yves

## Directie

Algemeen Directeur a.i. : Jean-Pierre Closon

Adjunct-Algemeen Directeur a.i. : Gert Peeters

## Contact

Federaal Kenniscentrum voor de Gezondheidszorg (KCE)
Administratief Centrum Kruidtuin, Doorbuilding (10e verdieping)
Kruidtuinlaan 55
B-1000 Brussel
Belgium

Tel: +32 [0]2 287 33 88
Fax: +32 [0]2 287 33 85

Email : info@kce.fgov.be
Web : http://www.kce.fgov.be

# Detectie van adverse events in administratieve databanken

*KCE reports 93A*

PIERRE GILLET, PHILIPPE KOLH, WALTER SERMEUS, ARTHUR VLEUGELS,
JESSICA JACQUES, KOEN VAN DEN HEEDE, STEPHAN DEVRIESE, FRANCE VRIJENS,
SANDRA VERELST

*Federaal Kenniscentrum voor de Gezondheidszorg*
*Centre fédéral d'expertise des soins de santé*
*2008*

**KCE REPORTS 93A**


| | |
|---|---|
| Titel: | Detectie van adverse events in administratieve databanken |
| Auteurs: | Pierre Gillet (CHU Liège), Philippe Kolh (CHU Liège), Walter Sermeus (KULeuven), Arthur Vleugels (KULeuven), Jessica Jacques (CHU Liège), Koen Van den heede (KULeuven), Stephan Devriese (KCE), France Vrijens (KCE), Sandra Verelst (KULeuven) |
| Externe experten: | Eric Baert (UZ Gent), Pascal Meeus (RIZIV) |
| Acknowledgements: | Adelin Albert (Biostatistical Center, University of Liege), Thibaut Degrave (University Hospital (CHU) of Liege), Martine Frenay (CHPLT), Emmanuel Lesaffre (Biostatistical Center, Katholieke Universiteit Leuven), Nathalie Maes (University Hospital (CHU) of Liege), Michel Meessen (University Hospital (CHU) of Liege), Sandrina von Winckelmann (Hospital Pharmacy Division, Katholieke Universiteit Leuven) |
| Externe validatoren: | Xavier De Béthune (Mutualité Chrétienne), Martine De Bruyne (EMGO-VUmc / Nivel, NL), Johan Hellings (Ziekenhuis Oost-Limburg) |
| Conflict of interest: | Geen gemeld |
| Disclaimer: | De externe experten hebben aan het wetenschappelijke rapport meegewerkt dat daarna aan de validatoren werd voorgelegd. De validatie van het rapport volgt uit een consensus of een meerderheidsstem tussen de validatoren. Alleen het KCE is verantwoordelijk voor de eventuele resterende vergissingen of onvolledigheden alsook voor de aanbevelingen aan de overheid. |
| Layout: | Ine Verhulst, Wim Van Moer |

Brussel, 17 November 2008

Studie nr 2006-21

Domein: Health Services Research (HSR)

MeSH: Adverse Effects ; Hospital Records ; Databases as Topic ; Risk Assessment

NLM classificatie: W 26.55.I4

Taal: Nederlands, Engels

Formaat: Adobe® PDF™ (A4)

Wettelijk depot: D/2008/10.273/73




Hoe refereren naar dit document?

Gillet P, Kolh P, Sermeus W, Vleugels A, Jacques J, Van den heede K, et al. Detectie van adverse events in administratieve databanken. Health Services Research (HSR). Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE); 2008. KCE reports 93A (D/2008/10.273/73)

# VOORWOORD

We willen dat alles goed gaat wanneer we geconfronteerd worden met ernstige ziekte en we zijn gewoon aan een hoge kwaliteit van zorg wanneer we gehospitaliseerd worden voor de behandeling van een ziekte. Echter zoals in alle menselijke activiteiten, gaat het niet altijd zoals we zouden willen. Dit wordt in de medische zorg vaak beschreven als adverse events. Voorvallen die niet gepland waren, maar zich niettemin voordeden.

De medische wetenschap beoogt een steeds betere zorg door het zoeken naar betere behandelingen, maar ook door het verhogen van de kwaliteit in het toepassen van deze behandelingen. Veel inspanningen gaan, en moeten gaan, naar preventie van adverse events. Echter, preventie is maar de eerste, weliswaar grote, stap in het omgaan met adverse events in de dagelijkse medische praktijk. Daarnaast heeft men goede instrumenten nodig om adverse events te detecteren en goede protocollen om te reageren op adverse events die ontdekt worden.

Het nauwgezet nakijken van alle medische dossiers wordt beschouwd als de techniek bij uitstek om adverse events te detecteren. Dit vergt echter uitgebreide middelen in termen van speciaal opgeleide medewerkers en van tijd. Het kunnen gebruiken van reeds beschikbare en min of meer gestandaardiseerde administratieve databanken kan een veel beter alternatief zijn met betrekking tot vereiste middelen en haalbaarheid.

Het doel van deze studie was om de accuraatheid van Belgische administratieve gegevens over ziekenhuisopnames te evalueren voor het detecteren van adverse events. In deze studie werd voor België pionierswerk verricht in dit domein en dit maakt dat de resultaten dan ook nog voorlopig zijn. We willen van de gelegenheid gebruik maken om de acht geselecteerde ziekenhuizen uitvoerig te bedanken voor hun wil om deel te nemen en voor het voorzien van de nodige data. Deze gegevens lieten toe eerste conclusies te trekken en gaven de richting aan voor verder onderzoek in dit gevoelige domein.

Gert Peeters

Adjunct Algemeen directeur a.i.

Jean Pierre Closon

Algemeen directeur a.i.

# Samenvatting

# INTRODUCTIE

Een adverse event [a] (AE) is een ongewild letsel of complicatie die resulteert in invaliditeit, overlijden of verlenging van het ziekenhuisverblijf van de patiënt, en dat eerder wordt veroorzaakt door het management van gezondheidszorg dan door de ziekte van de patiënt.

Op internationaal vlak wordt de prevalentie van adverse events in acute ziekenhuizen tussen 2,9% en 16,6% geschat, afhankelijk van het type adverse event. Dit beklemtoont nog de noodzaak om de zorgprocessen te verbeteren zodat het percentage complicaties in acute ziekenhuizen kan worden verminderd. Om dit doel te bereiken moeten we echter eerst een betrouwbaar proces definiëren om deze adverse events op te sporen. Aangezien diagnoses van gehospitaliseerde patiënten in administratieve databanken kunnen worden geïdentificeerd door middel van ICD-9-CM codes, vormen deze codes een erg goedkope en vlot toegankelijke bron van klinische informatie. In de loop der jaren werden databanken grondig bestudeerd om de validiteit van ICD-9-CM codes voor complicaties te beoordelen om de prestaties van de zorgverstrekkers te vergelijken. Omdat het gebruik van de Minimale Klinische Gegevens (Belgian Hospital Discharge Dataset B-HDDS; MKG) verplicht is voor alle gehospitaliseerde patiënten in acute ziekenhuizen kan het system als representatief worden beschouwd voor de zorgverlening in de Belgische acute ziekenhuizen. Dit gegevensbestand bevat demografische gegevens van de patiënt, informatie over het verblijf in het ziekenhuis (datum en soort opname/ontslag, gegevens over de verwijzing, opnameafdeling, bestemming na ontslag,…), evenals klinische informatie (primaire en secundaire diagnoses en therapeutische procedures zoals beschreven in de ICD-9-CM).

Het doel van deze studie was de nauwkeurigheid te toetsen van de Belgische administratieve gegevens over ziekenhuisverblijven (MKG) voor het detecteren van adverse events. Uit alle in de literatuur gevonden indicatoren werden er vijf geselecteerd: doorligwonden (decubitus), diepe veneuze trombose of longembolie (DVT/PE), postoperatieve sepsis, ventilator geassocieerde pneumonie (VAP) and postoperatieve wondinfectie. Deze keuze werd gemaakt op basis van een voldoende hoge prevalentie van het adverse event, en de beschikbaarheid van een duidelijke klinische definitie en coderingsalgoritme. Het is onwaarschijnlijk dat bijna-ongevallen in de medische dossiers kunnen worden aangetroffen of in de administratieve gegevens worden gecodeerd en daarom richtten we onze aandacht op adverse events. Omdat incidenten met geneesmiddelen niet in de administratieve gegevens worden gecodeerd, werden dergelijke indicatoren evenmin bestudeerd.

[a] De term “adverse event” is ook in de Nederlandstalige literatuur sterk ingeburgerd en wordt daarom in deze executive summary niet vertaald.

# METHODOLOGIE

Alle 116 Belgische acute ziekenhuizen werden uitgenodigd om aan deze studie deel te nemen. Dertien Vlaamse en acht Waalse acute ziekenhuizen stemden in om mee te werken. Acht ziekenhuizen werden geselecteerd op basis van de grootte van het ziekenhuis en hun geografische ligging.

Voor het selecteren van gevallen (ziekenhuisverblijven) werden administratieve gegevens over ontslagen voor het registratiejaar 2005 gebruikt. Om een voldoende aantal gevallen te verkrijgen, werd, waar nodig, een beroep gedaan op de bijkomende registratiejaren 2004 and 2006.

De vijf AE-indicatoren werden berekend aan de hand van de MKG van de geselecteerde ziekenhuizen. We selecteerden willekeurig 20 gevallen en 20 controles per adverse event voor elk ziekenhuis zodat we een totaal van 800 gevallen en 800 controlegevallen bekwamen. De matching van de controlegevallen gebeurde aan de hand van de APR-DRG, de ernst van de ziekte (severity of illness), leeftijd, geslacht, jaar en semester van registratie.

Omdat een aantal patiënten niet wensten deel te nemen of omdat enkele medische dossiers niet beschikbaar waren, werden in totaal 741 gevallen en 774 controlegevallen weerhouden in de studie.

Twee teams die elk uit twee zorgverstrekkers bestonden beoordeelden de medische dossiers. Eén team beoordeelde alle medische dossiers van vier ziekenhuizen, terwijl het andere team alle medische dossiers van de resterende vier ziekenhuizen beoordeelde. Medische dossiers werden gescreend met behulp van een data abstractie instrument bedoeld als een gestandaardiseerde methode voor het verzamelen van gegevens. In dit instrument werden strikte klinische criteria gehanteerd voor de 5 geselecteerde indicatoren, zowel gebaseerd op de literatuur als op de mening van de deskundigen.

Door de MKG gegevens en de screening van de medische dossiers te vergelijken, werd de positieve predictieve waarde (PPW) berekend voor de vijf indicatoren afzonderlijk. Vervolgens werden de gevallen uitgesloten waarvan tijdens de screening van het medische dossier bleek dat al een indicator aanwezig was op het moment van opname, en daarna werd de PPW opnieuw berekend. De MKG werd beschouwd als de testwaarde terwijl de screening van de medische dossiers werd beschouwd als de referentiewaarde.

Voor elk adverse event werd de verantwoordelijkheid van het gezondheidszorgmanagement en de vermijdbaarheid van het adverse event geëvalueerd op een schaal van 1 (geen) tot 6 (volledig) door een van de teams.

# RESULTATEN

De MKG maken onvoldoende onderscheid tussen een adverse event dat zich voordeed tijdens het verblijf in het ziekenhuis en een adverse event die al aanwezig was bij de opname. De indicator 'doorligwonden' bleek de meest gevoelige indicator te zijn voor het uitsluiten van aanwezigheid bij opname tijdens de screening van het medische dossier, met een lagere positieve predictieve waarde (PPW; 74% versus 68%) (zie tabel 1). Met uitzondering van VAP en doorligwonden worden alle andere indicatoren echter gedefinieerd als postoperatieve complicaties en het is dus minder waarschijnlijk dat ze al aanwezig zouden zijn bij de opname. VAP is, per definitie, nooit aanwezig bij opname en dus niet gevoelig voor deze correctie. Daarom verwijst de rest van dit hoofdstuk naar waarden waarbij aanwezigheid bij opname wordt uitgesloten bij de screening van het medische dossier.

Voor adverse events afwezig bij opname kunnen alle positieve predictieve waarden beschouwd worden als laag tot gemiddeld. Geen enkele indicator scoort goed genoeg wanneer vetrokken wordt van een minimale PPW van 75% zoals vooropgesteld door het Agency for Healthcare Research and Quality. De enige indicator die in de buurt komt, is de indicator voor doorligwonden met een PPW van 68%.

**Tabel 1 Positieve Predictieve Waarde per aanwezige (of niet aanwezige) indicator bij opname (Present On Admission - POA)**

| Indicator | PPV | 95% BI (onder-grens) | 95% BI (boven-grens) |
|---|---|---|---|
| Overall | 64.91 | 61.48 | 68.35 |
| Doorligwonden | 74.52 | 68.60 | 80.44 |
| Postoperatieve DVT/PE | 58.52 | 45.67 | 63.27 |
| Postoperatieve sepsis | 45.00 | 36.10 | 53.90 |
| Ventilator geassocieerde pneumonie | 29.94 | 23.19 | 36.69 |
| Postoperatieve wondinfectie | 69.06 | 62.99 | 75.13 |
| Overall, not POA | 61.37 | 57.69 | 65.05 |
| Doorligwonden, niet POA | 68.07 | 60.98 | 75.16 |
| Postoperatieve DVT/PE, niet POA | 54.47 | 41.19 | 58.06 |
| Postoperatieve sepsis, niet POA | 44.54 | 35.61 | 53.47 |
| Ventilator geassocieerde pneumonie, niet POA | 29.94 | 23.19 | 36.69 |
| Postoperatieve wondinfectie, niet POA | 66.83 | 60.43 | 73.23 |

Met betrekking tot de door de teams toegeschreven oorzaak van het adverse event werd een discrepantie gevonden tussen de resultaten van beide beoordelende teams. Het ene team beoordeelde dat de adverse events vaker de verantwoordelijkheid waren van het gezondheidszorgmanagement en vaker vermijdbaar waren dan het andere team.

# DISCUSSIE

In het algemeen adviseerden voorgaande studies tegen het algemeen gebruik van ICD-9-CM codes om het voorkomen van adverse events te meten. Waarschijnlijk liggen meerdere factoren aan de basis van de zwakke gegevens over adverse events in administratieve dossiers. Mogelijke oorzaken zijn o.m. het gebrek aan stimulansen om de AE's te coderen en het onvermogen of de terughoudendheid van clinici om complicatiediagnoses bij het ontslag van de patiënt te noteren. Tevens is het in de huidige ICD-9-CM niet altijd mogelijk om een aandoening voldoende specifiek te coderen om een onderscheid te kunnen maken tussen een adverse event en een normale complicatie. De ICD-9-CM code voor postoperatieve sepsis bijvoorbeeld maakt geen onderscheid tussen postoperatieve sepsis, postoperatieve hemorragische shock en postoperatieve cardiogene shock. Bovendien slaagt men er vaak niet in op basis van de administratieve database een onderscheid te maken tussen een toestand die al aanwezig was bij de opname, en een adverse event dat optreedt tijdens het verblijf in het ziekenhuis. De nieuwe versie van MKG, opgestart tijdens de tweede helft van 2007 en beschikbaar vanaf 2010, zal deze informatie wel bevatten.

In deze studie bleek uit een gedetailleerde herevaluatie van de medische dossiers per indicator dat onder-rapportering, over-rapportering, en strikte klinische criteria voor de evaluatie van adverse events in de medische dossiers de meest voorkomende verklaringen waren voor een discrepantie tussen de AE-indicator en het medische dossier.

In het licht van de resultaten over de verantwoordelijkheid van het gezondheidszorgmanagement en de vermijdbaarheid van de adverse events, lijkt het er op dat beide teams de vragen op een verschillende manier hebben geïnterpreteerd, waardoor de reproduceerbaarheid en validiteit van dit deel van de studie in vraag gesteld kan worden.

# AANBEVELINGEN

## CODERING

- Als een gecodeerd adverse event al bij opname aanwezig is, zouden de administratieve gegevens deze informatie moeten bevatten.
- Beschikbaarheid en het systematische gebruik van het volledige medische dossier, met inbegrip van het verpleegkundige dossier samen met de ontslagbrief, kan de MKG codering in theorie verbeteren. Een gestandaardiseerd medisch dossier kan de codering eveneens vergemakkelijken. Een meer nauwkeurige "vertaling" van het medische dossier in de MKG zou ook het probleem van onder-rapportering verminderen.
- De codering moet gebeuren op basis van een geïnformatiseerd patiëntdossier (zowel medisch als verpleegkundig dossier) inclusief variabelen die toelaten valide indicatoren te kwantificeren. Deze codering moet verplicht worden.
- Het classificatiesysteem dat bij administratieve gegevens wordt gebruikt, moet een voldoende fijne codering mogelijk maken. De overstap naar de ICD-10 zou hierin verbetering kunnen brengen indien de meer gedetailleerde codes die beschikbaar zijn, ook inderdaad worden gebruikt.
- Voor elke indicator zou een grondige vergelijking van de berekende indicator met een andere bron (bijv. het medische dossier) moeten worden uitgevoerd. De huidige grote verschillen in positieve predictieve waarde tussen verschillende AE-indicatoren sluiten veralgemening met gelijkaardig opgebouwde indicatoren uit.
- Een standaard algoritme in generieke vorm waarin de berekening volledig gedetailleerd wordt gedefinieerd, zou nationaal moeten beschikbaar gesteld worden evenals een voorbeeldimplementatie in ten minste één computertaal naar keuze. Deze algoritmen moeten bij voorkeur beschikbaar worden gesteld via de Federale Overheidsdienst Volksgezondheid, Veiligheid van de voedselketen en Milieu.

## GEBRUIK VAN DE INDICATOREN

- Geen enkele van de indicatoren die hier worden bestudeerd hebben een voldoende hoge positieve predictieve waarde voor benchmarking gebruik in hun huidige vorm. Benchmarking vereist een gepaste standaardisatie van referentiepercentages adverse events, gebruik makend van leeftijd en co-mobiditeitsfactoren, die op hun beurt sterk worden beïnvloed door de case-mix van het ziekenhuis. Anders zal de vergelijking tussen ziekenhuizen sterk worden gekleurd. Ook komt de MKG pas beschikbaar twee jaar na datum waardoor een snelle respons op mogelijke kwaliteitsproblemen die door adverse events worden veroorzaakt, wordt uitgesloten.

- Geen enkele van de bestudeerde indicatoren een voldoende hoge positieve predictieve waarde te hebben om ze te gebruiken als het enige instrument om adverse events binnen een ziekenhuis op te sporen. De aanwezigheid van het medische dossier in het ziekenhuis laat echter wel toe dat ze worden gebruikt als een eerste en relatief snel opsporingsmiddel in een ruimer programma om adverse events te voorkomen, op te sporen en hierop te reageren in het dagelijkse kwaliteitsmanagement van de ziekenhuiszorg. Rekening houdend met alle hierboven besproken beperkingen, kunnen ze ook worden gebruikt als een, zij het verre van perfect, follow-up middel binnen een ziekenhuis aangezien de case-mix in een ziekenhuis over verloop van tijd relatief stabiel blijft.

# ONDERZOEKSAGENDA

- Verder onderzoek over de prevalentie van adverse events zou zeer welkom zijn, daar er hierover weinig Belgische gegevens bestaan.
- Gezien de resultaten van deze studie lijkt een belangrijk probleem de relatief lage positieve predictieve waarde van de indicatoren voor adverse events te zijn. Verder onderzoek is nodig om de huidige algoritmen te verfijnen en te standaardiseren. Samen met onderzoek naar verbetering van de codering, moet dit onderzoek de positieve predictieve waarde van de indicatoren voor adverse events verhogen.
- Verder onderzoek is ook nodig over risicoverevening voor deze indicatoren. Slechts wanneer dit punt verduidelijkt is, en samen met (veel middelen vragend) onderzoek waarmee sensitiviteit en specificiteit berekend kunnen worden, kunnen deze indicatoren bruikbare instrumenten worden bij de zelfevaluatie van ziekenhuizen op gebied van adverse events.
- In deze studie werd een poging ondernomen de vermijdbaarheid van en de verantwoordelijkheid van het gezondheidszorgmanagement voor adverse events te beoordelen. Echter, gezien de resultaten, is verder onderzoek, vooral m.b.t. een gepaste methodologie, noodzakelijk.

# Scientific summary

Table of contents

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AHRQ | Agency for Healthcare Research and Quality |
| APR-DRG | All-Patient Refined Diagnosis Related Group |
| B-HDDS | Belgian hospital discharge dataset |
| CDC/HICPAC | Centers for Disease Control and Prevention / Healthcare Infection Control Practices Advisory Committee |
| CSP | Complications Screening Program |
| DRG | Diagnosis Related Group |
| DVT | Deep Vein Thrombosis |
| FTE | Full Time Equivalents |
| HCFA DRG | Health Care Financing Administration Diagnosis Related Group |
| ICD-9-CM | International Classification of Disease, 9th revision, Clinical modification, |
| ICU | Intensive Care Unit |
| IOM | Institute Of Medicine |
| JCAHO | Joint Commission on Accreditation of Healthcare Organizations |
| MDC | Major Diagnostic Categories |
| MKG/RCM | Minimale Klinische Gegevens/Résumé Clinique Minimum |
| NCC MERP | National Coordinating Council for Medication Error Reporting and Prevention |
| NPV | Negative Predictive Value |
| NQF | National Quality Forum |
| OECD | Organisation for Economic Co-operation and Development |
| ORP | Operating Room Procedure |
| PE | Pulmonary Embolism |
| POA | Present On Admission |
| PPV | Positive Predictive Value |
| PSI | Patient Safety Indicators |
| PWI | Postoperative Wound Infection |
| RN | Registered-Nurse |
| SOI | Severity Of Illness |
| UTI | Urinary Tract Infection |
| VA | Veterans Affairs |
| VAP | Ventilator-Acquired Pneumonia |
| VTE | Venous Thromboembolism |
| WHO | World Health Organization |

# 1 INTRODUCTION

Since the report "To Err is Human" by the Institute of Medicine (IOM) in 1999, attention was brought to the general public that adverse events in medicine are common and are one of the leading causes of morbidity and mortality within the United States. The report estimates that 44 000 – 98 000 patients hospitalised in the United States die each year as a result of medical errors. According to the report, between 3% and 4% of patients admitted to the hospital have adverse events resulting in injury or disability. About 30% of these adverse events are thought to be preventable and represent suboptimal care [1].

The effects of the IOM report were evident in at least 3 important areas. First, the IOM report profoundly changed the way many health care professionals and managers think and talk about medical errors and injury. Few individuals now doubt that preventable medical injuries are a serious problem. The concept that bad systems, not bad people, lead to the majority of errors and injuries, which is a crucial scientific foundation for improvement of safety in all successful high-hazard industries, has become a mantra in health care. It is much clearer now that the most effective method to improve either safety or quality overall is to change the systems[2]. In this regard, Longo et al defined "patient safety systems" as the various policies, procedures, technologies, services, and numerous interactions among them necessary for the proper functioning of hospital care[3]. Safety is a characteristic of systems and not of their components. Healthcare organizations must therefore develop a systems orientation to patient safety, rather than one that finds and attaches blame to individuals. For example, root cause analysis – a technique developed in industries that take a systems approach – examines in detail medical errors in an attempt to find the real cause of the problem rather than simply continuing to deal with its symptoms, and to remove the root problem so the situation does not occur again[3]. Following the IOM report, thirteen cases of medical errors were presented in the "Quality Grand Rounds: The Case for Patient Safety" in the hope that doing so might prevent another error[4].

The second major effect of the IOM report was to enlist a broad array of stakeholders to advance patient safety. The first stakeholder was the federal government but after only 3 years of support, federal funding for patient safety research through the Agency for Healthcare Research and Quality (AHRQ) became almost entirely earmarked toward studies of information technology. A host of nongovernmental organizations have made safety a priority. The Joint Commission on Accreditation of Healthcare Organizations (JCAHO) has led the way, tightening up accountability within health care organizations and requiring hospitals to implement new safe practices. Regional coalitions have sprung up across the country to facilitate stakeholders to work together to set goals, collect data, disseminate information, and provide education and training to improve safety. The most important stakeholders however who have been mobilized are the thousands of devoted physicians, nurses, therapists, and pharmacists at the ground level – in the hospitals and clinics – who have become much more alert to safety hazards. Most are making changes, not primarily in response to mandates, but rather to improve the quality of care for their patients.

The third effect of the IOM report was to accelerate the changes in practice needed to make health care safe. Initially, adoption of new safe practices was entirely voluntary. The JCAHO in 2003 required hospitals to implement 11 of a list of 30 evidence-based safe practices ready for implementation, including improving patient identification, communication, and surgical-site verification. Furthermore, a major practice change occurred in teaching hospitals in 2003 when all residency training programs implemented new residency training work hour limitations.

In spite of the growing patient safety movement however, health care isn't demonstrably and measurably safer[2, 5]. The premium placed on autonomy, the drive for productivity, and the economics of the system may lead to severe constraints and adverse medical events. The unusual degree of stress that health care workers experience derives from at least 4 factors.

First, health care is one of the few risk-prone areas in which public demand considerably constricts the application of common-sense safety-enhancing solutions, such as limiting the flow and choice of incoming patients. Second, health care is also one of the few risk-prone areas in which the system is extensively supported by novices, such as students, interns, and residents. Third, health care is one of the few risk-prone areas in which so many obvious sources of human error exist in the system, yet little has been done to reduce them. Sources of error include excessive fatigue on the job, systematic working of overtime, overloaded work schedules, and chronic shortage of staff. Finally, an endemic source of errors in medicine is the shifting of more clinical care and technology to the ambulatory setting. An important lesson from other industries is the move from training, regulation, and assessment of individuals to that of teams of health care providers. Given the interdisciplinary nature of health care and the need for cooperation among those who deliver it, teamwork is critical to ensuring patient safety and recovery from and mitigation of error[6].

Another key barrier to making progress is a paucity of measures. Identifying problems, measuring progress, and demonstrating that improvement has been achieved all depend on the availability of robust measures[2]. In this regard, Thomas and colleagues presented a conceptual model of commonly used methods for measuring latent errors, active errors and adverse events[7]. According to the author, latent errors include system defects such as poor design, incorrect installation, faulty maintenance, poor purchasing decisions, and inadequate staffing. These are difficult to measure because they occur over broad ranges of time and space and they may exist for days, months, or even years before they lead to a more apparent error or adverse event directly related to patient care. Active errors in contrast occur at the level of the frontline provider and are easier to measure because they are limited in time and space. Therefore, some measurement methods are best for latent errors and others for active errors although some methods are able to detect both of them. Table 1 shows the strengths and weaknesses of 8 measurement methods that have been used to measure errors and adverse events in health care[7].

**Table 1 Advantages and disadvantages of methods used to measure errors and adverse events in health care[7]**

| Error Measurement Method | Advantages | Disadvantages |
|---|---|---|
| Morbidity and mortality conferences and autopsy | Can suggest latent errors<br>Familiar to health care providers and required by accrediting groups | Hindsight bias<br>Reporting bias<br>Focused on diagnostic errors<br>Infrequently and non-randomly utilized |
| Malpractice claims analysis | Providers multiple perspectives<br>Can detect latent errors | Hindsight bias<br>Reporting bias<br>Non-standardized source of data |
| Error reporting systems | Can detect latent errors<br>Provide multiple perspectives over time<br>Can be a part of routine operations | Reporting bias<br>Hindsight bias |
| Administrative data analysis | Utilizes readily available data<br>Inexpensive | May rely upon incomplete and inaccurate data<br>The data are divorced from clinical context |
| Chart review | Utilizes readily available data<br>Commonly used | Judgements about adverse events not reliable<br>Expensive<br>Medical records are incomplete<br>Hindsight bias |
| Electronic medical record | Inexpensive after initial investment | Susceptible to programming and/or data entry errors |

| Error Measurement Method | Advantages | Disadvantages |
|---|---|---|
| | Monitors in real time<br>Integrates multiple data sources | Expensive to implement<br>Not good for detecting latent errors |
| Observation of patient care | Potentially accurate and precise<br>Provides data otherwise unavailable<br>Detects more active errors than other methods | Expensive<br>Difficult to train reliable observers<br>Potential Hawthorne effect<br>Potential concerns about confidentiality<br>Possible to be overwhelmed with information<br>Potential hindsight bias<br>Not good for detecting latent errors |
| Clinical surveillance | Potentially accurate and precise for adverse events | Expensive<br>Not good for detecting latent errors |

The model suggests that a comprehensive monitoring system for patient safety might include combinations of the discussed measurement methods[7].

Michel and colleagues compared the effectiveness, reliability, and acceptability of estimating rates of adverse events and rates of preventable adverse events using three methods[8] : cross sectional (data gathered in one day), prospective (data gathered during hospital stay), and retrospective (review of medical records). An adverse event was defined as an unintended injury caused by medical management rather than by a disease process and which resulted in death, life threatening illness, disability at time of discharge, admission to hospital, or prolongation of hospital stay. Preventable adverse events were those that would not have occurred if the patient had received ordinary standards of care appropriate for the time of the study. Table 2 provides an overview of advantages and disadvantages of the three methods used to estimate adverse events rates.

**Table 2 Advantages and disadvantages of three methods used to estimate adverse event rates [8]**

| Method | Advantages | Disadvantages |
|---|---|---|
| Prospective method | Best effectiveness for identifying preventable errors<br>Good reliability of judgment of iatrogenic nature of events<br>Staff sufficiently involved<br>Good appreciation of chain of events and their consequences | Most expensive<br>Heaviest workload |
| Cross sectional method | Least expensive<br>Rapid and easily renewed<br>May be sufficient to justify implementation of risk reduction policy<br>Good reliability of judgement of iatrogenic nature of events | Consequences of lack of follow up during patient's hospital stay<br>Excessive workload<br>Inadequate to serve as initial estimation |
| Retrospective | Good effectiveness<br>Almost no workload for staff<br>Data collection easily planned | Difficulty to judge iatrogenic and preventable nature on basis of sometimes piecemeal data<br>Lower face validity of results, especially for preventability judgment |

According to Lilford et al, medical record review is the only method for which there are a substantial number of published estimates of reliability[9]. Estimates of reliability however are usually not calculated in a way which allows comparison of studies or understanding the relative contribution of reviewers, their training, or the difficulty of the decision task. It is known that the more the heterogeneity in the raters and the conditions studied, the lower will be the reliability. The levels of sickness and fragility among patients make it difficult both to identify errors and to disentangle their effects from the progression of patients' underlying diseases. Moreover, intrinsic vagaries of judgment regarding errors in chart review exist, manifested in poor reliability among reviewers about what constituted adverse events and preventability[5]. Explicit methods of error detection – in which the quality of care is assessed against predetermined criteria – are likely to have much better interobserver agreement but also considerably less sensitivity than implicit methods which are based on expert judgement[9]. Therefore, we might expect some backing away from the notion of preventing accidental injury and more of a tilt toward effectiveness. Gains in effectiveness, including compliance with guidelines, are more readily measured and compared and should lead to demonstrable improvements in morbidity and mortality across populations[5].

Given the growing interest in the safety of patients, the development of accurate methods for measuring the frequency, severity and preventability of adverse events remains an important area in health services research [10] Methods for finding events have included spontaneous voluntary reporting, solicited voluntary reporting, direct observation of health care personnel during routine clinical meetings, computerized screening algorithms and retrospective chart review.

Medical records have so far been the primary source for researching medical errors and are considered to be the gold standard for monitoring adverse events [11] [12]. They contain rich clinical details that allow identification of various medical injuries and near misses and analysis of circumstances and causes of errors. Table 3 shows an overview on studies regarding adverse events performed in acute hospitals based on retrospective medical record review. A significant limitation of this system is that medical records are mostly in paper format or electronic format that is not readily usable for research. Transforming medical records into research data is expensive, resource intensive and requires exceptional knowledge and skills in medical context and research [11].

**Table 3 : Adverse event rates in acute hospitals based on retrospective medical record review**

| Publication year | Country and Region | Study Sample | Patients with Adverse Events |
|---|---|---|---|
| 1991[13] | USA, New York | 51 hospitals (n=30,195) | 3.7% |
| 1995[14] | Australia, New South Wales | 28 hospitals (n=14,189) | 16.6% |
| 1999[15] | USA, Utah and Colorado | 28 hospitals (n= 14,700) | 2.9% |
| 2001[16] | Denmark | 17 hospitals (n=1,097) | 9.0% |
| 2001[17] | England, Greater London area | 2 hospitals (n=1,014) | 10.8% |
| 2002[18] | New Zealand | 13 hospitals (n=6,579) | 12.9% |
| 2004[19] | Canada | 20 hospitals (n=3,745) | 10.6% |
| 2007[20] | Netherlands | 21 hospital (n=7,926) | 5.7% |

Administrative data are a viable source and their potential in patient safety research is increasingly recognized. One approach was the development of screening measures based on routinely collected administrative data such as the patient safety indicators (PSI). The most promising indicators for use as a screening tool were selected in order to provide an accessible and low-cost approach to identify potential problems in the quality of care related to patient safety[21].

Administrative data are readily available, inexpensive, computer readable, typically continuous, and often provide insight into the characteristics of large populations of patients [22] [12]. Nevertheless, ICD-9-CM were originally created to assist in describing the prevalence of major causes of morbidity and mortality worldwide and adapted for use in hospital reimbursement with the advent of prospective payment in 1982 and are now being used for purposes for which they were never intended. Lacking in detailed standard clinical definitions universally applied by medical record coders, the coding system is open to clinical and coding interpretation. Medical records coders are dependent to some extent on what is dictated in the discharge summary by the physician or surgeon to guide them in coding both active diagnoses that constitute patient comorbid conditions and postoperative adverse events [22]. Furthermore, incentives exist for complete coding of diagnoses and procedures by hospitals because greater levels of severity and complexity often are rewarded by higher levels of reimbursement [23]. As a result, the accuracy and reliability of these data in describing diagnoses, procedures, operations, characteristics of individual patients, and adverse events has been repeatedly questioned [22] [23].

At present, administrative data are increasingly used for the detection of adverse events. For instance, a retrospective analysis based on administrative data of all Belgian acute hospitals by Van den Heede et al revealed a prevalence of adverse outcomes of 7.12% in the medical and 6.32% in the surgical group [12]. These data highlights the importance for the development and implementation of processes aimed at reducing the incidence and impact of preventable adverse events since they are a major cause of morbidity and mortality.

The main objective of the present study is to assess whether the B-HDDS is a reliable source for the detection of 5 selected adverse events in acute Belgian hospitals. The screened events are pressure ulcer, deep vein thrombosis/pulmonary embolism, postoperative sepsis, ventilator-associated pneumonia and postoperative wound infection.

Firstly, this report will present a review of the literature on definitions and classifications of adverse events. The methodology used for validation will be exposed, and then the results will be detailed. Finally, those results will be discussed.

## *Research questions of the project*

- **On the basis of the international literature, how to define, classify and select the "adverse events". This part has two objectives: defining and classifying the adverse event concept (search 1) and listing candidate adverse event measures that can potentially be deducted from administrative databases (search 2). The scope of the research only includes in-hospital stays in acute hospitals, except for obstetrical adverse events.**
- **To translate the selected indicators from the literature review into algorithms, which allow the deduction from the Belgian administrative databases**
- **To validate the methodology of screening Belgian administrative databases for a selection of adverse event measures.**

# 2 LITERATURE REVIEW

## 2.1 DEFINITION AND CLASSIFICATION OF ADVERSE EVENTS

### 2.1.1 Methodology

A systematic review of the literature was performed in PubMed from 1990 to December 2006. Whenever possible MeSH terms were used. The following search terms were used : *medical errors*[a] *(MeSH)*; in combination with either of the terms *classification (MeSH)* or *definition*.

A total of 516 articles were retained initially. 494 articles were excluded based on the title of the article. Another 8 and 6 articles were excluded after reading the abstract or complete article respectively. The reasons for exclusion were: article not about adverse events, classification or definition; articles only related to nursing practice, hospital pharmacy, hospital laboratory, family practice, anatomic pathology or paediatric patients. A total of 8 articles were of particular interest and thus selected.

Five articles from references of the 8 selected articles were valuable for this part. Another 16 articles were brought in by experts on this matter.

### 2.1.2 Definition of Adverse Events

No universal definitions for descriptive terminology used within patient safety literature currently exist. This is one of the factors resulting in varying estimates of the prevalence of adverse events and medical errors [24 25 26].

#### *2.1.2.1 General definitions of adverse events*

According to Zhang et al [27], Reason's definition of human error is the most widely accepted: an error is a failure of achieving the intended outcome in a planned sequence of mental or physical activities. According to Reason, human errors are divided into two major categories: (1) slips that result from the incorrect execution of a correct action sequence and (2) mistakes that result from the correct execution of an incorrect action sequence. Furthermore, the human error problem can be viewed in two ways : the person approach and the system approach. The longstanding and widespread tradition of the person approach focuses on the unsafe acts – errors and procedural violations – of people at the sharp end : nurses, physicians, surgeons, anaesthetists, pharmacists, and the like. Followers of this approach tend to treat errors as moral issues, assuming that bad things happen to bad people. The basic premise in the system approach on the other hand is that humans are fallible and errors are to be expected, even in the best organisations. Errors are seen as consequences rather than causes, having their origins not so much in the perversity of human nature as in 'upstream' systemic factors. The person approach remains the dominant tradition in medicine. Nevertheless, the person approach has serious shortcomings and is ill suited to the medical domain. Indeed, continued adherence to this approach is likely to thwart the development of safer healthcare institutions. Another serious weakness of the person approach is that by focusing on the individual origins of error it isolates unsafe acts from their system context.

---

[a] As no MeSH term is available for 'adverse event', we used 'medical errors' a MeSH term whose scope covers "Errors or mistakes committed by health professionals which result in harm to the patient. They include errors in diagnosis (DIAGNOSTIC ERRORS), errors in the administration of drugs and other medications (MEDICATION ERRORS), errors in the performance of surgical procedures, in the use of other types of therapy, in the use of equipment, and in the interpretation of laboratory findings. Medical errors are differentiated from MALPRACTICE in that the former are regarded as honest mistakes or accidents while the latter is the result of negligence, reprehensible ignorance, or criminal intent"
http://ovidsp.tx.ovid.com/spb/ovidweb.cgi

Defences, barriers, and safeguards occupy a key position in the system approach. In an ideal world each defensive layer would be intact. In reality, however, they are more like slices of Swiss cheese, having many holes which are continually opening, shutting, and shifting their location. The presence of holes in any one 'slice' does not normally cause a bad outcome. Usually, this can happen only when the holes in many layers momentarily line up to permit a trajectory of accident opportunity. The holes in the defences arise for two reasons : active failures and latent conditions. Nearly all adverse events involve a combination of these two sets of factors. Active failures are the unsafe acts committed by people who are in direct contact with the patient or system. Latent conditions are the inevitable 'resident pathogens' within the system. They have two kinds of adverse effect : they can translate into error provoking conditions within the local workplace and they can create long lasting holes or weaknesses in the defences. Unlike active failures, latent conditions can be identified and remedied before an adverse event occurs[28].

The World Health Organization (WHO) defined an adverse event as an incident which results in harm to a patient . Harm implied impairment of structure or function of the body and/or any deleterious effect arising there from. Harm included disease, injury, suffering, disability and death and may thus be physical, social or psychological. A near miss was an incident that did not cause harm (also known as a close call). Finally, preventability has been defined as being accepted by the community as avoidable in the particular set of circumstances.

In the Harvard Medical Practice Study I," an adverse event was defined as an injury that was caused by medical management (rather than the underlying disease) and that prolonged the hospitalization, produced a disability at the time of discharge, or both". They defined negligence as care that fell below the standard expected of physicians in their community[13]. In the Harvard Medical Practice Study II [29], an adverse event was considered an operative complication if it occurred within the first two weeks after surgery or if it was thought to have been caused by the operation, regardless of when it occurred. Operative complications were sub-classified as technical, non-technical, related to wound infections, caused by surgical failure or late. Non-operative categories of injuries included those that were related to a procedure (which were further classified in the same manner as the operative complications), diagnostic mishaps, therapeutic mishaps, and those related to drugs.

The Institute of Medicine (IOM) defines adverse events as 'injuries caused by medical management rather than by underlying disease or condition of the patient' [12 24, 30 17 31 32]. In contrast to the Harvard Medical Practice Study, this definition of adverse event did not require prolongation of hospitalization or disability on discharge. A non-preventable adverse event is an unavoidable injury due to appropriate medical care. A preventable adverse event is an injury due to a non-intercepted serious error in medical care [31].

Furthermore, the IOM defines a medical error as 'the failure of a planned action to be completed as intended or the use of a wrong plan to achieve an aim'. A preventable adverse event is an adverse event that results from an error. Medical errors occur much more frequently than adverse events and medication errors outnumber adverse drug events by 100-1 [24].

A serious medical error is a medical error that causes harm (or injury) or has the potential to cause harm. It includes preventable adverse events, intercepted serious errors, and non-intercepted serious errors. It does not include trivial errors with little or no potential for harm to non-preventable adverse events. An intercepted serious error is a serious medical error that is caught before reaching the patient. A non-intercepted serious error is a serious medical error that is not caught and therefore reaches the patient but because of good fortune or because the patient had sufficient reserve to buffer the error, it did not cause clinically detectable harm [31].

Handler et al claim that emergency medicine should adopt the definitions that are consistent with the Institute of Medicine report 'To Err Is Human', the USP (U.S. Pharmacopeial Convention) taxonomy, and the major studies in the medical literature. As a result, they recommend the following definitions[33]:

| | |
|---|---|
| Error | failure of a planned action to be completed as intended (error of execution) or use of a wrong plan to achieve an aim (error of planning); the accumulation of errors results in accidents |
| Active error | an error that occurs at the level of the frontline operator and whose effects are felt almost immediately |
| Latent error | errors in the design organization, training, or maintenance that lead to operator errors and whose effects typically lie dormant in the system for lengthy periods of time |
| Slip errors | an error of execution when the action conducted was not what was intended; the wrong action is observable |
| Lapse errors | an error of execution when the action conducted was not what was intended; the wrong action is not observable |
| Mistake | an error in which the action proceeds as planned but fails to achieve its intended outcome because the planned action was wrong; error of planning |
| Accident | an event that involves damage to a defined system that disrupts the ongoing or future output of the system |
| Patient safety | freedom from accidental injury; ensuring patient safety involves the establishment of operational systems and processes that minimize the likelihood of errors and maximize the likelihood of intercepting them when they occur |
| Quality of care | degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge |
| Adverse event | an injury resulting from a medical intervention |
| Preventable adverse event | an injury that occurs as a result of medical error; with standard medical care the injury would not have occurred |
| Potential preventable adverse event ('near miss') | a medical error that could have resulted in injury |

McNutt et al specifically separate adverse events, failures and errors [34]. They define error only at the deepest reaches of the medical care system, because they are concerned that examining only adverse events and their proximate failures may not lead to lasting and significant change in the systems of care. In their model for medical failure adverse events can be caused by multiple failures that, in turn, can be caused by multiple errors interacting in complex ways.

In the research paper by Considine [35] an adverse event is defined as "an unintentional injury or complication resulting in disability, death or prolonged hospital stay that is a result of health care management rather than the patient's underlying disease". A preventable adverse event was defined in the Quality in Australian Health Care Study as an "error in (patient) management due to failure to follow accepted practice at an individual or system level".

Kellogg and Havens [36] reviewed the literature of adverse events. According to Walshe an adverse event is "a happening, incident, or set of circumstances which exhibits three key characteristics to some degree:

| | |
|---|---|
| Negativity | an event that by its very nature, is undesirable, untoward or detrimental to the health care process or to the patient |
| Patient involvement/impact | a continuum along which definitions of adverse events may fall. For instance, definitions at one end of this spectrum include events with potential but no actual negative patient impact, whereas definitions at the opposite end of the spectrum require an identifiable negative impact to the patient |
| Causation | event must be a result of the health care process, not of a patient's actions or the disease process itself |

A review on patient safety by Etchells and al [37] describes an error as "the failure of a planned action to be completed as intended (error of execution) or the use of the wrong plan to achieve an aim (error of planning)". A close call is an event that almost leads to patient harm but is avoided because of luck or timely interception. An adverse event, or complication, is "any unintended result of medical treatment that results in prolonged hospital stay, morbidity, or mortality; it may also be an injury caused by medical management rather than by the underlying condition of the patient". If an adverse event is cause by error(s), it is preventable.

Grober and Bohnen [38] reviewed the literature on defining medical error. Historically, patient safety researchers investigating the impact of error in medicine have adopted outcome-dependant definitions of medical error and its surrogate terms, and have limited their focus to patients experiencing adverse outcomes or injury as a consequence of medical care. Outcome-dependant definitions of medical error have provided valuable insight into the costs, morbidity and magnitude of harm resulting from such events. Nonetheless, quality improvement initiatives require understanding of the processes that lead to such errors. Therefore, according to the author, a definition of medical error should capture process or system failures that cause errors, irrespective of outcome (a process-dependant approach). Ideally, process-dependant definitions of medical error should capture the full spectrum of medical errors, namely, errors that result in adverse patient outcomes as well as those that expose patients to risk but do not result in injury or harm. Errors that do not result in injury are often referred to as near misses, close calls, potential adverse events or warning events. The authors propose the following outcome- and process-dependant definition of medical error: "an act of omission or commission in planning or execution that contributes or could contribute to an unintended result". This definition of medical error includes explicitly the key domains of error causation (omission and commission, planning and execution), and captures faulty processes that can and do lead to errors, whether adverse outcomes occur or not.

In a review of medical records in New South Wales and South Australia, Wilson [14] defined an adverse event 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 disease".

Guse [39] employed the definition of medical injury as "any untoward harm associated with a therapeutic or diagnostic health care intervention".

### *2.1.2.2 Function related definition*

Johnstone and Kanitsaki [40] concentrated on nursing errors as opposed to errors in general. Here, a nursing error is defined as "a discipline-specific term that encompasses an unintended 'mishap' made by a nurse and where a nurse is the one who is situated at the 'sharp end' of an event that adversely affected – or could have adversely affected – a patient's safety and quality care". In short, a nursing error is that in which a nurse stands as being the last causally and critically linked person to an unintended 'effect'.

### *2.1.2.3 Disability, causation and preventability definitions*

Disability was temporary or permanent impairment of physical function (including disfigurement) or mental function or prolonged hospital stay (even in the absence of such impairment). Temporary disability included adverse events from which complete recovery occurs within 12 months. Permanent disability included adverse events which caused permanent impairment or which resulted in permanent institutional or nursing care or death.

Causation was present if the adverse event was caused by health care management rather than the disease process. It included acts of omission (failure to diagnose or treat) and acts of commission (incorrect treatment or management). A scale from 1 – 6 was used to determine whether an adverse event was caused by health care management or the disease process [14 41, 42].

| | |
|---|---|
| 1 | virtually no evidence for management causation |
| 2 | slight-to-modest evidence for management causation |
| 3 | management causation not likely, less than 50-50 but close call |
| 4 | management causation more likely than not, more than 50-50 but close call |
| 5 | moderate/strong evidence for management causation |
| 6 | virtually certain evidence for management causation |

To determine the incidence and types of preventable adverse events in elderly patients, Thomas [41, 42] defined an adverse event as "an injury caused by medical management (rather than the disease process) that resulted in either prolonged hospital stay or disability at discharge". A confidence score of four or greater was required from the reviewing physician to indicate the presence of an adverse event. An adverse event was considered preventable if it was avoidable by any means currently available unless that means was not considered standard care. Davis [18] used the same operational definition in a study in New Zealand.

Preventability of an adverse event was assessed as "an error in management due to failure to follow accepted practice at an individual or system level"; accepted practice was taken to be 'the current level of expected performance for the average practitioner or system that manages the condition in question'. The degree of preventability was scored on a 1 – 6 scale, grouped into 3 categories [14] :

| | |
|---|---|
| **No preventability** | |
| 1 | virtually no evidence for preventability |
| **Low preventability** | |
| 2 | slight-to-modest evidence for preventability |
| 3 | preventability not likely, less than 50-50 but close call |
| **High preventability** | |
| 4 | preventability more likely than not, more than 50-50 but close call |
| 5 | strong evidence for preventability |
| 6 | virtually certain evidence for preventability |

### 2.1.3 Classification of Adverse Events

Consensus about specific methods for measuring quality remains elusive. Donabedian's classic framework[43] delineated 3 dimensions:

| | | |
|---|---|---|
| 1 | structure, or the characteristics of a health care setting | for example, the physical plant, available technology, staffing patterns, credentialing procedures and decision support system |
| 2 | process, or what is done to patients; inclusive appropriateness of services | errors of omission (failing to do necessary things), errors of commission (doing unnecessary things or doing them wrongly), errors of execution ( the failure of a planned action to be completed as intended) and errors of planning (use of a wrong plan to achieve an aim) : medical error, medication error, inappropriate drug prescription, near miss |
| 3 | outcomes, or how patients do after health care interventions | medical injury, adverse event, adverse drug event, iatrogenic illness, nosocomial infection, complication |

The 3 dimensions are intertwined, but their relative utility depends on context [44 45]. Outcomes that are not linked to specific medical practices provide little guidance for developing quality-improvement strategies. However, only a few links between processes and outcomes are backed by solid evidence from well-controlled studies. Furthermore, comparing outcomes across groups frequently requires adjustment for patient risk and the recognition that some patients are sicker than others [44]. Process measures are highly acceptable to providers because they demonstrate clearly how providers can improve their outcomes. Clinicians are also more accountable for the process of care than outcomes, which are affected by many other factors[46]. In general, there is considerable debate regarding whether quality measures should evaluate processes or outcomes of care[47]. One attraction of outcome measurement is that it is a measure of something that is important in its own right. Furthermore, outcome measurement will reflect all aspects of the processes of care and not simply those that are measurable or measured. Finally, data to construct simple rates are available from routine information systems[48].

An advantage of process measures is the ability to provide feedback for quality improvement initiatives. Secondly, most process measures require less risk adjustment for patient illness than do most outcome measures. Thirdly, process measures can usually be collected more quickly than outcome data[47]. Fourtly, process measures are more sensitive than outcome measure to differences in the quality of care and they are easy to interpret[48]. On the other hand, there are several disadvantages to process measures. Firstly, to be valid, there must be a strong relationship between the process and outcome measures. This relationship may be weak or non-existent for many processes even when they are truly linked to outcomes. Secondly, demonstrating the link between process and outcome is prohibitively expensive and often impossible to achieve for any one organization. Thirdly, while providers may care about process measures, patients and non-clinicians generally place little value on them. Fourthly, most feasible process measures are usually indicators for a very specific element of the care process rather than comprehensive measures of how care is delivered[47]. Whatever health care quality measure is used, it is imperative that the measures are meaningful, scientifically sound, generalizable, and interpretable. In order to achieve this, Rubin et al proposed steps and issues in developing and testing process-based measures of health care quality. According to the author, initial steps required to develop process measures will include: (1) defining the audience and the purpose of measurement; (2) choosing the clinical area to evaluate; (3) organizing the measurement team; (4) selecting the process criterion; (5) writing the measure specifications; (6) performing preliminary tests; and (7) developing scoring and analytical specifications[46].

Reason [49] claims that cognitive factors are critical at various levels of the healthcare system hierarchy of medical errors. At the lowest core level, it is individuals who trigger errors. At the next level, errors can occur due to interactions between an individual and technology. This is an issue of human-computer interaction where cognitive properties of interactions between human and technology affect and sometimes determine human behaviour. At the next level, errors can be attributed to the social dynamics of interactions between groups of people who interact with complex technology in a distributed cognitive system. This is the issue of distributed cognition and computer-supported cooperative work. At the next few levels up, errors can be attributed to factors of organizational structures (e.g. coordination, communications, standardization of work process), institutional functions (e.g. policies and guidelines), and national regulations. In this system hierarchy of human errors in medicine, it is clear that individuals are at the last stage of the chain, although the individuals may not be the root cause of the error. If the chain of events can be stopped at the individual's stage through cognitive interventions, errors could be potentially prevented. Zhang claims that the cognitive theory of human action most appropriate for medical errors is the seven-stage action theory developed by Norman and refined by Zhang and colleagues. According to this theory, any action has seven stages of activities: (1) establishing the goal; (2) forming the intention; (3) specifying the action specification; (4) executing the action; (5) perceiving the system state; (6) interpreting the state; and (7) evaluating the system state with respect to the goals and intentions. Errors can occur at any of the seven stages of action and between any two adjacent stages : due to incorrect translation from goals to intentions, incorrect action specifications from intentions, incorrect execution of actions, misperception of system state, misinterpretation of data perceived, and misevaluation of interpreted information with regard to the goal of the task.

Chang et al developed and applied a method of classification that was based on evaluations of extant taxonomies and reporting systems with feedback from individuals who would use the taxonomy[50]. Their review of the literature reinforced the fact that various approaches used in the health care sector to define and classify near misses, adverse events, and other patient safety concepts have generally been fragmented. Homogeneous elements of previous models were categorized into five complementary root nodes, or primary classifications.

1. Impact – the outcome or effects of medical error and systems failure, commonly reffered to as harm to the patient.
2. Type – the implied or visible processes that were faulty or failed.
3. Domain – the characteristics of the setting in which an incident occurred and the type of individuals involved.

4. Cause – the factors and agents that led to an accident.
5. Prevention and mitigation – the measures taken or proposed to reduce incidence and effects of adverse occurrences.

The 'Impact' classification comprised three subclassifications that could discriminate between 18 types of outcomes or effects (harm). The harm index was based on the NCC-MERP Medication Error Taxonomy[51]. The 'Type' classification included three levels that address communication, patient management, and clinical performance. The 'Domain' classification included the types of health care professionals commonly involved in patient care and the demographics of patients in a variety of health care settings where events might have occurred. The principal nodes of the 'Cause' classification comprised two subclassifications : system (process/structure) failures and human failures. Finally, three types of 'Prevention and mitigation' were identified : universal, selective, and indicated. The 'universal' subclassification covered preventive and corrective measures that are designed for everyone in the eligible population. Prevention and mitigation measures that are directed toward a subgroup of the population whose risk of adverse evetns is above average were grouped in the 'selective' subclassification. Lastly, the 'indicated' subclassification combined interventions that are targeted to specific high-risk individuals identified as having a minimal but detectable risk for sustaining an adverse event[50].

According to Etchells and collaborators [37], most preventable adverse events are not only the result of human error but are due to defective systems that allow errors to occur or go undetected. Therefore, a reasonable approach is to break the causes down into organizational factors, situational factors, team factors, individual factors, task factors and patient factors.

| | |
|---|---|
| Organizational factors | adequate personnel and equipment, scheduling and timing of procedures, substitution of usual team members with new members |
| Situational factors | distractions, interruptions, physical conditions and equipment design, including monitors and displays |
| Team factors | communication, confidence in team members and the ability to deal with unexpected events |
| Individual factors | mental readiness, technical performance and fatigue |
| Task factors | relate to the clarity of the task at hand, including clear protocols and accurate available information; they are important causes of drug events |
| Patient factors | obesity, anatomic variation, disease severity and co-morbidity |

Johnstone and Kanitsaki[40] used 8 categories of nursing errors as described by Benner et al. Taxonomy of nursing errors is the following:

| | Examples |
|---|---|
| Lack of attentiveness | missed predictable complications, such as a postoperative haemorrhage |
| Lack of agency/fiduciary concern | failure to advocate for the patient's best interests/failure to question a doctor's inappropriate directives |
| Inappropriate judgement | failure to recognise the implications of a patient's signs and symptoms |
| Medication error | wrong drug, wrong route, wrong amount |
| Lack of intervention on the patient's behalf | failure to follow up on signs of hypovolemic shock |
| Lack of prevention | failure to prevent threats to patient safety such as via breaches of infection control precautions |
| Missed or mistaken doctor/health care provider's orders | carrying out inappropriate orders/mistaking orders, resulting in an erroneous intervention |
| Documentation errors | charting procedures or medications before they were completed/failure to chart observations |

Others[52 53] described 5 categories of harm based on the National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP[51]) classifications of errors:

| | | |
|---|---|---|
| Category A | no error | circumstances or events that have the capacity to cause error |
| Category B | error, no harm | an error occurred but the error did not reach the patient |
| Category C | error, no harm | an error occurred that reached the patient bud did not cause patient harm |
| Category D | error, no harm | an error occurred that reached the patient and required monitoring to confirm that it resulted in no harm to the patient and/or required intervention to preclude harm |
| Category E | | harm that contributed to or resulted in temporary harm to the patient and required intervention |
| Category F | | harm that contributed to or resulted in temporary harm to the patient and required initial or prolonged hospitalization |
| Category G | | harm that contributed to or resulted in permanent patient harm |
| Category H | | harm that required intervention to sustain life |
| Category I | | harm that contributed to or resulted in the death of a patient |

As stated before, preventability of an adverse event was assessed as "an error in management due to failure to follow accepted practice at an individual or system level". The degree of preventability was scored on a 1 – 6 scale, grouped into 3 categories[14]:

### 2.1.4 Current selection

Although 'near misses' are important regarding overall quality improvement, an injury which doesn't cause harm or no clinically detectable harm will unlikely be found in the medical record, nor will it be coded in administrative data. The same can be concluded for latent errors. Therefore, only adverse events which resulted in disability or prolongation of hospital stay were retained. In this regard, for the purpose of comparing the results on administrative data with those on medical record review, the definition of an adverse event by Wilson et al[14] was the most complete in evaluating the adverse events apart from detecting them in the medical records. According to Wilson, three conditions have to be met in order to conclude to an adverse event :

1. an unintended injury or complication which
2. results in disability, death or prolongation of hospital stay,
3. is caused by health care management rather than the patient's disease

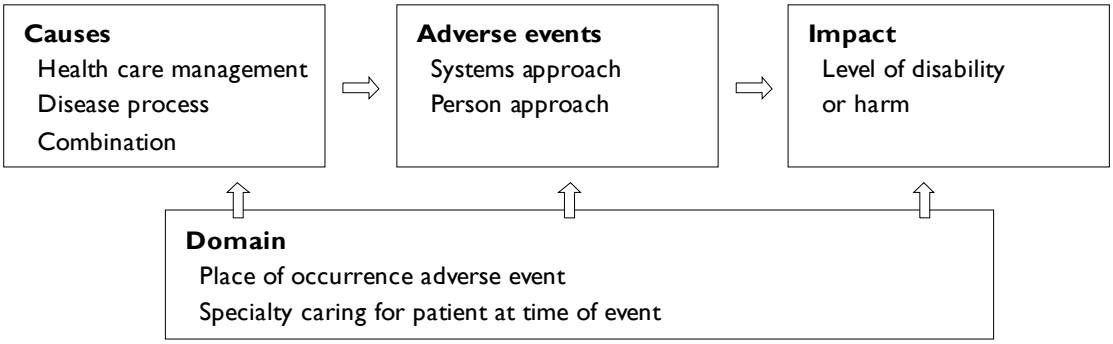
The 6-point Likert scale for causation and preventability will also be used [14].

Although the classification by the NCC (MERP)[51] was used for medication error reporting, it also can be used to describe categories of harm for error in general:

| | |
|---|---|
| Category E | harm that contributed to or resulted in temporary harm to the patient and required intervention |
| Category F | harm that contributed to or resulted in temporary harm to the patient and required initial or prolonged hospitalization |
| Category G | harm that contributed to or resulted in permanent patient harm |
| Category H | harm that required intervention to sustain life |
| Category I | harm that contributed to or resulted in the death of a patient |

Based on an analytical framework of the JCAHO patient safety event taxonomy, the linkages (Figure 1) between defining adverse events and their appreciation on causation and harm provides an organized approach to guide the retrospective process of identifying the factors (causes) that contribute to adverse events leading to a certain disability or harm (impact) to the patient in a certain domain of health care[50].

**Figure 1 Connection between adverse events – chart review- classification**



### 2.1.5 Summary of definitions and classifications of adverse events

**Table 4 Summary of definitions and classifications of adverse events in the literature**

| Type | Author | Definition |
|---|---|---|
| General definition | Reason | *Human error* : failure of achieving the intended outcome in a planned sequence of mental or physical activities<br>- Slips : incorrect execution of a correct action sequence<br>- Mistakes : correct execution of an incorrect action sequence<br>Human error as person approach : unsafe acts – errors and procedural violations – of people at the sharp end<br>Human error as system approach : errors as consequences rather than causes,having their origins not so much in the perversity of human nature as in 'upstream' systemic factors |
| | WHO | Adverse event : incident which results in harm to a patient, including disease, injury, suffering, disability and death<br>Near miss : incident that did not cause harm (close call)<br>Preventability : being accepted by the community as avoidable in the particular set of circumstances |
| | Brennan | *Adverse event* : an injury that was caused by medical management (rather thand the underlying disease) and that prolonged the hospitalization, produced a disability at the time of discharge, or both<br>*Negligence* : care that fell below the standard expected of physicians in the community |
| | IOM | *Adverse event* : injuries caused by medical management rather than by underlying disease or condition of the patient<br>- Non-preventable adverse event : unavoidable injury due to appropriate medical care<br>- Preventable adverse event : injury due to a non-intercepted serious error in medical care<br>*Medical error* : the failure of a planned action to be completed as intended or the use of a wrong plan to achieve an aim<br>- Serious medical error : medical error that causes harm (or injury) or has the potential to cause harm; includes preventable adverse events, intercepted serious errors, and non-intercepted serious errors |
| | Wilson | *Adverse event* : 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 disease |
| | Kellogg | *Adverse event* : a happening, incident, or set of circumstances which exhibits 3 key characteristics to some degree :<br>1) Negativity : undesirable, untoward or detrimental to the health care process or to the patient<br>2) Patient involvement/impact<br>3) Causation : event must be a result of the health care process |
| Function related definition | Johnstone | *Nursing error* : discipline-specific term; an unintended 'mishap' made by a nurse and where a nurse is the one who is situated at the 'sharp end' of an event that adversely affected - or could have adversely affected - a patient's safety and quality care |

| Type | Author | Definition |
|---|---|---|
| Disability | Wilson | *Disability* : temporary or permanent impairment of physical or mental function or prolonged hospital stay<br>- Temporary disability : adverse events with complete recovery within 12 months<br>- Permanent disability : adverse events which caused permanent impairment or which resulted in permanent institutional or nursing care or death |
| Causation | Wilson | *Causation* : adverse event caused by health care management rather than the disease process<br>- Acts of omission : failure to diagnose or treat<br>- Acts of commission : incorrect treatment or management |
| Preventability | Wilson | *Preventability* : an error in management due to failure to follow accepted practice at an individual or system level |
| Classification adverse events | Donabedian | 3 dimensions :<br>1) Structure : characteristics of a health care setting<br>2) Process : what is done to patients<br>3) Outcomes : how patients do after health care interventions |
| | Zhang | Cognitive factors : critical at various levels of the healthcare system hierarchy of medical errors<br>- Individuals triggering errors<br>- Errors due to interactions between an individual and technology<br>- Distrubed systems : interactions among individuals and interactions between groups of people and technology<br>- Organizational structures : coordination, communication, and standardization of work process, skills, input and output<br>- Institutional funcitons : policy, guidelines<br>- National regulations |
| | Etchells | Defective systems :<br>- Organizational factors<br>- Situational factors<br>- Team factors<br>- Individual factors<br>- Task factors<br>- Patient factors |
| | Johnstone | 8 categories of nursing errors :<br>- Lack of attentiveness<br>- Lack of agency/fiduciary concern<br>- Inappropriate judgement<br>- Medication error<br>- Lack of intervention on the patient's behalf |

| Type | Author | Definition |
|---|---|---|
| | | - Lack of prevention<br>- Missed or mistaken doctor/health care provider's orders<br>- Documentation errors |
| | NCC (MERP) | 5 categories of harm :<br>- Category E : harm that contributed to or resulted in temporary harm to the patient and required intervention<br>- Category F : harm that contributed to or resulted in temporary harm to the patient and required initial or prolonged hospitalization<br>- Category G : harm that contributed to or resulted in permanent patient harm<br>- Category H : harm that required intervention to sustain life<br>- Category I : harm that contributed to or resulted in the death of a patient |

## 2.2 INDICATORS IN THE LITERATURE

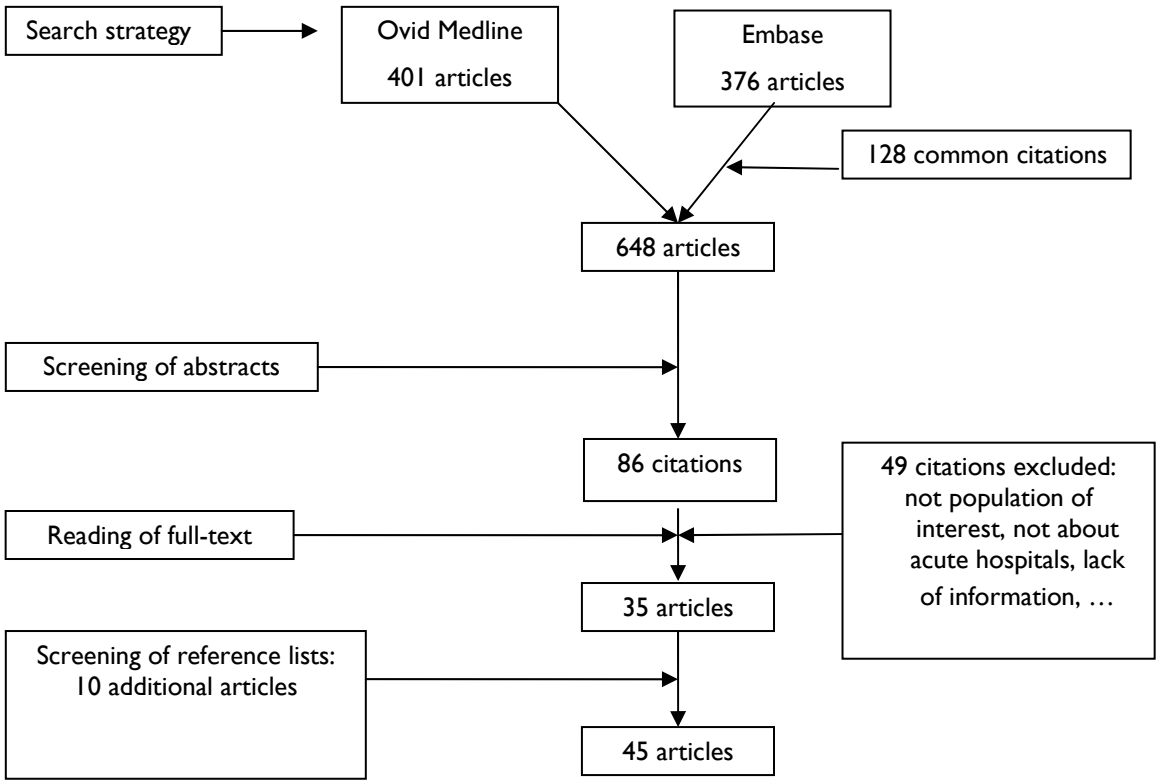
### 2.2.1 Methodology

A systematic review of the literature was performed in Ovid MEDLINE and EMBASE from 1990 to the first week of February 2007. Whenever possible, MeSH terms were used. The final formula used as search strategy was: ("administrative data" OR ""international classification of diseases" [MeSH] OR "ICD-9-CM") AND ("medical errors" [MeSH] OR "safety management" [MeSH] OR "outcome assessment (Health Care)" [MeSH] OR complication$ OR safety OR "adverse event" OR "adverse events" OR "iatrogenic disease" [MeSH]). The search was limited to articles written in English, French or Dutch.

The purpose of the study was to validate some "safety" indicators which can be obtained from administrative data available in hospitals. To be included, articles have to provide information about indicator purpose and construction in sufficient detail. Disease- and condition-specific papers without general importance were not included. We also excluded articles dealing with children or obstetric patients who are very specific populations or articles addressing only long term care or psychiatric facilities. Finally, articles describing ICD-9-CM algorithms were preferred because administrative data in Belgium use this international classification.

A total of 401 articles were obtained by searching on PubMed and 376 articles by searching on EMBASE (**Figure 2**). 128 citations were found in both databases. On the basis of the titles or abstracts, 86 articles were selected. After reading the full text of these articles, we kept 35 articles of interest. The screening of the reference lists provided another 10 articles for inclusion.

**Figure 2 Results of literature search in the indexed literature**



### 2.2.2 Results of the literature search

Many indicators of adverse events in acute hospitals were found in the literature. We selected 5 indicators for the medical record reviewing because it is beyond the scope of this study to evaluate all of them. The selection of indicators was based on several criteria: in order of importance, a sufficiently high prevalence, a clear clinical definition, coding validity in literature and expert opinion regarding the national and international literature.

The prevalence of events was found in previous works on Belgian administrative data [12] [54] or in the international literature. The prevalence limit for selecting an indicator was 5 cases per 1000 population at risk. An event must be found frequently enough to obtain a sufficient amount cases in the participating hospitals (see second part of this study).

It was also important that a clear definition of the selected adverse events was provided, and then identify clinical criteria describing the event in a similar fashion for the different reviewers. Clear clinical criteria have to be found in order to minimize possible interrater variability.

Validity and specificity of the indicators must have been analysed in the international literature and be sufficiently frequent to be considered.

Chosen adverse events were part of indicators selected for patient safety at the health systems level in OECD countries [55]. This project presents the consensus recommendations of an international expert panel on indicators for patient safety.

For this study, five indicators were selected on the basis of their prevalence, the availability of clinical criteria definition and their validity in the literature:

1. Postoperative pulmonary embolism or deep vein thrombosis
2. Postoperative sepsis
3. Decubitus ulcer
4. Ventilator-acquired pneumonia
5. Postoperative wound infection

Later in this chapter, we propose a summary as well as detailed evidence for each of the five selected indicators. For each of them, detailed evidence will be the definition of the algorithm, the prevalence of the event found in the literature and a summary of the literature in terms of importance of the event, coding validity, construct validity. Finally, we will describe the different sources of the indicator and present a table that summarizes the algorithm used in this project.

Thereafter, general table aims at summarising the evidence for most of the indicators found in the literature and not retained in this project. More detailed evidence for those indicators which is described in the appendix 1. All indicators related to the area obstetrics and pediatrics were not discussed in detail since this population was excluded from the actual retrospective study (see further). Indicators regarding medical equipment and technical difficulty were also not mentioned since these indicators might be less related to human error. Finally, all indicators related to 'adverse drug events', although often occurring, were not discussed further since detailed administrative data are not available for these events.

Accidental puncture or laceration, complication of anesthesia, foreign body left in during a procedure, iatrogenic pneumothorax, infection due to medical care, postoperative abdominopelvic wound dehiscence, postoperative hemorrhage or hematoma, postoperative hip fracture, postoperative physiologic and metabolic derangements, postoperative respiratory failure and transfusion reaction were events too rare to be selected. We did not find a clearly defined prevalence for aspiration pneumonia, postoperative infection (except wound and pneumonia) and reopening of surgical site.

Clinical criteria may be ambiguous for accidental puncture or laceration, aspiration pneumonia, complication of anesthesia, gastrointestinal hemorrhage, infection due to medical care, postoperative physiologic and metabolic derangements, postoperative respiratory failure, shock or cardiac arrest and urinary tract infection.

### 2.2.2.1 *Decubitus ulcer*

**Summary of the evidence**

| Indicator | Prevalence (/1000 at risk) | Source | Coding validity | Construct validity | Clinical criteria |
|---|---|---|---|---|---|
| Decubitus ulcer | Belgium[54]: 15.39<br><br>ARHQ[56] : 22.66<br>Zhan[11]: 21.51<br>Romano[57]: 21.30<br>Coffey[58]: 23.89<br>Rosen[59]: 18.36<br><br>Belgium[12]: 12.2 (Surgical) - 12.4 (Medical)<br><br>Needleman[60]: 58 (Surgical) - 72 (Medical)<br><br>Smith[61]: 29.86 | PSI<br><br>Needleman<br><br>Other | No evidence on validity is available from CSP studies | Weak association between PSI[59]<br><br>No association between staffing and rate[60]<br><br>Inversed relation between hours of care delivered and rate[62]<br><br>Nursing mix related to rate[63] | Available |

### Definition of indicator Decubitus Ulcer

*Numerator.* Discharges with ICD-9-CM code of decubitus ulcer in any secondary diagnosis field amongst the cases meeting the inclusion and exclusion rules for the denominator.

*Denominator.* All medical and surgical discharges age 16 years and older. Include only patients with a length of stay of more than four days. Exclude patients in MDC9 (skin, subcutaneous tissue, and breast), patients with a primary diagnosis of decubitus ulcer or with any diagnosis of hemiplegia, paraplegia, or quadriplegia. Exclude patients with a spina bifida or an anoxic brain damage and patients with a debridement or a pedicle graft before or on the same day as the major operating room procedure (surgical cases only). Exclude patients admitted from long-term care facility or transferred from an acute care facility and all obstetric admissions (MDC 14: pregnancy, childbirth, and puerperium)

### International prevalence figure in literature

On the basis of PSI's algorithm, previous study in Belgium estimated 15.39 decubitus ulcer per 1,000 discharges at risk between 1999 and 2004 [54]. Zhan and colleagues found 21.51 patients with decubitus ulcer per 1,000 discharges at risk [11]. Taking race and ethnicity into account, Coffey et al found a rate of 23.89 per 1,000 discharges [58], which according to the authors, is a higher rate for blacks and Hispanics in comparison to the white population. Rosen et al obtained a risk-adjusted rate of 18.36 per 1,000 eligible discharges [59].

Needleman et al found an adverse outcome rate of 7.2% and 5.8% in medical and surgical discharges, respectively [60]. In Belgium, Van den Heede and colleagues used the same definition and found a crude adverse outcome rate per 1,000 discharges of 12.4 for medical patients and 12.2 for surgical patients [12].

Rothschild et al reviewed preventable medical injuries in older patients [64]. They found that, amongst the high-risk hospitalised patients, the incidence of decubitus ranged up to 30%. A review by Smith describes a national average amongst at-risk Medicare patients of 29.86 per 1,000 hospital admissions [61].

### Summary of the indicator

This indicator is intended to flag cases of in-hospital decubitus ulcers. It was developed as part of the Complications Screening Program (CSP), which consists of a computer algorithm that screens for potential complications of adult discharge abstracts [65]. While several authors included in their definition of Decubitus cellulitus [65] [60] [63] [12], it was omitted from the PSI. Decubitus ulcer was investigated as an adverse outcome potentially sensitive to nurse staffing levels by Needleman et al [60] and Lichtig et al [63].

In order to separate conditions that arise during the hospitalization from those present on admission, debubitus ulcer labelled as the primary diagnosis was excluded from the denominator. Also excluded were patients who were particularly vulnerable to decubitus ulcer, namely patients with major skin disorders (MDC 9), patients admitted from a long term care facility or another acute care and patients with any form of paralysis. Finally, the indicator excludes patients that have a length of stay less than five days since it would be unlikely for a decubitus ulcer to develop within this period of time.

**Literature review/evidence levels**

*Importance of the indicator.* For the experts' panel of the OECD [55], a decubitus ulcer in a hospitalized patient has a serious negative impact on the individual's health and often leads to a significantly prolonged hospital stay. The economic impact of extended hospital stays makes this indicator important for financial improvement. Decubitus ulcers are preventable with good quality nursing care. It is a common complication of inadequate care for immobilized patients. This indicator has great clinical relevance as a patient safety measure.

*Coding validity.* The committee of OECD estimated that the biggest threat to construct validity is the inability to precisely distinguish, on the basis of administrative data, with decubitus ulcer present on admission or hospital-acquired [55].

After reviewing safety initiatives in the health systems of the United Kingdom, Canada, Australia and the United States, Arah et al found conflicting evidence for the validity of this indicator [66].

*Construct validity.* No evidence on validity is available from CSP studies.

Needleman et al found no association between the measures of registered-nurse staffing and pressure ulcers amongst the medical patients [60]. In contrast, Blegen et al concluded that the proportion of hours of care delivered on a patient care unit by registered nurses was inversely related to the unit rates of decubitus [62]. However, the total hours of patient care was associated with higher rates of decubitus.

The total hours of care was determined to a great extent by the average acuity level of patients on these units. Given the high correlation between acuity and total nursing care hours, the interpretation of these coefficients must be done with care. Another authors concluded that for all four data sets nursing skill mix was related to lower pressure ulcer rates [63]. Each additional percentage point of nursing personnel that were registered nurses was associated with a reduction in the pressure ulcer rate of between 0.79% and 1.77%. In two of the four data sets, additional hours of nursing per nursing intensity weight were significantly related to lower rates of pressure ulcers. The results also show that hospitals in large urban areas had higher rates of pressure ulcers. Conversely, teaching hospitals in California were found to have a more than 35% lower pressure ulcer rate than other hospitals.

McCloskey et al examined in a retrospective study the effects of New Zealand's Health Reengineering on nursing and patient outcomes [67]. During that period the combined registered nurses and enrolled nurses full time equivalents (FTEs) decreased with 36%, as did hours worked per 1,000 medical/surgical discharges. By 2000, there was an 18% increase in skill mix, with registered nurses labour representing 93% of nursing FTEs and hours worked by medical and surgical nurses. There were statistically significant increases in the rates for decubitus ulcers after reengineering 1993 implementation with an increase of 88% for medical discharges and an increase of 258% in the surgical group.

Rosen, who implemented the PSI software on Veterans Health Administration data, concluded that additional evidence was provided of PSIs having good construct validity [59]. Although correlations amongst the indicators were generally weak, these finding suggested that each indicator most likely reflects a unique dimension of quality.

Murff et al determined whether an association existed between patient complaints and surgical complications [68]. They found no statistically significant difference in complaint categories between patients who experienced a decubitus ulcer and those who did not. Major complications occurred in 19.2% of surgical admissions associated with a patient complaint and in 12.5% admissions not associated with complaints. Surgical admissions associated with a complication had an odds ratio of 1.74 of being associated with a patient complaint. This relationship remained significant after adjusting for patient length of stay, patient age, co-morbid illness, surgical sub-speciality and patient race.

Mattke et al evaluated the impact of alternative definitions of exclusion rules for defining patient samples used to construct measures of patient outcomes sensitive to nurse staffing in in-patient units of acute care hospitals [69].

Relaxing the length of stay restriction added a large number of patients with lower risk. Thus, the rate in the full surgical and medical samples fell from 1.97% to 1.09% and 4.00% to 2.48% respectively. These findings provide evidence that the patient groups affected by the exclusion rules have different clinical characteristics and thus a different propensity to experience hospital-acquired complications.

Romano et al determined how accurately postoperative complications are reported in administrative data, whether accuracy varies systematically across hospitals, and whether serious complications were more consistently reported amongst the adults who underwent elective lumbar diskectomies [70] . The sensitivity of reporting for all complications was < 35%, the specificity was 98%, the positive predictive value was 82% and the negative predictive value was 84%.

Geraci et al confirmed only 2 of 9 episodes of pressure ulcers reported on discharge abstracts of Veterans Affairs (VA) patients hospitalized in 1978-89 for congestive heart failure, chronic obstructive pulmonary disease or diabetes [71]. The sensitivity for an iatrogenic ulcer was 40%. Berlowitz et al found that the sensitivity of a discharge diagnosis of pressure ulcer amongst all patients transferred from VA hospitals to VA nursing homes in 1996 was 31% overall, or 54% for deep ulcers [72]. The overall sensitivity increased modestly since 1992 and was slightly but statistically significantly better amongst the medical patients than amongst the surgical patients (33% versus 26%).

**Sources**

This indicator was originally proposed in 1992 by Iezzoni et al. as part of the CSP [65]. It's also retained as Patient Safety Indicator from AHRQ [56]. Needleman et al. identified decubitus ulcer as an 'outcome potentially sensitive to nursing' [60].

**Specification of numerator/denominator of Decubitus Ulcer**

| | |
|---|---|
| Numerator | Discharges, 16 years and older, with ICD-9-CM code of decubitus ulcer in any secondary diagnosis field amongst cases meeting the inclusion and exclusion rules for the denominator |
| Denominator | All medical and surgical discharges age 16 years and older defined by specific APR-DRGs<br><br>Exclude cases:<br>- with length of stay of less than 5 days<br>- with primary diagnosis of decubitus ulcer<br>See Appendix 2.B Decubitus Ulcer<br>- MDC 9 (Skin, Subcutaneous Tissue, and Breast)<br>- MDC 14 (pregnancy, childbirth, and puerperium)<br>- with any diagnosis of hemiplegia, paraplegia, or quadriplegia<br>See Appendix 2.C Hemiplegia, Paraplegia, or Quadriplegia<br>- with ICD-9-CM code of spina bifida or anoxic brain damage<br>See Appendix 2.D Spina Bifida or Anoxic Brain Damage<br>- with an ICD-9-CM procedure code for debridement or pedicle graft before or on the same day as the first operating room procedure (surgical cases only)<br>See Appendix 2.E Procedure code for debridement or pedicle graft<br>- admitted from a long-term care facility<br>- transferred from an acute care facility |

### 2.2.2.2 *Postoperative pulmonary embolism (PE) or deep vein thrombosis (DVT)*

**Summary of the evidence**

| Indicator | Prevalence (/1000 at risk) | Source | Coding validity | Construct validity | Clinical criteria |
|---|---|---|---|---|---|
| Deep Vein Thrombosis - Pulmonary Embolism (DVT/PE) | *Postoperative Pulmonary Embolism or Deep Vein Thrombosis (PSI)*<br>Belgium[54]: 5.41<br><br>ARHQ[56] : 9.83<br>Zhan[11]: 9.34<br>Romano[57]: 9.19<br>Coffey[58]: 9.0<br>Rosen[59]: 10.62<br><br>Belgium[12]: 3.39 (Surgical) - 6.25 (Medical)<br><br>Needleman[60]: 4 (Surgical) - 5 (Medical)<br><br>Smith[61]: 13.14<br>White[73]: 8 | PSI<br><br>Needleman [60]<br><br>other | *DVT/PE (CSP)* [74 75 76 77]<br><u>Surgical cases</u> (n = 41 ) :<br>PPV : 89.6%<br>NPV : 98.1%<br>Present on admission : 22%<br>Complication confirmed by reabstraction : 88 %<br>Confirmed cases (timing&complication) by reabstraction : 59 %<br>Complication confirmed by physicians : 70 %<br>Complication confirmed by physicians with at least one potential quality problem : 60.7%<br>At least one process problem : 72.2%<br><br><u>Medical cases</u> (n = 53 ) :<br>PPV : 75.7%<br>NPV : 98.5%<br>Present on admission : 57%<br>Complication confirmed by reabstraction: 78%<br>Confirmed cases (timing&complication) by reabstraction: 32%<br>Complication confirmed by physicians: 28.2%<br>Complication confirmed by physicians with at least one potential quality problem: 60.0%<br>At least one process problem : 69.1%<br><br>Sensitivity <35%, specificity 98%, PPV 82%, NPV 84% [70] | *DVT/PE (CSP)* [74]<br><u>Surgical cases</u> (n = 36 ) :<br>clinical evidence : 67%<br>physician notes : 8%<br>no evidence : 25%<br><br><u>Medical cases</u> (n = 42 ) :<br>clinical evidence : 55%<br>physician notes : 12%<br>no evidence : 33%<br><br>Nurse staffing is independent of the occurrence of DVT/PE [60]<br><br>Inverse relation between registered-nurse hours and non-RN hours with rate (surgery) [78]<br><br>Weak association between PSI [59]<br><br>No statistically significant difference in complaint categories between patients who experienced an adverse event and those who did not [68] | Available |

**Definition of indicator Postoperative DVT/PE**

Numerator. Discharges with ICD-9-CM codes for deep vein thrombosis (DVT) or pulmonary embolism (PE) in any secondary diagnosis field amongst cases meeting the inclusion and exclusion rules for the denominator.

Denominator. All surgical discharges age 16 and older defined by surgical APR-DRG and an ICD-9-CM code for an operating room procedure. Exclude cases : with principal diagnosis of deep vein thrombosis or pulmonary embolism; where the interruption of vena cava occurs before or on the same day as the first operating room procedure; all obstetric admissions (MDC 14 : pregnancy, childbirth and puerperium).

**International prevalence figure in literature**

Different studies applied the AHRQ definition of postoperative pulmonary embolism or deep vein thrombosis on several populations. Rates ranged from 9.00 [58] to 10.62 [59] per 1,000 discharges at risk. In Belgium, this rate was 5.41 per 1,000 discharges at risk between 1999 and 2004 [54]. According to some authors, the rate is higher for African American and Non-Hispanics in comparison to the white population [58].

Needleman et al found pulmonary embolism or deep vein thrombosis in 0.5% of medical patients and in 0.4% of surgical patients [60]. With the same definition, the crude adverse outcome rate per 1,000 discharges in Belgium was 6.25 in medical patients and 3.39 in surgical patients [12].

The complication rate per 1,000 amongst at-risk Medicare patients hospitalized during 2000-2002 was 13.14 according to Smith [61]. A retrospective study by White et al found 13,533 cases diagnosed with venous thromboembolism (VTE); 6,005 cases with VTE diagnosed during the index hospitalization and 7,528 cases after discharge but within 91 days of the day of surgery, an overall incidence of 0.8% [73]. In an editorial, Kearon et al[79] concluded that malignancy was a convincing risk factor for post-operative VTE with about a 70% increase of risk. Previous VTE 'stands out' as the single most potent risk factor for post-operative VTE (about 7-fold higher than malignancy).

**Summary of the indicator**

Iezzoni and colleagues introduced this indicator as part of the Complications Screening Program [65]. Rates of complications for individual hospitals were calculated using 'risk pools' as the population denominators. Risk pools identified those patients at risk for specific types of complications. Defined by DRGs or ICD-9-CM procedure codes, risk pools were of 6 types: major surgery (A); minor and miscellaneous surgery (B); invasive cardiology and radiology procedures (C); endoscopy (D); medical patients (E) and complications applicable for all patients (F). The screen DVT/PE was assigned to all risk pools in the study. Studies done by Weingart et al [75], McCarthy et al [74], Van den Heede et al [12] and Needleman et al [60] also included surgical and medical discharges in their denominator. Adverse outcome rates were separately measured for surgical and medical patients.

The PSI indicator from AHRQ limited the denominator to surgical discharges only.

All studies excluded patients with a principal diagnosis of DVT or PE. All obstetric admissions were also excluded, as well as patients with a secondary procedure code 38.7 ("Interruption of vena cava") when this procedure occurred on the day of or previous to the day of the principal procedure.

White et al used a broader definition for the numerator and exclusion criteria [73 80].

**Literature review/evidence levels**

*Importance of the indicator.* Panel of experts from OCDE reported about the importance of this indicator[55]. The symptoms of postoperative PE/DVT ranged from mild to devastating clinical consequences including pain, respiratory distress and death. It causes unnecessary prolongation of hospital stay, unnecessary pain, suffering and death. This event can be prevented through the appropriate use of anticoagulants and other preventive measures.

The panel estimated this indicator has important financial and quality improvement implications[55]. Since the recent implementation of medication fee in Belgium, appropriate use of anticoagulants and other preventive measures are important to supervise.

*Coding validity*. For the purpose of identifying in-hospital events, the indicator has better validity for surgical cases than for medical cases in a study by Lawthers et al on Medicare beneficiaries of 65 years of age or older [76]. With a positive predictive value (PPV) of 89.6%, Lawthers concluded that this indicator was a good-to-excellent candidate as screening for complications in the major surgical risk pool. The negative predictive value (NPV) was 98.1%. The PPV and NPV in the medical risk pool were 75.7% and 98.5%, respectively.

Arnason et al measured the accuracy of a broader range of ICD-9-CM codes for thromboembolism using a random sample of patients discharged in a tertiary care hospital in Ottawa, Canada[81]. Compared to the gold standard chart review, the ICD-9-CM codes were, in general, sensitive (97%) but not specific (75%). This resulted in a positive predictive value of only 63%. The negative predictive value was 98%. By selecting a sub-group of ICD-9-CM codes for thromboembolism, the positive predictive value increased to 87%.

According to Arah et al, there is conflicting validity evidence for this indicator [66]. Finally, panel expert of OECD reported that coding of those events should be unambiguous, but PE/DVT is known to frequently go undiagnosed [55].

*Construct validity*. The validation study by Lawthers et al [76] was performed in two different states and was limited to the major surgical and medical risk pools. In the major surgical risk pool, the proportion of cases with trigger codes corroborating on record review was 88% versus 78% in the medical risk pool. The overall proportion of cases confirmed as in-hospital events was 59% for the major surgical risk pool and only 32% for the medical risk pool. In cases flagged for PE/DVT, the diagnosis appeared to be present on admission in 22% of the major surgical risk pool and in 57% in the medical risk pool. Weingart found similar figures [75]: physicians confirmed the flagged CSP screen in 70.0% of surgical and in 28.2% of medical cases. With this, DVT/PE is an event with one of the highest rates of confirmed complication and potential quality problem in the study.

Using the California Patient Discharge Data Set and specific ICD-9-CM surgical procedures codes, White et al found a complex relationship between age and the incidence of thromboembolism that varied with the surgical procedure[80]. Advancing age was a significant predictor for VTE following surgeries performed for conditions not inherently associated with significant co-morbidity. This study suggests that the type of surgery or perhaps the underlying pathology associated with the specific surgical procedure may override or eliminate age as a risk factor for thromboembolism.

McCloskey et al examined in a retrospective study the effects of New Zealand's Health Reengineering on nursing and patient outcomes [67]. The rate for DVT/PE initially increased by 9% in the medical group and by 91% in the surgical group, but later returned to rates near or below pre-reengineering levels.

Needleman et al found that nurse staffing was independent of the occurrence of DVT/PE amongst both major surgical or medical patients [60]. However, Kovner et al reported that having more registered nurse hours and non-RN hours was associated with a lower rate of DVT/PE after major surgery [78].

McCarthy et al created objective, explicit chart review instruments itemizing key clinical criteria confirming coded diagnoses [74]. Consensus on clinical indicators was reached through discussion with the other clinicians. Only confirmatory clinical criteria that were supported by the literature were included, although the literature was limited for certain conditions. In this study, medical records contained no clinical evidence or physicians' notes to support the coded condition in 25% of surgical cases and in 33.3% of medical cases. Objective clinical evidence was present in 66.7% and in 54.8% of surgical and medical cases respectively.

In 8.3% of surgical cases and 11.9% of medical cases, only physician notes supported the condition but had no specific objective clinical evidence to confirm the complication (Table 5).

**Table 5 Presence of clinical factors confirming a complication of deep vein thrombosis and pulmonary embolism (surgical cases n=36)[74]**

| Clinical factors | Presence of Clinical Factor, n (%) | Type of Clinical Evidence, n (%) |
|---|---|---|
| New pulmonary embolism based on abnormal pulmonary | 12 (33.3) | … |
| arteriogram or high-probability VQ scan | | |
| New deep venous thrombosis based on abnormal unilateral | 10 (27.8) | … |
| impedance plethysmography, Doppler flow velocity ultrasound | | |
| (above knee/popliteal fossa), duplex (real-time B-mode) | | |
| ultrasound, positive venogram (above knee/popliteal fossa), or | | |
| positive MRI | | |
| New pulmonary embolism based on moderate-probability VQ | 5 (13.9) | … |
| scan, physician diagnosis of pulmonary embolism, or clinical | | |
| signs highly suggestive of pulmonary embolism | | |
| Had at least 1 objective clinical factor | … | 24 (66.7) |
| Physician note but no objective clinical factor | … | 3 (8.3) |
| No clinical factor or objective physician note | … | 9 (25.0) |

VQ indicates ventilation perfusion; MRI, magnetic resonance imaging.

One of the most frequent PSI events in the study by Rosen et al was postoperative DVT/PE [59]. The authors concluded that additional evidence was provided that the PSI had good construct validity.

Murff et al determined whether an association existed between patient complaints and surgical complications [68]. They found no statistically significant difference in complaint categories between patients who experienced a DVT/PE and those who did not. Surgical admissions associated with a complication had an odds ratio of 1.74 of being associated with a patient complaint. This relationship remained significant after adjusting for patient length of stay, patient age, co-morbid illness, surgical sub-speciality and patient race.

Explicit process of care failures were relatively frequent amongst both major surgical and medical cases with DVT/PE, respectively in 72% and 69% of cases in which event was not present at admission[77].

Romano et al determined how accurately postoperative complications are reported in administrative data, whether accuracy varies systematically across hospitals, and whether serious complications are more consistently reported [70] amongst the patients who underwent elective lumbar diskectomies. The sensitivity of reported complications was <35%, the specificity was 98%, the positive predictive value was 82% and the negative predictive value was 84%.

**Sources**

This indicator was originally proposed by Iezzoni et al as part of the Complications Screening Program [65]. It is one of the AHRQ's Patient Safety Indicators [56]. Needleman assessed the indicator DVT/PE as an outcome which is potentially sensitive to the extent or quality of nursing care [60].

| **Specification of numerator/denominator of Postoperative DVT/PE** | |
|---|---|
| Numerator | Discharges amongst the cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM codes for deep vein thrombosis or pulmonary embolism in any secondary diagnosis field.<br>See Appendix 2.F Pulmonary Embolism/Deep Vein Thrombosis |
| Denominator | All surgical discharges age 16 and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure.<br>See Appendix 2.A Operating Room Procedure<br>Surgical Discharge APR-DRGs: procedural APR-DRGs<br>Exclude cases:<br>- with principal diagnosis of deep vein thrombosis or pulmonary embolism<br>See Appendix 2.F Pulmonary Embolism/Deep Vein Thrombosis<br>- where a procedure for interruption of vena cava is the only operating room procedure<br>*ICD-9-CM Interruption Of Vena Cava procedure code:*<br>387 INTERRUPTION OF VENA CAVA<br>- where a procedure for interruption of vena cava occurs before or on the same day as the first operating room procedure<br>- MDC 14 (pregnancy, childbirth, and puerperium) |

### 2.2.2.3 *Postoperative Sepsis*

**Summary of the evidence**

| Indicator | Prevalence (/1000 at risk) | Source | Coding validity | Construct validity | Clinical criteria |
|---|---|---|---|---|---|
| Postoperative sepsis | Belgium[54]: 14.30<br><br>ARHQ[56] : 11.8<br>Zhan[11]: 11.25<br>Romano[57]: 10.91<br>Coffey[58]: 12.12<br>Rosen[59]: 6.62<br><br>Belgium[12] : 13.2 (Surgical) - 13.8 (Medical)<br><br>Needleman[60]:<br>10 (Surgical) - 13 (Medical)<br><br>Smith[61]: 13.2 | PSI<br><br>Needleman [60]<br><br>other | No evidence | No association between nurse staffing and complication [60]<br><br>Sepsis is an outcome sensitive to nursing [67]<br><br>Weak association between PSI [59]<br><br>No statistically significant difference in complaint categories between patients who experienced an adverse event and those who did not [68] | Available |

### Definition of indicator Postoperative Sepsis

*Numerator.* Discharges with ICD-9-CM codes for sepsis in any secondary diagnosis field amongst cases meeting the inclusion and exclusion rules for the denominator.

*Denominator.* All elective surgical discharges age 16 and older with a length of stay of more than three days. Exclude cases: with principal diagnosis of sepsis or infection, with any code for immunocompromised state or cancer, all obstetric admissions (MDC 14 : pregnancy, childbirth, and puerperium).

### International prevalence figure in literature

Different studies applied the AHRQ definition of postoperative sepsis on several populations. In Belgium, postoperative sepsis between 1999 and 2004 was 14.30 per 1,000 discharges at risk [54]. Zhan and colleagues [11] found 11.25 events per 1,000 discharges at risk. Taking race and ethnicity into account, Coffey et al found a rate of 12.12 per 1,000 discharges [58]. The authors concluded that each of the minority groups (African American, non-Hispanic; Hispanic; Asian and Pacific Islander) had higher rates of postoperative sepsis compared to the white population. Rosen et al implemented the PSI on the Veterans Health Administration and found a risk-adjusted rate of 6.62 per 1,000 eligible discharges [59].

Needleman et al made a distinction between medical and surgical patients and found an adverse outcome rate of 1.3% and 1.0% respectively [60]. In Belgium, Van den Heede and colleagues used the same definition and found a crude adverse outcome rate per 1,000 discharges of 13.8 for medical patients and 13.2 for surgical patients [12].

### Summary of the indicator

Iezzoni and colleagues introduced this indicator as part of the Complications Screening Program using the risk pools 'major surgery' and 'minor or miscellaneous surgery' as the population denominators [65]. AHRQ limited PSI to postoperative sepsis. Studies done by Van den Heede et al [12] and Needleman et al [60] also included surgical and medical discharges in their denominator.

In order to better differentiate with sepsis present on admission, the indicator limits its definition to secondary diagnosis. Furthermore, patients who are particularly susceptible for sepsis are excluded, namely patients in immunocompromised state due to any cause and cancer patients. Patients with a length of stay less than four days are also excluded since it is unlikely that hospital acquired sepsis will develop in such a short time. Patients who had expired during their admission were eliminated in the study performed by Murff et al as this outcome would have reduced the opportunity to generate a complaint in the construct of their study [68].

## Literature review/evidence levels

*Importance of the indicator*. For panel of experts from OECD [55], the occurrence of sepsis following surgery is a severe complication with a mortality rate of up to 30%. Even less severe cases will require prolonged ICU treatment for organ failure. For this committee, many cases of postoperative sepsis can be prevented, primarily through a reduction of hospital infection rates, and they estimated it is a good measure of quality. It is also relevant to cost containment as prolonged hospital stays due to postoperative sepsis have considerable economic impact. Many cases of postoperative sepsis can be prevented through the appropriate use of prophylactic antibiotics, good surgical site preparation, careful and sterile surgical techniques and appropriate postoperative care.

*Coding validity*. Considering the dramatic nature of this complication, it is usually reliably coded in administrative data sources [55].

*Construct validity.* For the panel experts of OECD [55], sepsis after elective surgery is considered a severe complication. It usually results from less severe infective complications, such as urinary tract infections, pneumonia and wound infection, which should be avoided and/or properly treated.

Needleman found no association between the measures of registered-nurse staffing and the occurrence of sepsis amongst both surgical and medical patients [60]. Sepsis in medical and surgical cases was one of the nurse sensitive clinical outcomes in the study by McCloskey [67]. There was a statistically significant increase in the rate for sepsis after the reengineering's 1993 implementation, respectively 95% for medical discharges and 172% in the surgical group.

Rosen, who implemented the PSI software on Veterans Health Administration data, concluded that additional evidence was provided of PSI having good construct validity [59]. Although correlations amongst the indicators were generally weak, these findings suggested that each indicator most likely reflects a unique dimension of quality.

Murff et al determined whether an association existed between patient complaints and surgical complications using administrative data [68]. They found no statistically significant difference in complaint categories between patients who experienced septicaemia and those who did not. Surgical admissions associated with a complication had an odds ratio of 1.74 of being associated with a patient complaint. This relationship remained significant after adjusting for patient length of stay, patient age, co-morbid illness, surgical sub-speciality and patient race.

After reviewing safety initiatives in the health systems of the United Kingdom, Canada, Australia and the United States, Arah et al concluded that there was an unclear construct validity for this indicator [66].

Romano et al determined how accurately postoperative complications were reported in administrative data, whether accuracy varies systematically across hospitals, and whether serious complications were more consistently reported amongst patients who underwent elective lumbar diskectomies. The sensitivity of reporting for all complications was <35%, the specificity was 98%, the positive predictive value was 82% and the negative predictive value was 84% [70].

**Sources**

This indicator was originally proposed by Iezzoni et al as part of the Complications Screening Program [65]. It is a part of Patient Safety indicators set from the AHRQ[56]. Needleman identified this adverse outcome for both medical and surgical patients as an 'outcome potentially sensitive to staffing by nurses' [60].

**Specification of numerator/denominator of Postoperative Sepsis**

| | |
|---|---|
| Numerator | Discharges amongst cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM code for sepsis in any secondary diagnosis field.<br>See Appendix 2.G Sepsis |
| Denominator | All elective* surgical discharges age 16 and older defined by specific APR-DRGs and an ICD-9-CM code for an operating room procedure.<br>See Appendix 2.A Operating Room Procedure<br>Surgical Discharge APR-DRGs: procedural APR-DRGs<br>*Elective - Admission type is recorded as elective<br><br>Exclude cases:<br>- with principal diagnosis of sepsis or infection<br>See Appendix 2.G Sepsis<br>- with any code for immunocompromised state or cancer<br>See Appendix 2.K Immunocompromised States<br>See Appendix 2.L Cancer<br>See Appendix 2.N Cancer APR-DRGs<br>- MDC 14 (pregnancy, childbirth, and puerperium)<br>- with length of stay of less than 4 days |

### 2.2.2.4 *Postoperative wound infection (PWI)*

**Summary of the evidence**

| Indicator | Prevalence (/1000 at risk) | Source | Coding validity | Construct validity | Clinical criteria |
|---|---|---|---|---|---|
| Postsurgical wound infection | Belgium[12] : 7.92 (Surgical)<br><br>Needleman[60]: 8 (Surgical) | Needleman [60] | *Postoperative wound infection (CSP)* [75 76]<br>Surgical cases (n = 44 ) :<br>PPV : 91.5%<br>NPV : 95.0%<br>Present on admission : 0%<br>Complication confirmed by reabstraction : 93%<br>Confirmed cases (timing&complication) by reabstraction : 91%<br>Complication confirmed by physicians : 60.5 %<br>Complication confirmed by physicians with at least one potential quality problem : 26.9%<br><br>Medical cases (n = 39 ) :<br>PPV : 80%<br>NPV : 99.3%<br>Present on admission : 68%<br>Complication confirmed by reabstraction : 95%<br>Confirmed cases (timing&complication) by reabstraction : 28%<br>Complication confirmed by physicians : 24.2%<br>Complication confirmed by physicians with at least one potential quality problem : 0%<br><br>Sensitivity <35%, specificity 98%, PPV 82%, NPV 84% [70] | *Postoperative wound infection (CSP)* [74]<br>Surgical cases (n = 41 ) :<br>clinical evidence : 63.4%<br>physician notes : 0%<br>no evidence :36.6%<br><br>Postoperative wound infection is an outcome sensitive to nursing [67]<br><br>No statistically significant difference in complaint categories between patients who experienced an adverse event and those who did not [68] | Available |

### Definition of indicator Postoperative Wound Infection

For all surgical discharges age 16 years and older, discharges with ICD-9-CM code of wound infection in any secondary diagnosis field (exclude case with principal diagnosis of wound infection)

### International prevalence figure in literature

In Belgium, on the basis of the Needleman's definition, Van den Heede and colleagues found a crude adverse outcome of 7.92 per 1,000 surgical discharges [12] Needleman et al found a mean adverse outcome rate of 8 per 1,000 surgical patients [60].

### Summary of the indicator

This indicator is intended to flag cases of postoperative wound infection. The indicator was developed as part of the Complications Screening Program (CSP)[65]. This indicator is limited to surgical discharges only, aged 16 year and older. In order to separate conditions that arise during the hospitalization from those present on admission, wound infection labelled as the primary diagnosis was excluded from the denominator.

### Literature review/evidence levels

*Importance of the indicator*. For OECD committee [55], symptoms of a wound infection ranged from minor insignificant inflammation to considerable pain and suffering, wound disruption, septicaemia and even death. This event often required re-operation and prolonged hospitalisation. Occurrence of such an infection can be reduced by proper pre-, intra- and post-operative care, in particular strict hygiene. Hospital staff tends to neglect simple measures like hand washing and use of disinfectants. Various clinical processes are linked to wound infections and, considering the economic impacts of this event, it is important to reduce it.

*Coding validity*. Some studies[76] [75] [74] performed validation on CSP indicator on Medicare beneficiaries age 65 or older. For the purpose of identifying in-hospital events, the indicator has better validity in surgical cases than in medical cases in a study by Lawthers et al[76]. With a positive predictive value (PPV) of 91.5%, the authors concluded that this indicator was a good-to-excellent candidate as screening for complications in the major surgical risk pool. The negative predictive value (NPV) was 95.0%. The PPV and NPV in the medical risk pool were respectively 80% and 99.3%.

*Construct validity*. The validation study by Lawthers [76] was performed in 2 states, Connecticut and California, using Medicare's fiscal year 1994. In the major surgical risk pool cases with trigger codes corroborating on record review was 93% versus 95% in the medical risk pool. The overall proportion of cases confirmed as in-hospital events was 91% for the major surgical risk pool but only 28% for the medical risk pool. In cases flagged for the indicator, the diagnosis appeared to be present on admission in 0% of the major surgical risk pool and in 68% in the medical risk pool. In the study of Weingart et al[75], physicians confirmed the flagged CSP indicator in 60.5% of surgical and in 24.2% of medical cases. Amongst cases with confirmed in-hospital complications, physician reviewers identified at least one potential quality problem in 26.9% of surgical and 0.0% of medical flagged cases. The prevalence of physician-identified potential quality problems amongst flagged cases was only 25.6% in surgical and 3.0% in medical cases. The author concluded that this CSP would be a poor quality-of-care indicator.

McCarthy et al created objective, explicit chart review instruments itemizing key clinical criteria confirming coded diagnoses[74]. Consensus on clinical indicators was reached through discussion with other clinicians. Only confirmatory clinical criteria that were supported by the literature were included, although the literature was limited for certain conditions. The clinical criteria for surgical wound infection were evaluated in 41 surgical cases. Medical records contained objective clinical evidence in 63.4% of surgical cases. No clinical evidence or physicians' notes to support the coded condition were present in 36.6% of surgical cases.

In 0.0% of surgical cases, only physician notes supported the condition but had no specific objective clinical evidence to confirm the complication.

**Table 6 Presence of clinical factors confirming a complication of Postoperative wound infection (n=26) (Unpublished, obtained from E McCarthy)**

| Clinical factor confirming complication (n = 26) | (%) |
|---|---|
| Incisional infection as evidenced by superficial drainage and positive gram stain for white blood cells | 4 (15.4) |
| Incisional infection as evidenced by documentation or red (erythema) and hot or swollen and painful incision site, and clinician note of purulent drainage of infection site | 9 (34.6) |
| Incisional infection as evidenced by superficial drainage, positive gram stain for white blood cells, and clinician note of purulent drainage of infection site | 8 (30.8) |
| Incisional infection as evidenced by documentation of red (erythema) and hot or swollen and painful incision site, and fever, leukocytosis, or left shift | 11 (42.3) |
| Deep infection as evidenced by drainage and positive gram stain for wbc | 6 (23.1) |
| Deep infection as evidenced by fever, leukocytosis, or left shift and x-ray, CT scan or ultrasound evidence of abscess at anatomical site of surgical incision | 5 (19.2) |
| Deep infection as evidenced by creptitus in the wound on physical exam or x-ray, CT scan, or ultrasound evidence of gas at anatomical site or surgical incision, and documentation of red (erythema) and hot or swollen and painful incision site with fever | 0 (0.00) |

Murff et al determined whether an association existed between patient complaints and surgical complications [68]. They found no statistically significant difference in complaint categories between patients who experienced a wound infection and those who did not. Surgical admissions associated with a complication had an odds ratio of 1.74 of being associated with a patient complaint. This relationship remained significant after adjusting for patient length of stay, patient age, co-morbid illness, surgical sub-speciality and patient race.

Surgical wound infection was one of the nurse sensitive clinical outcomes in the study performed by McCloskey et al [67]. There was a statistically significant increase in the rate for surgical wound infection after the reengineering's 1993 implementation: increase of 134% in the surgical group. The authors concluded that the increase in skill mix was not large enough to overcome the decrease in full time equivalents and hours worked nor to compensate for the additional burden that a decreased length of stay poses on nursing staff.

Romano et al determined how accurately postoperative complications were reported in administrative data, whether accuracy varies systematically across hospitals, and whether serious complications were more consistently reported [70]. Therefore, 991 randomly sampled adults who underwent elective lumbar diskectomies at 30 non-federal acute care hospitals in California in 1990 to 1991 were selected. The sensitivity of reporting for this complication was < 35%, the specificity was 98%, the positive predictive value was 82% and the negative predictive value was 84%.

**Sources**

This indicator was originally proposed in 1992 by Iezzoni et al. as part of the CSP[65]. Needleman identified this indicator as a nursing-sensitive patient outcome, in which it was restricted to surgical patients only [60].

**Specification of numerator/denominator of PWI**

| | |
|---|---|
| Numerator | Discharges amongst cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM code of wound infection in any secondary diagnosis field.<br>See Appendix 2.H Wound infection |
| Denominator | All surgical discharges age 16 years and older defined by specific APR-DRG and an operating room procedure.<br>Surgical Discharge APR-DRGs: procedural APR-DRGs<br>See Appendix 2.A Operating Room Procedure<br><br>Exclude cases:<br>with principal diagnosis of wound infection<br>See Appendix 2.H Wound infection |

### *2.2.2.5 Ventilator-Acquired Pneumonia (VAP)*

**Summary of the evidence**

| Indicator | Prevalence (/1000 at risk) | Source | Coding validity | Construct validity | Clinical criteria |
|---|---|---|---|---|---|
| Ventilator-acquired pneumonia | Belgium[82] : 141.2 (surgical) | Needleman [60] | No evidence found in the literature | | Available |

**Definition of indicator Ventilator-Acquired Pneumonia**

*Numerator*. Discharges amongst cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM code pneumonia in any secondary diagnosis field.

*Denominator*. All surgical and medical discharges age 16 years and older defined that were ventilated. Exclude cases with pneumonia in principal diagnosis, any diagnosis for viral pneumonia, with diseases or disorders of respiratory system (MDC 4) or with immunocompromised state, all obstetrics admissions (MDC 14: pregnancy, childbirth, and puerperium)

**International prevalence figure in literature**

In Belgium, the prevalence of ventilator-acquired pneumonia (VAP) in surgical patients only was 14.12% for year 2003[82]. The OECD committee estimated that incidence of VAP ranged from 6% to 52% of intubated patients depending on patient risk factors [55]. After publishing guidelines for the prevention of VAP an analysis of studies published since 2004 was conducted by Gastmeier et al[83]. Five cohort studies using multi-module programmes to improve VAP were discovered with a range of improvement from 31 to 57%. However, these success stories were often reported from hospital wards with particularly high baseline rates and methodological weaknesses making it difficult to draw any conclusions. Furthermore, different definitions for VAP have been used in the studies investigated. On the other hand, a review of randomized controlled trials, meta-analyses or systematic reviews showed that many VAP cases are preventable and there is still room for improvement[83]. Regardless of the published guidelines for the prevention of VAP it can be assumed that the prevalence of VAP is still above the range of 5 cases per 1000 population at risk.

**Summary of the indicator**

This indicator is intended to capture cases of ventilator associated pneumonia and is limited to secondary diagnosis codes of pneumonia to eliminate complications present on admission. Also patients with principal diagnosis of pneumonia, an immunocompromised state or patients with MDC 4 were excluded as these patients were likely to suffered from aspiration pneumonia already on admission.

### Literature review/evidence levels

*Importance of the indicator*. Panel of experts from OECD estimated that VAP was a leading cause of morbidity and mortality in the ICU [55]. Overall VAP is associated with a mortality of up to 30%.

*Coding validity*. No evidence found in the literature.

*Construct validity*. Patients with mechanically-assisted ventilation have a high risk of developing nosocomial pneumonia. Prevention and control of nosocomial pneumonia is discussed in the CDC/HICPAC document, Guidelines for Prevention of Nosocomial Pneumonia. The Guideline strongly recommends that surveillance be conducted for bacterial pneumonia in ICU patients who are mechanically ventilated to facilitate the identification of trends and comparative analysis. High rates may suggest the need to examine the clinical and organizational processes related to the care of patients on ventilators including adherence to recommended guidelines. Ventilator associated pneumonia was suggested as an indicator potentially sensitive to nursing care by: National Quality Forum (2005) [84]. Collard and Saint[85] reviewed four evidence based practices that carried the potential to reduce the occurrence of VAP in patients receiving mechanical ventilation, including randomised clinical trials. The Committee of OECD [55] also emphasized that literature identified only a small number of explicit processes of care that have been proven in randomised clinical trials to prevent this complication. Given the grave consequences of VAP and the efforts that ICUs undertake to prevent them, VAP rates appear to be a plausible indicator of patient safety.

### Sources

This indicator was originally proposed by NQF and complemented with information from the pneumonia indicator of PSI.

**Specification of numerator/denominator of VAP**

| | |
|---|---|
| Numerator | Discharges amongst cases meeting the inclusion and exclusion rules for the denominator with ICD-9-CM code pneumonia in any secondary diagnosis field.<br>See Appendix 2.I Pneumonia |
| Denominator | All surgical and medical discharges age 16 years and older defined that were ventilated. Ventilated patients are selected by a RIZIV-code from Belgian nomenclature for procedure codes: 211046:<br>Artificial ventilation from day 2 until day 21.<br>Surgical Discharge APR-DRGs: procedural APR-DRGs<br>Medical Discharge APR-DRGs: medical APR-DRGs<br><br>Exclude cases:<br>with principal diagnosis of pneumonia or 997.3<br>See Appendix 2.I Pneumonia<br>with any diagnosis code for viral pneumonia<br>See Appendix 2.J Viral Pneumonia<br>MDC 4 (diseases/disorders of respiratory system)<br>MDC 14 (pregnancy, childbirth, and puerperium)<br>with any diagnosis of immunocompromised state.<br>See Appendix 2.K Immunocompromised states |

### 2.2.2.6 *Literature review for the non-selected indicators*

Accidental puncture or laceration, complication of anesthesia, foreign body left in during a procedure, iatrogenic pneumothorax, infection due to medical care, postoperative abdominopelvic wound dehiscence, postoperative hemorrhage or hematoma, postoperative hip fracture, postoperative physiologic and metabolic derangements, postoperative respiratory failure and transfusion reaction were events too rare to be selected. We did not find a clearly defined prevalence for aspiration pneumonia, postoperative infection (except wound and pneumonia) and reopening of surgical site.

Clinical criteria may be ambiguous for accidental puncture or laceration, aspiration pneumonia, complication of anesthesia, gastrointestinal hemorrhage, infection due to medical care, postoperative physiologic and metabolic derangements, postoperative respiratory failure, shock or cardiac arrest and urinary tract infection.

**Table 7 Literature review for the indicators not selected in this project**

| Indicator | Prevalence (/1000 at risk) | Source | Coding validity | Construct validity | Clinical criteria |
|---|---|---|---|---|---|
| Accidental puncture or laceration | Belgium [54]: 2.65<br><br>ARHQ[56] : 3.55<br>Zhan [11]: 3.32<br>Romano [57]: 3.24<br>Coffey [58]: 3.27<br>Rosen [59]: 2.82<br><br>Belgium: 13.5 (Surgical) - 14.1 (Medical) *[12]*<br><br>Needleman [60]: 12 (Surgical) - 23 (Medical) | PSI<br><br>Needleman [60] | *Procedure-related perforation or laceration (CSP)* [74-76]<br><u>Surgical cases</u><br>PPV : 81.6%<br>NPV : 99.1%<br>Present on admission : 24%<br>Complication confirmed by reabstraction : 94%<br>Confirmed cases (timing&complication) by reabstraction : 71%<br>Complication confirmed by physicians : 58.3%<br>Complication confirmed by physicians with at least one potential quality problem: 61.9%<br><br>Sensitivity <35%, specificity 98%, PPV 82%, NPV 84% [70] | *Procedure-related perforation or laceration (CSP)* [74]<br><u>Surgical cases</u> (n = 30 ) (procedure related) :<br>clinical evidence : 83.3%<br>physician notes : 6.7%<br>no evidence : 10%<br><br>Weak association between PSI [59] | Clinical criteria may be ambiguous |
| Aspiration Pneumonia | no prevalence | CSP | *[74 75 76, 77]*<br><u>Surgical cases</u> (n = 35 ) :<br>PPV : 85.7%<br>NPV : 97.4%<br>Present on admission : 15%<br>Complication confirmed by reabstraction : 94%<br>Confirmed cases (timing&complication) by reabstraction : 77%<br>Complication confirmed by physicians : 58.8%<br>Complication confirmed by physicians with at least one potential quality problem : 30%<br>At least one process problem : 68.8% | [74]<br><u>Surgical cases</u> (n = 32 ) :<br>clinical evidence : 53%<br>physician notes : 37.5%<br>no evidence : 9.4% | Clinical criteria may be ambiguous |

| Indicator | Prevalence (/1000 at risk) | Source | Coding validity | Construct validity | Clinical criteria |
|---|---|---|---|---|---|
| Complication of anesthesia | Belgium [54]: 0.58<br><br>ARHQ[56] : 0.81<br>Zhan [11]: 0.71<br>Romano [57]: 0.56<br>Coffey [58]: 0.69<br>Rosen [59]: 0.59 | PSI | no evidence | Weak association between PSI [59]<br><br>Decreased 18% between 1995 and 2000 [57] | Clinical criteria may be ambiguous |
| Failure to rescue (FTR) | Belgium [54]: 176.24<br><br>ARHQ[56] : 127.69<br>Zhan [11]: 169.13<br>Romano [57]: 174.24<br>Rosen [59]: 156.16<br><br>Belgium: 211 (Surgical) - 240 (Medical) [12]<br><br>Needleman [60]: 197 (Surgical) - 186 (medical)<br><br>Smith [61]: 155 | PSI<br><br>Needleman [60]<br><br>other | no evidence | Association between higher proportion of registered-nurse-hours and lower rate of FTR (medical cases) [60]<br><br>Association between patient-to-nurse ratio and FTR; each additional patient per nurse associated with a 7% increase in the odds of FTR [86]<br><br>Weak association between PSI [59]<br><br>No statistically significant difference in complaint categories between patients who experienced an adverse event and those who did not [68] | Available |
| Foreign body left in during procedure | Belgium [54]: 0.07<br><br>ARHQ[56] : 0.08<br>Zhan [11]: 0.09<br>Romano [57]: 0.08<br>Coffey [58]: 0.09<br>Rosen [59]: 0.17 | PSI | no evidence | no evidence<br><br>Weak association between PSI [59] | Available |

| Indicator | Prevalence (/1000 at risk) | Source | Coding validity | Construct validity | Clinical criteria |
|---|---|---|---|---|---|
| Gastrointestinal hemorrhage | Belgium: 8.2 (Medical) - 3.6 (Surgical) [12]<br><br>Needleman [60]: 10 (Medical) - 5 (Surgical) | Needleman [60] | *Postoperative gastrointestinal hemorrhage (CSP)* [75 76]<br>Surgical cases (n = 41) :<br>PPV : 81.4%<br>NPV : 99.1%<br>Present on admission : 15%<br>Complication confirmed by reabstraction : 81 %<br>Confirmed cases (timing&complication) by reabstraction : 66 %<br>Complication confirmed by physicians : 72.5%<br>Complication confirmed by physicians with at least one potential quality problem : 48.3% | *Postoperative gastrointestinal hemorrhage (CSP)* [74]<br>Surgical cases (n = 39 )<br>clinical evidence : 69.2%<br>physician notes : 7.7%<br>no evidence : 23.1%<br><br>Higher proportion of licensed-nurse[b] care provided by registered nurses and more registered-nurse-hours per day were associated with lower rate of Upper gastrointestinal bleeding (medical cases) [60] | Clinical criteria may be ambiguous |
| Hospital-acquired Pneumonia | Belgium: 13.5 (Surgical) - 14.1 (Medical) [12]<br><br>Needleman [60]:<br>12 (Surgical) - 23 (Medical)<br><br>Kovner [78]: 7.5 in 1990; 12.4 in 1996 | Needleman [60]<br><br>Other | *Postoperative pneumonia (CSP)* [74 75 77]<br>Surgical cases (n= 42)<br>PPV : none<br>NPV : none<br>Present on admission : none<br>Complication confirmed by reabstraction : none<br>Confirmed cases (timing&complication) by reabstraction : none<br>Complication confirmed by physicians : 64.3%<br>Complication confirmed by physicians with at least one potential quality problem : 7.4%<br>At least one process problem : 82.5% | *Postoperative pneumonia (CSP)* [74]<br>Surgical cases (n = 40 ) :<br>clinical evidence : 50%<br>physician notes : 30%<br>no evidence : 20%<br><br>Higher proportion of licencied-nurse care provided by registered nurses and more registered-nurse-hours per day associate with lower rate (medical cases) [60]<br><br>Inverse relation between registered-nurse staffing and rate (surgery) [78] | Available |

[b] Licensed nurses are registered nurses and licensed practical nurses [60]

| Indicator | Prevalence (/1000 at risk) | Source | Coding validity | Construct validity | Clinical criteria |
|---|---|---|---|---|---|
| Iatrogenic pneumothorax | Belgium [54]: 0.35<br><br>ARHQ[56] : 0.56<br>Zhan [11]: 0.67<br>Romano [57]: 0.67<br>Coffey [58]: 0.72<br>Rosen [59]: 1.2 | PSI | no evidence | no evidence<br><br>Weak association between PSI [59] | Available |
| Infection due to medical care | Belgium [54]: 1.64<br><br>ARHQ[56] : 2.14<br>Zhan [11]: 1.99<br>Romano [57]: 1.93<br>Coffey [58]: 2.13<br>Rosen [59]: 2.37<br><br>Smith [61]: 2.84 | PSI | no evidence | no evidence<br><br>Weak association between PSI [59] | Clinical criteria may be ambiguous |
| Postoperative abdominopelvic wound dehiscence | Belgium [54]: 1.20<br><br>ARHQ[56] : 2.00<br>Zhan [11]: 2.05<br>Romano [57]: 1.93<br>Coffey [58]: 2.13<br>Rosen [59]: 4.49<br><br>Smith [61]: 3.76 | PSI | no evidence | no evidence<br><br>Weak association between PSI [59] | Available |

| Indicator | Prevalence (/1000 at risk) | Source | Coding validity | Construct validity | Clinical criteria |
|---|---|---|---|---|---|
| Postoperative hemorrhage or hematoma | Belgium [54]: 2.07<br><br>ARHQ[56] : 2.12<br>Zhan [11]: 2.06<br>Romano [57]: 2.06<br>Coffey [58]: 2.27<br>Rosen [59]: 2.90 | PSI | *Postprocedural hemorrhage or hematoma (CSP)* [74 75 76]<br><br><u>Surgical cases</u> (n = 46 ) :<br>PPV : 89.7%<br>NPV : 93.6%<br>Present on admission : 2%<br>Complication confirmed by reabstraction : 91 %<br>Confirmed cases (timing&complication) by reabstraction : 83 %<br>Complication confirmed by physicians : 56.5 %<br>Complication confirmed by physicians with at least one potential quality problem : 61.9%<br><br><u>Medical cases</u> (n = 53 ) :<br>PPV : 90.6%<br>NPV : 98.6%<br>Present on admission : 31%<br>Complication confirmed by reabstraction : 91%<br>Confirmed cases (timing&complication) by reabstraction : 49%<br>Complication confirmed by physicians : 54.9%<br>Complication confirmed by physicians with at least one potential quality problem : 46.4%<br><br>Sensitivity <35%, specificity 98%, PPV 82%, NPV 84% [70] | *Postprocedural hemorrhage or hematoma (CSP)* [74]<br><br><u>Surgical cases</u> (n = 44 ) :<br>clinical evidence : 54.5%<br>physician notes : 25.0%<br>no evidence : 20.5%<br><br><u>Medical cases</u> (n = 45 ) :<br>clinical evidence : 33.3%<br>physician notes : 40.0%<br>no evidence : 26.7%<br><br>Weak association between PSI [59]<br><br>No statistically significant difference in complaint categories between patients who experienced an adverse event and those who did not [68] | Available |

| Indicator | Prevalence (/1000 at risk) | Source | Coding validity | Construct validity | Clinical criteria |
|---|---|---|---|---|---|
| Postoperative hip fracture<br>In-hospital hip fracture | *Postoperative hip fracture*<br>Belgium [54]: 0.31<br><br>ARHQ[56] : 0.28<br>Zhan [11]: 0.77<br>Romano [57]: 0.80<br>Coffey [58]: 0.76<br>Rosen [59]: 1.14 | PSI | *In-hospital hip fractures and falls (CSP)* [75 76 77]<br>Surgical cases (n = 21)<br>PPV : 85.0%<br>NPV : 99.2%<br>Present on admission : 21%<br>Complication confirmed by reabstraction : 91%<br>Confirmed cases (timing&complication) by reabstraction : 57%<br>Complication confirmed by physicians : 71.4%<br>Complication confirmed by physicians with at least one potential quality problem : 26.7%<br>At least one process problem : 76.2%<br><br>Medical cases (n = 64)<br>PPV : 60.6%<br>NPV : 99.5%<br>present on admission : 87%<br>Complication confirmed by reabstraction : 97%<br>Confirmed cases (timing&complication) by reabstraction : 11%<br>Complication confirmed by physicians : 10.9%<br>Complication confirmed by physicians with at least one potential quality problem : 28.6%<br>At least one process problem : 53.8% | Weak association between PSI [59] | Available |
| Postoperative infection (except wound and pneumonia) | no | CSP | [74 75 76]<br>Surgical cases (n = 30 ) :<br>PPV : 96.8%<br>NPV : 98.3%<br>Present on admission : 23%<br>Complication confirmed by reabstraction : 94 %<br>Confirmed cases (timing&complication) by reabstraction : 72 %<br>Complication confirmed by physicians : 73.3 %<br>Complication confirmed by physicians with at least one potential quality problem : 50% | [74]<br>Surgical cases (n = 27 ) :<br>clinical evidence : 81.5%<br>physician notes : 3.7%<br>no evidence : 14.8%<br><br>No statistically significant difference in complaint categories between patients who experienced an adverse event and those who did not [68] | Available |

| Indicator | Prevalence (/1000 at risk) | Source | Coding validity | Construct validity | Clinical criteria |
|---|---|---|---|---|---|
| Postoperative physiologic and metabolic derangements | Belgium [54]: 1.9<br><br>ARHQ[56] : 1.04<br>Zhan [11]: 1.00<br>Romano [57]: 0.89<br>Coffey [58]: 1.43<br>Rosen [59]: 1.81 | PSI | Sensitivity <35%, specificity 98%, PPV 82%, NPV 84% [70] | Weak association between PSI [59]<br><br>No association between nurses staffing and rate [60] | Clinical criteria may be ambiguous |
| Postoperative respiratory failure | *Postoperative respiratory failure*<br>Belgium [54]: 3.64<br><br>ARHQ[56] : 9.29<br>Zhan [11]: 3.57<br>Romano [57]: 3.58<br>Coffey [58]: 4.01<br>Rosen [59]: 3.43<br><br>Belgium: 14 [12]<br><br>Needleman [60]: 12 | PSI<br><br>Needleman [60] | *Postoperative pulmonary compromise CSP* [75] [76] [77]<br><u>Surgical cases</u> (n = 46 ) :<br>PPV : 92.5%<br>NPV : 96.2%<br>Present on admission : 12%<br>Complication confirmed by reabstraction : 91 %<br>Confirmed cases (timing&complication) by reabstraction : 72 %<br>Complication confirmed by physicians : 75%<br>Complication confirmed by physicians with at least one potential quality problem : 27.3%<br>At least one process problem : 52.3% | Weak association between PSI [59]<br><br>No association between nurse staffing and rate [60] | Clinical criteria may be ambiguous |
| Reopening of surgical site | no | CSP | [74] [75] [76]<br><u>Surgical cases</u> (n = 32) :<br>PPV : 88.2%<br>NPV : 98%<br>Present on admission : 0%<br>Complication confirmed by reabstraction : 97 %<br>Confirmed cases (timing&complication) by reabstraction : 97 %<br>Complication confirmed by physicians : 61.3 %<br>Complication confirmed by physicians with at least one potential quality problem : 42.1% | [74] [75] [76]<br><u>Surgical cases</u> (n = 44 ) :<br>clinical evidence : 86.7%<br>physician notes : 0%<br>no evidence : 13.3%<br><br>No statistically significant difference in complaint categories between patients who experienced an adverse event and those who did not [68] | Available |

| Indicator | Prevalence (/1000 at risk) | Source | Coding validity | Construct validity | Clinical criteria |
|---|---|---|---|---|---|
| Shock - Cardiac Arrest | Belgium: 5.32 (Surgical) - 6.67 (Medical) [12]<br><br>Needleman [60]: 5 (Surgical) - 6 (Medical) | Needleman [60] | *Postoperative shock or cardiorespiratory arrest (CSP)* [75 76]<br>Surgical cases (n = 40 ) :<br>PPV : 89.3%<br>NPV : 97.8%<br>Present on admission : 29%<br>Complication confirmed by reabstraction : 85 %<br>Confirmed cases (timing&complication) by reabstraction : 53 % Complication confirmed by physicians : 74.4 %<br>Complication confirmed by physicians with at least one potential quality problem: 20.7% | Higher proportion of registered-nurse-hours and lower rates (medical; no association in surgical group)no association between nursing and rate [60]<br><br>No statistically significant difference in complaint categories between patients who experienced an adverse event and those who did not [68] | Clinical criteria may be ambiguous |
| Transfusion reaction | Belgium [54]: 0.01<br>ARHQ[56] : 0.005<br>Zhan [11]: 0. 004<br>Romano [57]: 0.004<br>Coffey [58]: 0. 005<br>Rosen [59]: 0.007 | PSI | No evidence | Weak association between PSI [59] | Available |
| Urinary tract infection (UTI) | Belgium: 17.6 (Surgical) - 32.3 (Medical) [12]<br><br>Needleman [60]: 33 (Surgical) - 63 (Medical)<br><br>Quan [87]: 43 | Needleman [60]<br><br>Other | No evidence on validity is available from CSP studies<br><br>Sensitivity 55.6%; specificity 99.8%; PPV 62.5%; NPV 99.7% [87] | Higher proportion of licensed-nurse care provided by registered nurses and more registered-nurse-hours per day were associated with lower rate of UTI (medical cases) [60]<br><br>Association between higher proportion of registered-nurse-hours and lower rate of UTI (surgical cases) [60]<br><br>Association not statistically significant [62]<br><br>Association between skill mix and UTI [63] | Clinical criteria may be ambiguous |

# 3 METHODOLOGY

In this chapter, we first present the algorithms used and the adaptations made to the Belgian context. We then describe the data used in the algorithms, the selection of participating hospitals, the selection of medical files reviewed and the screening of the files. A brief explanation of analyses is developed followed by a description of the approval of the ethics committee.

## 3.1 ALGORITHM SOURCES AND ADAPTATIONS

The selected indicators are were all (integral or partial ) Patient Safety Indicators of the Agency for Healthcare Research and Quality (AHRQ)[56], except postoperative wound infection which is was based on the OECD-algorithm [55]. The algorithms used in this study are were based on the technical manual provided by the AHRQ (version 3.1) or the OECD patient safety indicators. For indicators of the AHRQ[56], algorithms were the same as those used in the recent feedback[54] sent to all Belgian hospitals by the Ministry of Public Health.

In Belgium the APR-DRG v15.0 is used in the hospital financing system. The internationally published algorithms use HCFA DRGs. This implies that the most important adjustment that needs to be made is the mapping of HCFA DRG to APR-DRG v15.0 . However, currently there is no such crosswalk between the APR-DRG and HCFA DRG for use with the patient safety indicators. There is not really a one-to-one mapping, although many of the DRG are very similar. Van den Heede et al. (in progress) did a cross-mapping for the DRGs used by the AHRQ in the Cancer-DRG algorithm and the infection-DRG algorithm to the APR-DRG v15.0 system (see appendix 2.M and appendix 2.N). Sermeus et al.[82] developed a crosswalk between the APR-DRG and HCFA DRG for use with the patient safety indicators. This cross-mapping was applied, in the present study, in the Cancer-DRG algorithm and the infection-DRG algorithm (see appendix 2.M and appendix 2.N).

A second important modification concerned the procedures. In the Belgian administrative data the principal procedure is not identified, like it is done in US hospital discharge data. In the algorithms used in the present study we replaced 'principal procedure' by identifying the first operating room procedure in chronological order (Postoperative DVT/PE, Decubitus Ulcer).

## 3.2 SAMPLE DATA

### 3.2.1 Data source

The federal government service of Public Health, Food chain safety and Environment collects data concerning the Belgian health care organizations. The registration of these data is important for the guidance to determine the health policy and financing of the Belgian health care organizations. One of the registration systems in acute hospitals is the Belgian Hospital Discharge Dataset (B-HDDS).

The Belgian Hospital Discharge Dataset includes an administrative clinical database gathering information transmitted by each hospital to the Ministry of Public Health, called MKG/RCM or Minimale Klinische Gegevens/Résumé Clinique Minimum. All non-psychiatric hospitals must participate to this data collection since 1990. The available information concerning outpatient or inpatient stay discharged are mainly year of birth, sex, domicile zip code, length of stay, year and month of admission and discharge, in addition to all diagnoses and procedures coded in ICD-9-CM (International Classification of Disease, 9th revision, Clinical modification, published in October 2005). The Ministry runs the APR-DRG version 15th grouper program to assign an APR-DRG (All-Patient Refined Diagnosis Related Group). The registration is continuous and every registration period lasts 6 months.

The purposes of RCM/MKG are:

- to determine the need for hospital facilities;
- to define the qualitative and quantitative recognition standards of hospitals and their services;
- to organize the financing of hospitals;
- to determine the policy concerning the practice of medicine;
- to outline a policy in relation with the epidemiology;
- to help the hospitals in their internal management (feedbacks on their data).

Because of the frequency of registration, data from RCM/MKG are available with one year delay, after a numerical validation process.

Although the database contains complete information from every Belgian hospital, the reliability of the information is a major concern as there is a lack of external validation of the registered data[88]. However the Ministry of Public Health has audited the data and concluded, for the period 2004-2007, that even in case of upcoding of the pathologies, the phenomenon tends to decrease compared to the last analysis[89].

The selection of patient's files reviewed during this project was based on the Belgian Hospital Discharge Dataset (B-HDDS). The B-HDDS were provided by the selected hospitals in an anonymous form.

The quality of the data is audited by the Ministry of Public Health in two ways. Firstly, a software program checks the data for missing, illogical, and outlier values. Secondly, by regular hospital visits, a random selection of patient records is reviewed to ensure that data were recoded correctly.

### 3.2.2 Hospital selection

All 116 Belgian acute hospitals were invited to join this study. Thirteen Flemish acute hospitals and eight Walloon acute hospitals agreed to participate. A selection of eight hospitals was done according to the hospital size, the geographical spread and the moment they expressed interest in participating the study.

Four hospitals each were selected for the Flemish region and the French-speaking region. Amongst the eight selected hospitals, there was one university teaching hospital for each region. One hospital had less than 300 beds, 2 hospitals had between 300 and 450 beds and 5 hospitals had more than 450 beds. All the hospitals were located in different provinces.

### 3.2.3 Study population

The study targeted adult non-obstetric patients admitted in an acute care hospital in Belgium. Long-term care and rehabilitation facilities as well as psychiatric hospitals and one-day clinics were excluded. Patients who were transferred from another acute care facility were also excluded.

Basically, administrative data on discharges for the registration year 2005 were used for the selection of cases. Whenever necessary, additional registration years 2004 and 2006 were used in order to become a sufficient number of cases.

### 3.2.4 Development and description of the file selection

This project was focused on five events being likely to occur during an in-hospital stay. Adverse events are known to have a small prevalence, therefore a random selection in the administrative datasets wouldn't reach a sufficient number of cases to validate the chosen indicators unless a very large sample was taken[47]. It was not possible to review all the medical files of all stays in the eight selected hospitals because of time and resource constraints.

We aimed at validating five adverse event indicators chosen on the basis of the evidence available in the literature and on their importance. Therefore, we focused on a selection of sufficient cases for each indicator. To ensure a more objective review of medical files and of the information available, a selection of controls was required. For the cases population, medical records were selected randomly. Moreover, a stratified method for the selection of control files was used to enhance the blinded review of medical records. For each 'flagged' case - that is a hospital discharge case which was positive for at least one of the selected indicators - a matched control case was selected. The matching was based on the APR-DRG, the severity of illness (SOI), age, gender, year and semester of registration. The severity of illness was classified in four categories (SOI 1, SOI 2 and 3, SOI 4). Age was also divided in five categories (<30 years old, [30-49], [50-64], [65-79] and age >=80 years). As adverse events are known to be associated with co-morbidity and the age of the patient, a simple method to standardize the populations with the available data was to select patients on the basis of patient age and of APR-DRG/SOI. The populations with or without flagged adverse event were therefore more homogeneous.

For the registration year 2005 of the B-HDDS, we have randomly selected 20 flagged cases and 20 control cases per adverse event for each hospital in order to obtain a total of 200 patients per hospital. In case of an insufficient number of patients, another 50 reserve patients were randomly drawn from the pool of not selected cases. However, those could not be balanced for the 5 indicators due to the unequal prevalence of adverse events. In order to reach the targeted number of cases, additional registration years 2004 and 2006 were requested whenever necessary. The reserve of 50 patients could not be obtained in one Flemish hospital due to its small size.

### 3.2.5 Medical record review

The medical records in each hospital selected were evaluated independently by 2 teams during a one stage review process. The choice of two teams instead of one was mainly made in order to overcome the language barrier. Even in the case of being bilingual, some detailed – yet important – information could get lost by reviewing the medical records, serving as a potential bias in the search and evaluation of adverse events. In the Flemish region, the team was composed of one internal medicine specialist and one clinical pharmacist. The latter was experienced after similar medical record reviewing on adverse drug events. In the French-speaking region, it was one surgeon and one nurse. The two members of both teams had a broad clinical experience which was essential for the objective of this study. The members of the two teams were trained to use the abstraction tool together after a pre-test was performed on 20 medical records to get used to the tool.

In each team, both reviewers were blinded for the result of the screening procedure based on administrative data. In other words, they were unaware whether a case was positively flagged for one of the 5 indicators or not. Given the stratified method for medical file selection – already leading to a group of 'flagged' cases and control cases – only a one stage medical record review was performed. Whenever the reviewers disagreed on the occurrence of an adverse event, they debated on the case until a consensus was reached. If no consensus could be obtained, the medical record was presented to a panel of experts. The same experts were also involved in determining the specific clinical criteria of the five selected adverse events. In this regard, the medical record review process was partly an implicit procedure (expert opinion) and partly an explicit procedure (adherence to specific clinical criteria).

### 3.2.6 Data abstraction

Medical records were screened using a data abstraction tool (see appendix 2.O). The tool aimed at providing a standardized method of recording and data collection. It consisted of an anonymous number, admission and discharge dates, and appreciation of the completeness of the medical file. The initial aim was to proceed with the medical record review on the condition of the combined presence of nursing notes, procedures notes and discharge notes.

Since especially nursing notes were incomplete or totally missing up to 54.24 % of the cases in the hospital 7, the medical file was considered as complete if the following 2 out of 3 items were present: nursing notes, procedures notes and discharge note.

The particularly high number of missing nurses' notes in hospital 7 was due to a reorganization of the archiving department which has been moved to another site. Table 8 provides the distribution of present parts in the medical file per hospital.

**Table 8 Distribution of present parts – nursing notes, procedures notes, and discharge notes – per hospital**

| | Complete | | Nurse notes missing | | Procedures missing | | Discharge missing | | Total file number |
|---|---|---|---|---|---|---|---|---|---|
| | N | % | N | % | N | % | N | % | |
| **Hosp 1** | 147 | 90.19 | 2 | 1.23 | 7 | 4.29 | 7 | 4.29 | 163 |
| **Hosp 2** | 171 | 85.50 | 23 | 11.50 | 2 | 1.00 | 4 | 2.00 | 200 |
| **Hosp 3** | 142 | 71.00 | 55 | 27.50 | 3 | 1.50 | 0 | 0.00 | 200 |
| **Hosp 4** | 188 | 94.00 | 4 | 2.00 | 3 | 1.50 | 5 | 2.50 | 200 |
| **Hosp 5** | 153 | 80.10 | 12 | 6.28 | 17 | 8.90 | 9 | 4.71 | 191 |
| **Hosp 6** | 133 | 59.91 | 17 | 7.66 | 63 | 28.38 | 9 | 4.05 | 222 |
| **Hosp 7** | 48 | 31.37 | 83 | 54.24 | 19 | 12.42 | 3 | 1.96 | 153 |
| **Hosp 8** | 153 | 82.26 | 0 | 0.00 | 11 | 5.91 | 22 | 11.83 | 186 |

Some patient information was provided through administrative data and was verified by the review team: age and sex, admission type (elective or emergency), place before admission and destination after discharge, length of stay and co-morbidities. The latter were derived from the Elixhauser co-morbidity measurement method.

Furthermore, strict clinical criteria were defined for the 5 indicators selected, based both on the literature and expert opinion. Those criteria can be consulted in appendix 2.O in the abstraction tool.

Finally, if an adverse event was found in the medical record a detailed analysis was made about the occurrence of the event. Based on the framework provided in Figure 1 the reviewer assessed whether the adverse event was the result of health care management rather than the patient's disease itself, the degree of disability and extra length of stay due to the adverse event, the potential quality problems which caused the adverse event, the degree of responsibility of health care management (scale from 1 to 6) and the preventability of the event (scale from 1 to 6), the specialty caring for the patient and the place of occurrence of the adverse event and, finally the additional care which was necessary because of the adverse event. The extra length of stay due to the occurrence of an adverse event was estimated based on the national average length of stay in Belgium for a given disease and can therefore only be seen as a subjective figure. If an indicator was judged to be present on admission, no further detailed analysis on the indicator was made.

### 3.2.7 Statistical analysis

All analyses were performed using SAS v9.1 (SAS Institute, Inc, Cary, NC).

The first part of the results was the comparison between the studied population and the non-consenting population in order to determine whether the patient approval has introduced a bias in the study. In this way, we compared the mean age (Student t), the distribution of sex (Chi square), and the distribution of non flagged/ flagged cases (Chi square) in the two populations.

Thereafter, descriptive analyses were calculated for the 5 indicators. A correction was made whenever an exclusion rule applied for a certain indicator in the B-HDDS. In other words, whenever an exclusion rule was present in the B-HDDS, the found indicator during the medical record screening was neglected. Adverse events rates were reported as the percentage of hospitalizations during which adverse events were detected.

Comparing the B-HDDS and the medical record screening positive predictive values (PPV) were calculated for the group of 5 indicators in general and per indicator separately. PPV is calculated as a simple ratio of the cases which are flagged as positive by the test and are identified ad diseased by the gold standard (Medical record review) divided by all cases flagged as positive. That is:

$$PPV = \frac{\text{Cases positive on the indicator and in the medical record (true positve)}}{\text{Cases positive on the indicator regardless of the medical record (true + false positive)}}$$

Subsequently, cases which had an indicator present on admission during medical record screening were excluded and then revealed a new set of PPV. The B-HDDS was considered to be the test-value while the medical record screening was considered to be the true value.

The second part of analyses concerned the characteristic of events found in the medical files. If an adverse event was present in the medical file the reviewer had to assess whether it depended on the medical management and whether it was preventable.

This assessment was reflected in the following two questions of the abstraction tool:

- Consider the extent to which health care management rather than the disease process is responsible for the AE[14, 41, 42]:
  1. Virtually no evidence for management causation/system failure
  2. Slight-to-modest evidence for management causation
  3. Management causation not likely; less than 50-50 but close call
  4. Management causation more likely than not, more than 50-50 but close call
  5. Moderate/strong evidence for management causation
  6. Virtually certain evidence for management causation
- Degree of preventability of the adverse event[14]:
  1. *No preventability* : virtually no evidence for preventability
  2. *Low preventability* : slight-to-modest evidence for preventability
  3. *Low preventability* : preventability not likely, less than 50-50 but close call
  4. *High preventability* : preventability more likely than not, more than 50-50 but close call
  5. *High preventability* : strong evidence for preventability
  6. *High preventability* : virtually certain evidence for preventability

If the score to one of these was higher than 3, the event was considered as being an evident adverse event[14].

### 3.2.8 Ethical approval

The study was approved by the ethics committee of all hospitals. The research team obtained informed consent from the selected patients. A letter was send by post by the participating hospitals to all selected patients (see appendix 2.P). This letter contained information about the study and a reply-card. Patients had the opportunity to answer during a period of 30 days to give or refuse consent. If no answer was received after 30 days, informed consent was assumed. The management of each participating hospital received a letter with information concerning the study, which could possibly be spread to their hospital physicians.

# 4 RESULTS

In this chapter, we first compare the non-consenting population with the study population in order to determine whether the patient approval has introduced a bias in the study or not. After that, we present results of positive predictive values globally and for each indicator. We also describe the reason why cases don't correspond with the administrative data. Finally, we observe for all adverse events found in the medical review how cases were preventable or due to health care management.

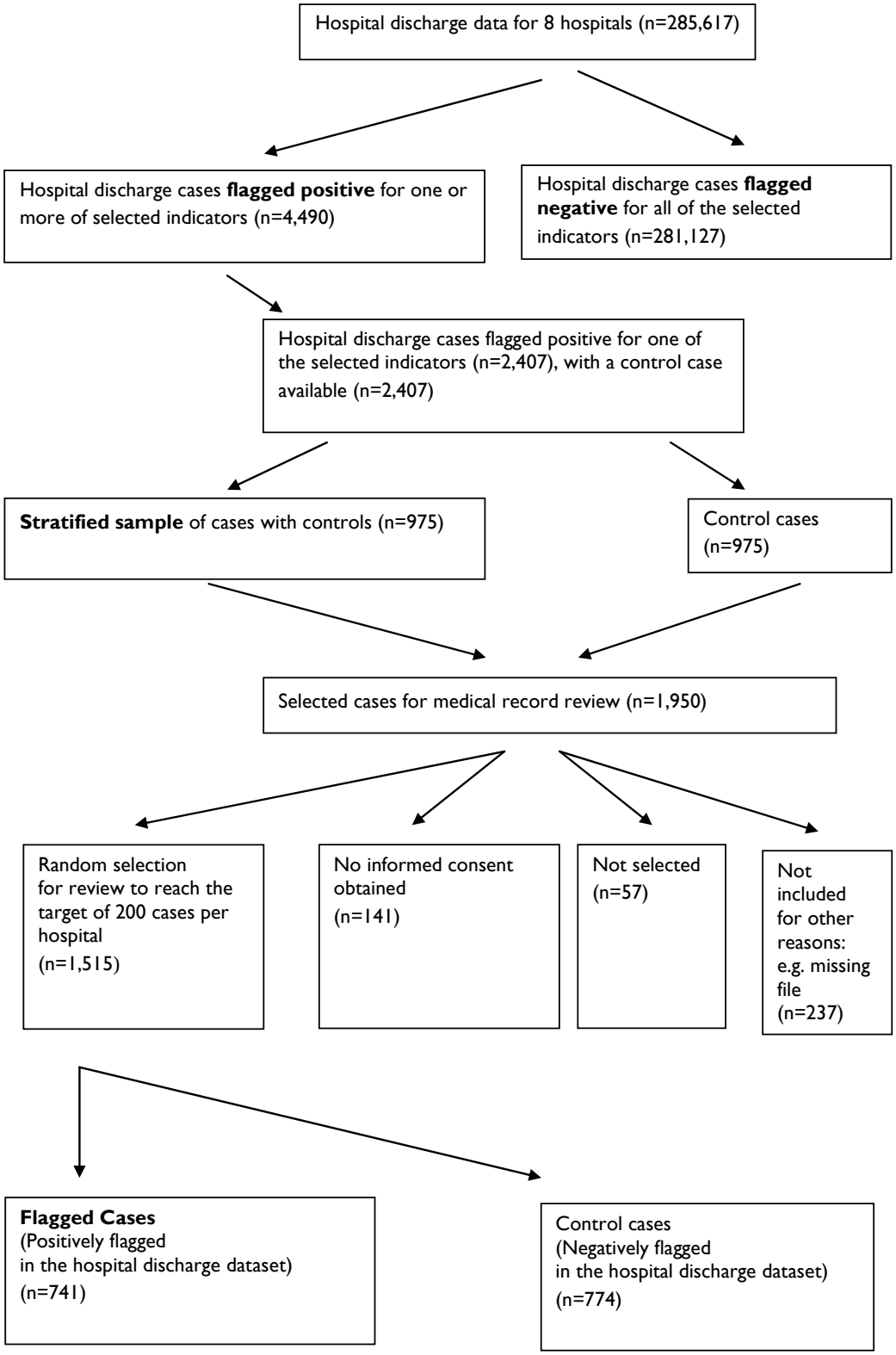
## 4.1 SELECTION PROCEDURE OF CASES

Administrative data obtained from the eight hospitals selected revealed a total of 285,617 hospital stays (Figure 3). Hospital discharge data were flagged positive for one or more of the selected indicators in 4,490 cases and flagged negative for all of the selected indicators in 281,127 cases. For 2,407 positive flagged cases, a control case was available. After a stratified sample of cases according to the predefined indicators, a total of 975 flagged hospital stays and 975 control cases were withheld. Consequently, 1,950 hospital stays were potentially meant for medical record review.

No informed consent was obtained in 141 cases. A total of 237 files could not be included due to other reasons: 173 files were not available or could not be found at the time of the study, 47 files were excluded because of incompleteness, 13 files weren't available due to an address change of the patient and 4 cases due to the transfer of patients from another acute hospital.

As a result, a total of 1,515 medical files were randomly selected for medical record review in order to reach the target of 200 cases per hospital. Of these, 741 files were positively flagged in the hospital discharge dataset for at least one of the selected indicators and 774 cases were controls and hence negatively flagged.

**Figure 3 Selection of cases**



## 4.2 COMPARISON OF THE STUDIED AND NON-CONSENTING POPULATION

Amongst the 1950 selected cases for medical record review, 141 patients refused to participate to the study.

As shown in table 9, the population studied was statistically younger than the population who refused to participate to this project. The mean age was 68.2+/-15.9 years (range 16 to 102) in the studied population and 72.4+/-11.3 years (range 33 to 94) in the non-consenting population (p<0.0001).

**Table 9 : Mean age in studied and non-consenting population**

| | N | Lower CL Mean | Mean | Upper CL Mean | Lower CL Std Dev | Std Dev | Upper CL Std Dev | Min | Max |
|---|---|---|---|---|---|---|---|---|---|
| Studied population | 1515 | 67.38 | 68.18 | 68.979 | 15.312 | 15.857 | 16.442 | 16 | 102 |
| Non-consenting population | 141 | 70.509 | 72.39 | 74.271 | 10.115 | 11.297 | 12.795 | 33 | 94 |

Likewise (Table 10) the repartition of sex was different in both populations. Women were statically more represented in the non-consenting population (57.45 %) than in the studied population (48.71 %) (p<0.05).

**Table 10: Distribution of sex in studied and non-consenting population (1=male; 2=female)**

| | SEX | | |
|---|---|---|---|
| Frequency<br>Row Pct | 1 | 2 | Total |
| Studied population | 777<br>51.29 | 738<br>48.71 | 1515 |
| Non-consenting population | 60<br>42.55 | 81<br>57.45 | 141 |
| Total | 837<br>50.54 | 819<br>49.46 | 1656<br>100.00 |

Finally, we compared the proportion of flagged cases and control cases between the two populations (Table 11). Although the proportion of cases flagged was larger in the non-consenting population, the difference was not statistically significant (p>0.05).

**Table 11 : Distribution of the cases flagged or not in studied and in non-consenting population**

| | Flagged in B-HDDS | | |
|---|---|---|---|
| Frequency<br>Row Pct | No | Yes | Total |
| Studied population | 774<br>51.09 | 741<br>48.91 | 1515 |
| Non-consenting population | 64<br>45.39 | 77<br>54.61 | 141 |
| Total | 838 | 818 | 1656 |

## 4.3 RESULTS OF PREDICTIVE VALUES

We reviewed 1,515 (77.7%) out of the 1,950 sampled cases for medical record review. The target of 200 cases was reached in half of the selected hospitals. Hospital 1 lacked 37 files; hospitals 5, 7 and 8 lacked respectively 9, 47 and 14 cases mainly because files were not available at the time of the review. In order to partially correct this shortage of files an extra 22 files were reviewed in hospital 6.

Hence, 1,515 (94.7%) out of the 1,600 targeted records were reviewed. Out of the 1,515 records, 280 (18.5%) were registered in 2004, 1198 (79.1%) in 2005 and the remaining 37 (2.4%) files were registered in 2006.

The mean age of the reviewed patients was 68.2+/- 15.9 years with a median age of 72 years. Mean age by hospital varied from 62.0 years (hospital 3) to 75.1 years (hospital 1) (Table 12). The mean length of hospital stay was 27.3+/-31.2 days with a median length of stay of 16 days. The mean length of stay per hospital varied from 18.4 days (hospital 1) to 38.6 days (hospital 6) (Table 13).

**Table 12 : Age distribution per hospital**

| | Age | | | | | | |
|---|---|---|---|---|---|---|---|
| | Mean | Std | Min | P25 | Median | P75 | Max |
| **hosp** | | | | | | | |
| 1 | 75.13 | 12.65 | 33.00 | 70.00 | 77.00 | 84.00 | 97.00 |
| 2 | 65.50 | 14.32 | 16.00 | 59.00 | 68.50 | 76.00 | 99.00 |
| 3 | 62.03 | 16.44 | 16.00 | 52.50 | 65.50 | 75.00 | 92.00 |
| 4 | 66.00 | 17.36 | 29.00 | 52.00 | 69.50 | 82.00 | 99.00 |
| 5 | 66.55 | 17.36 | 16.00 | 55.00 | 71.00 | 81.00 | 93.00 |
| 6 | 70.13 | 16.00 | 21.00 | 64.00 | 73.00 | 81.00 | 98.00 |
| 7 | 72.40 | 11.61 | 40.00 | 65.00 | 74.00 | 80.00 | 96.00 |
| 8 | 69.81 | 15.32 | 23.00 | 59.00 | 75.00 | 82.00 | 102.00 |
| Total | 68.18 | 15.86 | 16.00 | 59.00 | 72.00 | 80.00 | 102.00 |

**Table 13 : Length of stay distribution per hospital**

| | Length of stay | | | | | | |
|---|---|---|---|---|---|---|---|
| | Mean | Std | Min | P25 | Median | P75 | Max |
| **hosp** | | | | | | | |
| 1 | 18.39 | 14.55 | 1.00 | 9.00 | 15.00 | 25.00 | 103.00 |
| 2 | 27.33 | 35.32 | 1.00 | 8.50 | 13.00 | 30.50 | 200.00 |
| 3 | 30.20 | 40.82 | 1.00 | 10.00 | 16.00 | 32.00 | 297.00 |
| 4 | 25.43 | 26.38 | 1.00 | 9.00 | 15.00 | 35.50 | 197.00 |
| 5 | 24.82 | 27.47 | 0.00 | 11.00 | 17.00 | 33.00 | 249.00 |
| 6 | 38.64 | 39.42 | 0.00 | 14.00 | 26.00 | 50.00 | 242.00 |
| 7 | 21.99 | 21.92 | 1.00 | 9.00 | 15.00 | 25.00 | 131.00 |
| 8 | 27.56 | 24.53 | 1.00 | 12.00 | 19.00 | 33.00 | 133.00 |
| Total | 27.33 | 31.19 | 0.00 | 10.00 | 16.00 | 33.00 | 297.00 |

Most cases (79%) had a surgical diagnostic category which could be explained by the specific choice of the selected indicators. As shown in Table 14, the main major diagnostic categories (MDC) were the circulatory system (30%) and the musculoskeletal system (19%). The digestive system was represented in 11% of the cases, respiratory system in 7% and the nervous system in 5%. The remaining 27% of cases came from 17 other categories.

**Table 14 : Five most frequent Major Diagnostic Categories (MDC) in studied population**

| MDC | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| 05 Circulatory System & related condition | 459 | 30.30 | 459 | 30.30 |
| 08 Musculoskeletal System and Connective Tissue & related condition | 290 | 19.14 | 749 | 49.44 |
| 06 Digestive System & related condition | 168 | 11.09 | 917 | 60.53 |
| 04 Respiratory System & related condition | 112 | 7.39 | 1029 | 67.92 |
| 01 Nervous System & related condition | 82 | 5.41 | 1111 | 73.33 |

### 4.3.1 Combined results

The number of adverse events per indicator and per hospital based on the screening of the B-HDDS is described in Table 15.

**Table 15: Distribution of selected adverse events according to the B-HDDS**

| | Pressure Ulcers | | DVT | | Sepsis | | VAP | | Wound infection | | Tot |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Flagged (%) | Not Flagged (%) | Flagged (%) | Not Flagged (%) | Flagged (%) | Not Flagged (%) | Flagged (%) | Not Flagged (%) | Flagged (%) | Not Flagged (%) | N |
| Hosp 1 | 28 (17,2%) | 135 (82,8%) | 10 (6,1%) | 153 (93,9%) | 2 (1,2%) | 161 (98,8%) | 8 (4,9%) | 155 (95,1%) | 36 (22,1%) | 127 (77,9%) | 163 |
| Hosp 2 | 27 (13,5%) | 173 (86,5%) | 18 (9,0%) | 182 (91,0%) | 18 (9,0%) | 182 (91,0%) | 25 (12,5%) | 175 (87,5%) | 28 (14,0%) | 172 (86,0%) | 200 |
| Hosp 3 | 26 (13,0%) | 174 (87,0%) | 19 (9,5%) | 181 (90,5%) | 20 (10,0%) | 180 (90,0%) | 28 (14,0%) | 172 (86,0%) | 35 (17,5%) | 165 (82,5%) | 200 |
| Hosp 4 | 27 (13,5%) | 173 (86,5%) | 20 (10,0%) | 180 (90,0%) | 19 (9,5%) | 181 (90,5%) | 18 (9,0%) | 182 (91,0%) | 34 (17,0%) | 166 (83,0%) | 200 |
| Hosp 5 | 24 (12,6%) | 167 (87,4%) | 15 (7,9%) | 176 (92,1%) | 22 (11,5%) | 169 (88,5%) | 26 (13,6%) | 165 (86,4%) | 24 (12,6%) | 167 (87,4%) | 191 |
| Hosp 6 | 29 (13,1%) | 193 (86,9%) | 19 (8,6%) | 203 (91,4%) | 13 (5,9%) | 209 (94,1%) | 28 (12,6%) | 194 (87,4%) | 24 (10,8%) | 198 (89,2%) | 222 |
| Hosp 7 | 23 (15,0%) | 130 (85,0%) | 8 (5,2%) | 145 (94,8%) | 12 (7,8%) | 141 (92,2%) | 20 (13,1%) | 133 (86,9%) | 15 (9,8%) | 138 (90,2%) | 153 |
| Hosp 8 | 24 (12,9%) | 162 (87,1% | 26 (14,0%) | 160 (86,0%) | 14 (7,5%) | 172 (92,5%) | 24 (12,9%) | 162 (87,1%) | 27 (14,5%) | 159 (85,5%) | 186 |
| Total | 208 (13,7%) | 1307 (86,3%) | 135 (8,9%) | 1380 (91,1%) | 120 (7,9%) | 1395 (92,1%) | 177 (11,7%) | 1338 (88,3%) | 223 (14,7%) | 1292 (85,3%) | 1515 |

In general, postoperative wound infection and pressure ulcers were over-represented, with respectively 14.7% and 13.7% of all hospital discharges. Postoperative sepsis and deep vein thrombosis were less present amongst the selected medical records, respectively 7.9% and 8.9%.

The same calculation was made for the 5 selected indicators per hospital based on the medical record screening (Table 16). However, in order to compare results from the administrative data screening and the medical record screening, a correction was made for every case where an exclusion rule for a certain indicator in the B-HDDS applied. In other words, whenever an exclusion rule was present in the B-HDDS, the found indicator during medical record screening was neglected (Table 17).

**Table 16: Distribution of selected adverse events according to the medical record screening**

| | Pressure Ulcers | | DVT | | Sepsis | | VAP | | Wound infection | | Tot |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Flagged (%) | Not Flagged (%) | Flagged (%) | Not Flagged (%) | Flagged (%) | Not Flagged (%) | Flagged (%) | Not Flagged (%) | Flagged (%) | Not Flagged (%) | N |
| Hosp 1 | 54 (33,1%) | 109 (68,9%) | 6 (3,7%) | 157 (96,3%) | 7 (4,3%) | 156 (95,7%) | 3 (1,8%) | 160 (98,2%) | 38 (23,3%) | 125 (76,7%) | 163 |
| Hosp 2 | 33 (16,5%) | 167 (83,5%) | 9 (4,5%) | 191 (95,5%) | 16 (8,0%) | 184 (92,0%) | 13 (6,5%) | 187 (93,5%) | 29 (14,5%) | 171 (85,5%) | 200 |
| Hosp 3 | 45 (22,5%) | 155 (77,5%) | 7 (3,5%) | 193 (96,5%) | 30 (15,0%) | 170 (85,0%) | 14 (7,0%) | 186 (93,0%) | 40 (20,0%) | 160 (80,0%) | 200 |
| Hosp 4 | 43 (21,5%) | 157 (78,5%) | 16 (8,0%) | 184 (92,0%) | 31 (15,5%) | 169 (84,5%) | 5 (2,5%) | 195 (97,5%) | 30 (15,0%) | 170 (85,0%) | 200 |
| Hosp 5 | 36 (18,9%) | 155 (81,2%) | 11 (5,8%) | 180 (94,2%) | 25 (13,1%) | 166 (86,9%) | 13 (6,8%) | 178 (93,1%) | 27 (14,1%) | 164 (85,9%) | 191 |
| Hosp 6 | 43 (19,4%) | 179 (80,6%) | 11 (5,0%) | 211 (95,0%) | 25 (11,3%) | 197 (88,7%) | 10 (4,5%) | 212 (95,5%) | 20 (9,0%) | 202 (91,0%) | 222 |
| Hosp 7 | 21 (13,7%) | 132 (86,3%) | 6 (3,9%) | 147 (96,1%) | 19 (12,4%) | 134 (87,6%) | 10 (6,5%) | 143 (93,5%) | 11 (7,2%) | 142 (92,8%) | 153 |
| Hosp 8 | 26 (14,0%) | 160 (86,0%) | 18 (9,7%) | 168 (90,3%) | 19 (10,2%) | 167 (89,8%) | 14 (7,5%) | 172 (92,5%) | 28 (15,1%) | 158 (84,9%) | 186 |
| Total | 301 (19,9%) | 1214 (80,1%) | 84 (5,5%) | 1431 (94,5%) | 172 (11,4%) | 1343 (88,6%) | 82 (5,4%) | 1433 (94,6%) | 223 (14,7%) | 1292 (85,3%) | 1515 |

**Table 17: Distribution of adverse events according to the medical record screening corrected for B-HDDS exclusion rules**

| | Pressure Ulcers | | DVT | | Sepsis | | VAP | | Wound infection | | Tot |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Flagged (%) | Not Flagged (%) | Flagged (%) | Not Flagged (%) | Flagged (%) | Not Flagged (%) | Flagged (%) | Not Flagged (%) | Flagged (%) | Not Flagged (%) | N |
| Hosp 1 | 42 (25,8%) | 121 (74,2%) | 6 (3,7%) | 157 (96,3%) | 2 (1,2%) | 161 (98,8%) | 2 (1,2%) | 161 (98,8%) | 37 (22,7%) | 126 (77,3%) | 163 |
| Hosp 2 | 26 (13,0%) | 174 (87,0%) | 9 (4,5%) | 191 (95,5%) | 4 (2,0%) | 196 (98,0%) | 13 (6,5%) | 187 (93,5%) | 27 (13,5%) | 173 (86,5%) | 200 |
| Hosp 3 | 38 (19,0%) | 162 (81,0%) | 7 (3,5%) | 193 (96,5%) | 11 (5,5%) | 189 (94,5%) | 10 (5,0%) | 190 (95,0%) | 39 (19,5%) | 161 (80,5%) | 200 |
| Hosp 4 | 34 (17,0%) | 166 (83,0%) | 16 (8,0%) | 184 (92,0%) | 15 (7,5%) | 185 (92,5%) | 4 (2,0%) | 196 (98,0%) | 30 (15,0%) | 170 (85,0%) | 200 |
| Hosp 5 | 29 (15,2%) | 162 (84,8%) | 11 (5,8%) | 180 (94,2%) | 17 (8,9%) | 174 (91,1%) | 10 (5,2%) | 181 (96,8%) | 27 (14,1%) | 164 (85,9%) | 191 |
| Hosp 6 | 36 (16,2%) | 186 (83,8%) | 11 (5,0%) | 211 (95,0%) | 6 (2,7%) | 216 (97,3%) | 9 (4,1%) | 213 (95,9%) | 19 (8,6%) | 203 (91,4%) | 222 |
| Hosp 7 | 20 (13,1%) | 133 (86,9%) | 6 (3,9%) | 147 (96,1%) | 13 (8,5%) | 140 (91,5%) | 8 (5,2%) | 145 (94,8%) | 9 (5,9%) | 144 (94,1%) | 153 |
| Hosp 8 | 20 (10,8%) | 166 (89,2%) | 18 (9,7%) | 168 (90,3%) | 11 (5,9%) | 175 (94,1%) | 12 (6,5%) | 174 (93,6%) | 26 (14,0%) | 160 (86,0%) | 186 |
| Total | 245 (16,2%) | 1270 (83,8%) | 84 (5,5%) | 1431 (94,5%) | 79 (5,2%) | 1436 (94,2%) | 68 (4,5%) | 1447 (95,5%) | 214 (14,1%) | 1301 (85,9%) | 1515 |

Table 18 shows overall PPV in circumstances where the indicator might be present on admission when reviewing the medical record. A match is defined as an indicator being positive in the B-HDDS and in the medical record screening, regardless of the kind of indicator.

**Table 18: Overall PPV, POA not excluded from medical record screening**

| Table of B-HDDS by Medical Record Review (MRR) | | | |
|---|---|---|---|
| **B-HDDS** | **MRR** | | |
| Frequency<br>Row Pct | **No** | **Yes** | **Total** |
| No | 695<br>89.79 | 79<br>10.21 | 774 |
| Yes | 260<br>35.09 | 481<br>64.91 | 741 |
| Total | 955 | 560 | 1515 |

| cat | rate | Normal Approximation (L) | 95% CI (Upper) |
|---|---|---|---|
| PPV | 64.9123 | 61.4761 | 68.3485 |

Table 19 presented PPV when excluding the indicator from the medical record screening when it was present on admission. Given the fact that fewer indicators were judged as being a truly adverse event instead of comorbidity, the PPV diminished to 61.37%.

**Table 19: Overall positive predictive values, POA excluded from medical record screening**

| Table of B-HDDS by Medical Record Review (MRR) | | | |
|---|---|---|---|
| **B-HDDS** | **Medical Record Review** | | |
| Frequency<br>Row Pct | **No** | **Yes** | **Total** |
| No | 777<br>92.28 | 65<br>7.72 | 842 |
| Yes | 260<br>38.63 | 413<br>61.37 | 673 |
| Total | 1037 | 478 | 1515 |

| cat | rate | Normal Approximation (L) | 95% CI (Upper) |
|---|---|---|---|
| PPV | 61.3670 | 57.6884 | 65.0457 |

## 4.3.2 Results per indicator

For each indicator, the distribution of cases was presented by hospital and PPV was calculated for each indicator. In this context, there had to be an exact match between the specific indicator in the B-HDDS and the medical record screening.

### *4.3.2.1 Pressure ulcers*

86,897 cases, out of 285,617 hospital stays, were included in the denominator of indicator 'pressure ulcers'. Pressure ulcer was flagged positive in 1,777 cases included (2%).

As shown in Table 20, 13.73% of the overall files included in the medical review were flagged for pressure ulcers. This indicator was under-represented in hospitals 5 (12.57%) and 8 (12.90%) while it was overrepresented in hospitals 1 (17.18%) and 7 (15.03%).

**Table 20: Frequency flagged cases for pressure ulcers in the B-HDDS per hospital**

| | Number of flagged cases for pressure ulcers | Number of files for the medical record review | Rate (%) |
|---|---|---|---|
| Hosp 1 | 28 | 163 | 17.18 |
| Hosp 2 | 27 | 200 | 13.50 |
| Hosp 3 | 26 | 200 | 13.00 |
| Hosp 4 | 27 | 200 | 13.50 |
| Hosp 5 | 24 | 191 | 12.57 |
| Hosp 6 | 29 | 222 | 13.06 |
| Hosp 7 | 23 | 153 | 15.03 |
| Hosp 8 | 24 | 186 | 12.90 |
| Total | 208 | 1515 | 13.73 |

Table 21 and Table 22 represent the correspondence of B-HDDS and the medical record review for pressure ulcers. Table 21 presents results for the events not excluded from the medical record review. In other words, if the pressure ulcer was present on admission the event was not excluded from the results. On the other hand, Table 22 shows results with the exclusion of events present on admission on the basis of the medical record screening.

When including all cases flagged in the B-HDDS, the PPV was 74.52%. When cases found present on admission during the medical record review were excluded, the PPV was lower with 68.07%.

**Table 21 Positive predictive value for pressure ulcer, POA not excluded from medical record screening**

| Pressure ulcer - Table of B-HDDS by Medical Record Review (MRR) | | | |
|---|---|---|---|
| **B-HDDS** | **MRR** | | |
| Frequency<br>Row Pct | **No** | **Yes** | **Total** |
| No | 1220<br>93.34 | 87<br>6.66 | 1307 |
| Yes | 53<br>25.48 | 155<br>74.52 | 208 |
| Total | 1273 | 242 | 1515 |

| cat | rate | Normal Approximation (L) | 95% CI (Upper) |
|---|---|---|---|
| PPV | 74.5192 | 68.5974 | 80.4411 |

**Table 22: Positive predictive value for pressure ulcer, POA excluded from medical record screening**

| Pressure ulcer - Table of B-HDDS by Medical Record Review (MRR) | | | |
|---|---|---|---|
| **B-HDDS** | **MRR** | | |
| Frequency<br>Row Pct | **No** | **Yes** | **Total** |
| No | 1268<br>94.00 | 81<br>6.00 | 1349 |
| Yes | 53<br>31.93 | 113<br>68.07 | 166 |
| Total | 1321 | 194 | 1515 |

| cat | rate | Normal Approximation (L) | 95% CI (Upper) |
|---|---|---|---|
| PPV | 68.0723 | 60.9804 | 75.1642 |

### 4.3.2.2 *Postoperative Deep vein thrombosis and pulmonary embolism (DVT/PE)*

For the second indicator 'postoperative deep vein thrombosis and pulmonary embolism', 99,967 cases were included in the denominator. In total, only 451 (0.45%) of cases were flagged positive for this indicator.

Amongst the overall files included in the medical record review, 8.91% of cases were flagged in the B-HDDS for DVT/PE (Table 23). Hospitals 7 (5.23%) and 1 (6.13%) were under-represented for this indicator. DVT/PE is most represented in hospitals 8 (13.98%) and 4 (10.00%).

**Table 23: Frequency flagged cases for postoperative deep vein thrombosis/pulmonary embolism (DVT/PE) in the B-HDDS per hospital**

| | Number of flagged cases for DVT/PE | Number of files for the medical record review | Rate (%) |
|---|---|---|---|
| Hosp 1 | 10 | 163 | 6.13 |
| Hosp 2 | 18 | 200 | 9.00 |
| Hosp 3 | 19 | 200 | 9.50 |
| Hosp 4 | 20 | 200 | 10.00 |
| Hosp 5 | 15 | 191 | 7.85 |
| Hosp 6 | 19 | 222 | 8.56 |
| Hosp 7 | 8 | 153 | 5.23 |
| Hosp 8 | 26 | 186 | 13.98 |
| Total | 135 | 1515 | 8.91 |

Results on the second indicator, deep venous thrombosis and pulmonary embolism, are shown in Table 24 and Table 25. When disregarding the present on admission from the medical record screening (Table 24), the PPV was relatively low (58.52%). The rate decreased even further when excluding the present on admission form the medical record screening (54.47%) meaning that approximately half of the cases were present in the medical file when it was flagged in the B-HDDS.

**Table 24: Positive predictive value for postoperative DVT/PE, POA not excluded from medical record screening**

| DVT/PE - Table of B-HDDS by Medical Record Review (MRR) | | | |
|---|---|---|---|
| **B-HDDS** | **MRR** | | |
| Frequency<br>Row Pct | **No** | **Yes** | **Total** |
| No | 1391<br>99.93 | 1<br>0.07 | 1392 |
| Yes | 56<br>45.53 | 67<br>54.47 | 123 |
| Total | 1447 | 68 | 1515 |

| cat | rate | Normal Approximation (L) | 95% CI (Upper) |
|---|---|---|---|
| PPV | 54.4715 | 45.6708 | 63.272 |

**Table 25: Positive predictive value for postoperative DVT/PE, POA excluded from medical record screening**

| DVT/PE - Table of B-HDDS by Medical Record Review (MRR) | | | |
|---|---|---|---|
| **B-HDDS** | **MRR** | | |
| Frequency<br>Row Pct | **No** | **Yes** | **Total** |
| No | 1379<br>99.93 | 1<br>0.07 | 1380 |
| Yes | 68<br>50.37 | 67<br>49.63 | 135 |
| Total | 1447 | 68 | 1515 |

| cat | rate | Normal Approximation (L) | 95% CI (Upper) |
|---|---|---|---|
| PPV | 49.6296 | 41.1955 | 58.064 |

### *4.3.2.3 Postoperative sepsis*

26,718 cases were included in the denominator for postoperative sepsis. Out of these, 348 (1.30%) were positively flagged. Postoperative sepsis represented 7.92% of all cases reviewed (Table 26). In hospital 1, only 2 cases were reviewed (1.23%). Postoperative sepsis was the most represented in hospitals 5 (11.52%) and 3 (10.00%).

**Table 26: Frequency flagged cases for Postoperative sepsis in the B-HDDS per hospital**

| | Number of flagged cases for Postoperative sepsis | Number of files for the medical record review | Rate (%) |
|---|---|---|---|
| Hosp 1 | 2 | 163 | 1.23 |
| Hosp 2 | 18 | 200 | 9.00 |
| Hosp 3 | 20 | 200 | 10.00 |
| Hosp 4 | 19 | 200 | 9.50 |
| Hosp 5 | 22 | 191 | 11.52 |
| Hosp 6 | 13 | 222 | 5.86 |
| Hosp 7 | 12 | 153 | 7.84 |
| Hosp 8 | 14 | 186 | 7.53 |
| Total | 120 | 1515 | 7.92 |

Results for the indicator postoperative sepsis are shown in the Table 27 and Table 28. On the basis of the file record, postoperative sepsis was judged to be present on admission only for one case. Chances of finding postoperative sepsis in the medical record while it was indicated in administrative data were low (respectively 45.00% and 44.53%, respectively if POA included or not).

**Table 27: Positive predictive value for Postoperative sepsis, POA not excluded from medical record screening**

| **Postop sepsis - Table of B-HDDS by Medical Record Review (MRR)** | | | |
|---|---|---|---|
| **B-HDDS** | **MRR** | | |
| Frequency<br>Row Pct | **No** | **Yes** | **Total** |
| No | 1370<br>98.21 | 25<br>1.79 | 1395 |
| Yes | 66<br>55.00 | 54<br>45.00 | 120 |
| Total | 1436 | 79 | 1515 |

| **cat** | **rate** | **Normal Approximation (L)** | **95% CI (Upper)** |
|---|---|---|---|
| PPV | 45.0000 | 36.0989 | 53.9011 |

**Table 28: Positive predictive value for postoperative sepsis, POA excluded from medical record screening**

| **Postop sepsis - Table of B-HDDS by Medical Record Review (MRR)** | | | |
|---|---|---|---|
| **B-HDDS** | **MRR** | | |
| Frequency<br>Row Pct | No | Yes | **Total** |
| No | 1371<br>98.21 | 25<br>1.79 | 1396 |
| Yes | 66<br>55.46 | 53<br>44.54 | 119 |
| Total | 1437 | 78 | 1515 |

| cat | rate | Normal Approximation (L) | 95% CI (Upper) |
|---|---|---|---|
| PPV | 44.5378 | 35.6081 | 53.4675 |

### *4.3.2.4 Ventilator associated pneumonia (VAP)*

4,282 cases were included in the denominator for this indicator out of which 1,025 (23.94%) were flagged positive.

Amongst all the files reviewed during the screening, 177 (11.68%) concerned ventilator associated pneumonia (Table 29). VAP was less represented in hospitals 1 (4.91%) and 4 (9.00%) while it was most represented in hospitals 3 (14.00%) and 5 (13.61%).

**Table 29: Frequency flagged cases for Ventilator-associated pneumonia (VAP) in the B-HDDS per hospital**

| | Number of flagged cases for VAP | Number of files for the medical record review | Rate (%) |
|---|---|---|---|
| Hosp 1 | 8 | 163 | 4.91 |
| Hosp 2 | 25 | 200 | 12.50 |
| Hosp 3 | 28 | 200 | 14.00 |
| Hosp 4 | 18 | 200 | 9.00 |
| Hosp 5 | 26 | 191 | 13.61 |
| Hosp 6 | 28 | 222 | 12.61 |
| Hosp 7 | 20 | 153 | 13.07 |
| Hosp 8 | 24 | 186 | 12.90 |
| Total | 177 | 1515 | 11.68 |

Results on ventilator associated pneumonia are exactly the same when in- or excluding the present on admission from the medical record screening (Table 30 and Table 31). Chances of finding real ventilator associated pneumonia in the medical record when it is indicated in administrative data are very low (29.94%).

**Table 30: Positive predictive value for Ventilator associated pneumonia**

| VAP - Table of B-HDDS by Medical Record Review (MRR) | | | |
|---|---|---|---|
| B-HDDS | MRR | | |
| Frequency<br>Row Pct | **No** | **Yes** | **Total** |
| No | 1323<br>98.88 | 15<br>1.12 | 1338 |
| Yes | 124<br>70.06 | 53<br>29.94 | 177 |
| Total | 1447 | 68 | 1515 |

| cat | rate | Normal Approximation (L) | 95% CI (Upper) |
|---|---|---|---|
| PPV | 29.9435 | 23.1961 | 36.6909 |

### *4.3.2.5 Postoperative wound infection*

99,787 cases were included in the denominator for the indicator 'postoperative wound infection'. In 1,302 (1.30%) cases, postoperative wound infection was flagged positive.

Amongst all medical files screened, 223 (14.72%) concerned an event of postoperative wound infection (Table 31). This indicator was less represented in hospitals 6 (10.81%) and 5 (12.57%) and more represented in hospitals 7 (28.30%) and 1 (22.86).

**Table 31: Frequency flagged cases for Postoperative wound infection (PWI) in the B-HDDS per hospital**

| | Number of flagged cases for PWI | Number of files for the medical record review | Rate (%) |
|---|---|---|---|
| Hosp 1 | 36 | 163 | 22.86 |
| Hosp 2 | 28 | 200 | 14.00 |
| Hosp 3 | 35 | 200 | 17.50 |
| Hosp 4 | 34 | 200 | 17.00 |
| Hosp 5 | 24 | 191 | 12.57 |
| Hosp 6 | 24 | 222 | 10.81 |
| Hosp 7 | 15 | 153 | 28.30 |
| Hosp 8 | 27 | 186 | 14.52 |
| Total | 223 | 1515 | 14.72 |

Results for postoperative wound infection are shown in Table 32 and Table 33. Chances of finding a real postoperative wound infection when it was indicated by administrative data was 69.06%. When excluding present on admission from the medical screening, it decreased to 62.33%.

**Table 32: Positive predictive value for postoperative wound infection (PWI), POA not excluded from medical record screening**

| PWI - Table of B-HDDS by Medical Record Review (MRR) | | | |
|---|---|---|---|
| B-HDDS | MRR | | |
| Frequency<br>Row Pct | **No** | **Yes** | **Total** |
| No | 1234<br>95.51 | 58<br>4.49 | 1292 |
| Yes | 69<br>30.94 | 154<br>69.06 | 223 |
| Total | 1303 | 212 | 1515 |

| cat | rate | Normal Approximation (L) | 95% CI (Upper) |
|---|---|---|---|
| PPV | 69.0583 | 62.9913 | 75.1253 |

**Table 33: Positive predictive value for postoperative wound infection (PWI), POA excluded from medical record screening**

| PWI - Table of B-HDDS by Medical Record Review (MRR) | | | |
|---|---|---|---|
| B-HDDS | MRR | | |
| Frequency<br>Row Pct | **No** | **Yes** | **Total** |
| No | 1260<br>96.40 | 47<br>3.60 | 1307 |
| Yes | 69<br>33.17 | 139<br>66.83 | 208 |
| Total | 1329 | 186 | 1515 |

| cat | rate | Normal Approximation (L) | 95% CI (Upper) |
|---|---|---|---|
| PPV | 66.8269 | 60.4283 | 73.2255 |

### 4.3.3 Results per hospital and per team

Logistic regression was performed in the 6 models – one per indicator and one for all indicators. The response variable was the correspondence between the medical record screening and the screening of the B-HDDS. If the screening result of the B-HDDS were confirmed by the medical record screening, the value of the response variable was '1', otherwise this value was set to '0'. We included the 'medical record screening team' and the hospital as regressors in the model.

The results of these analyses, illustrated in Table 34, show that there was no significant effect between the regressors and the response variable in all 6 models.

**Table 34 Effect of hospital or medical record screening team on the correspondence of results between B-HDDS screening and medical record screening, tested by logistic regression analyses**

| | Pressure Ulcers | | DVT | | Sepsis | | VAP | | Wound infection | | Overall | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Odds ratio (95% CI) | p-value | Odds ratio (95% CI) | p-value | Odds ratio (95% CI) | p-value | Odds ratio (95% CI) | p-value | Odds ratio (95% CI) | p-value | Odds ratio (95% CI) | p-value |
| **Effect of hospital under study** | 0.859 (0.733-1.006) | 0.0594 | 1.072 (0.848-1.355) | 0.5598 | 0.939 (0.775-1.138) | 0.5227 | 0.963 (0.823-1.128) | 0.6422 | 0.999 (0.848-1.178) | 0.9929 | 0.942 (0.845-1.051) | 0.2875 |
| **Effect of medical record screening team** | 7.891 (0.933-3.832) | 0.0770 | 0.927 (0.326-2.634) | 0.8866 | 1.346 (0.572-3.167) | 0.4963 | 1.018 (0.504-2.057) | 0.9605 | 1.076 (0.517-2.239) | 0.8452 | 1.171 (0.719-1.906) | 0.5258 |

### 4.3.4 Codification analysis

Comparing the results obtained from administrative data versus medical record review, a detailed analysis was made on all false negative and false positive cases in the selected hospitals. In this regard, the aim was to try to found an answer why mismatches occurred between the two sets of data. Therefore, whenever a false positive or negative result originated, a re-evaluation was performed on gained information from the medical record review.

Amongst the 1515 medical files reviewed, teams found 194 cases (12%) with an indication of pressure ulcer not present on admission. As shown in Table 22, pressure ulcer was not referenced in administrative data in 81 cases (6%). In 34 out of 81 cases (42%) it concerned pressure ulcers of grade one, in 24 cases (30%) grade and in 13 cases (16%) grade three pressure ulcers. In 10 cases (12%) no specific grade of pressure ulcer could be found during medical record review. Pressure ulcers were just mentioned in the medical files without clinical description. Fifty-three cases were false positive, in which no indication of pressure ulcers was found during medical record review. All of them were judged as being over reported.

After excluding cases in which postoperative DVT/PE were present on admission, one case was found in the medical file and was not present in the administrative data (Table 25). On the other hand, 68 cases were false positive. For 41 of these (60.3%), it was considered as a problem of over reporting. Twelve cases (17.7%) appeared to be present on admission and thus per definition no adverse event was considered. In 6 cases (8.8%) no DVT/PE could be scored due to the stringency of the used clinical criteria. In 8 cases (11.8%) the ICD-9-CM code for this indicator didn't seem to be specific enough, in these files an indication for a superficial phlebitis was found. Finally, in one case (0.1%), reviewers found a pulmonary embolism that occurred during the inhospital stay but before the intervention.

For postoperative sepsis excluding event present on admission, a total of 67 false positive cases were found of which 32 (47.8%) were a matter of over reporting. Twelve cases (17.6%) had an indication of postoperative infection, but clinical criteria appeared to be too stringent to evaluate the file as postoperative sepsis during medical record review. In an important part of false positive files (21 out of 37[c]), the ICD-9-CM code wasn't specific enough to distinguish postoperative sepsis from postoperative haemorrhagic shock (15 cases - 71%) and postoperative cardiogenic shock (6 cases – 29%). For one case (0.2%), the sepsis occurred during the inhospital stay but before the surgical intervention. In one (0.2%) case the patient presented with sepsis without a previous recent surgical procedure.

As regards ventilator associated pneumonia[d], 8 cases appeared to be false negative and judged as underreporting. In 58 cases, a false positive result was withheld of which only in 5 files (8.6%) no indication of respiratory infection could be found so that these were judged as being a clear matter of over reporting. In contrast, 34 cases (58.6%) had a clear diagnosis of respiratory infection in the medical file, but the clinical criteria used for ventilator associated pneumonia were too stringent to conclude to this indicator. In another 11 cases (19.0%) there was an indication of aspiration pneumonia during medical record screening, indicating a lack of specificity of the ICD-9-CM code for this indicator. Finally, in 8 cases (13.8%) pneumonia or aspiration pneumonia appeared to be present on admission.

For postoperative wound infection, 84 cases were false positive. Amongst those cases, 25 (29.8%) had an infection present on admission. For 46 cases (54.8%), reviewers found no information in the medical files or mention of absence of infection, it was considered as over reporting. In 11 cases (13.1%) a postoperative wound infection could not be scored during medical record review due to the stringency of clinical criteria although there was an indication of wound infection in the files. Finally, in 2 cases (6.1%) the ICD-9-CM didn't seem to be specific enough.

## 4.4 CHARACTERISTICS OF EVENTS

When analysing the answers to the two questions on which basis an event was defined as an evident adverse event, we noticed that the distribution of answers was different between the two teams who reviewed the medical files. Table 35 shows how the two teams answered to a more general question: "By what was the patient's injury/complication caused?" Note that its results don't exclude adverse events present on admission. Team 1 generally attributed cause of adverse events to the health care management. On the other hand, in more than a third of the cases, team 2 estimated that they didn't get sufficient information to decide if an adverse event was due to the health care management or not.

**Table 35 Distribution of the cause of adverse events found in medical files by team**

| | | Team 1 | | Team 2 | |
|---|---|---|---|---|---|
| | | Frequency | Percent | Frequency | Percent |
| 1 | Health care management | 197 | 75.19 | 50 | 20.04 |
| 2 | Health care management interacting with disease process | 63 | 24.05 | 45 | 18.37 |
| 3 | Solely by disease process | 2 | 0.76 | 52 | 21.22 |
| 4 | Not documented | 0 | 0.00 | 98 | 40.00 |
| | Total | 262 | 100.00 | 245 | 100.00 |
| | *Missing* | *0* | | *0* | |

As shown in Table 36, team 1 estimated that more than 85% of cases presented moderate to strong evidence that health care management was responsible for the adverse event. Team 2 presented less evident cases of management causation as only 22 % of cases have moderate to strong evidence.

[c] Only one team noted the diagnosis for this event
[d] Available only for one team

**Table 36 Distribution by team of degree of evidence that health care management is responsible for the adverse event**

| | | Team 1 Frequency | Percent | Team 2 Frequency | Percent |
|---|---|---|---|---|---|
| 1 | Virtually no evidence for management causation/system failure | 1 | 0.38 | 64 | 27.23 |
| 2 | Slight-to-modest evidence for management causation | 0 | 0.00 | 50 | 21.28 |
| 3 | Management causation not likely; less than 50-50 but close call | 9 | 3.41 | 20 | 8.51 |
| 4 | Management causation more likely than not, more than 50-50 but close call | 22 | 8.33 | 49 | 20.85 |
| 5 | Moderate/strong evidence for management causation | 70 | 26.52 | 42 | 17.87 |
| 6 | Virtually certain evidence for management causation | 162 | 61.36 | 10 | 4.26 |
| | *Missing* | | *34* | | *27* |

Evaluation of the degree of preventability of the adverse events was estimated equally by the two teams (Table 37). Team 1 usually attributed a strong evidence of preventability of events found in medical files. Inversely, team 2 was more moderate and cautious.

**Table 37 Distribution by team of degree of preventability of the event**

| | | Team 1 Frequency | Percent | Team 2 Frequency | Percent |
|---|---|---|---|---|---|
| 1 | Virtually no evidence for preventability | 14 | 5.30 | 38 | 16.38 |
| 2 | Slight-to-modest evidence for preventability | 6 | 2.27 | 43 | 18.53 |
| 3 | Preventability not likely, less than 50-50 but close call | 15 | 5.68 | 18 | 7.76 |
| 4 | Preventability more likely than not, more than 50-50 but close call | 27 | 10.23 | 57 | 24.57 |
| 5 | Strong evidence for preventability | 85 | 32.20 | 63 | 27.16 |
| 6 | Virtually certain evidence for preventability | 117 | 44.32 | 13 | 5.60 |
| | *Missing* | | *34* | | *30* |

Distribution of causation of management of care (Table 38) and degree of preventability (Table 39) can also be observed for each indicator by team.

**Table 38 Distribution, by team and indicator, of degree of evidence that health care management is responsible for the adverse event**

| HCM | | Decubitus ulcer Frequency | Percent | DVT/PE Frequency | Percent | VAP Frequency | Percent | Postoperative sepsis Frequency | Percent | PWI Frequency | Percent |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Team1 | 1 | 0 | 0.00 | 0 | 0.00 | 0 | 0.00 | 0 | 0.00 | 1 | 0.79 |
| | 2 | 0 | 0.00 | 0 | 0.00 | 0 | 0.00 | 1 | 1.47 | 1 | 0.79 |
| | 3 | 0 | 0.00 | 7 | 23.33 | 1 | 2.94 | 2 | 2.94 | 5 | 3.97 |
| | 4 | 3 | 2.33 | 14 | 46.67 | 17 | 50.00 | 13 | 19.12 | 10 | 7.94 |
| | 5 | 28 | 21.71 | 6 | 20.00 | 16 | 47.06 | 15 | 22.06 | 44 | 34.92 |
| | 6 | 98 | 75.97 | 3 | 10.00 | 0 | 0.00 | 37 | 54.41 | 65 | 51.59 |
| Team2 | 1 | 24 | 26.67 | 21 | 53.85 | 10 | 27.03 | 15 | 24.59 | 11 | 15.49 |
| | 2 | 16 | 17.78 | 10 | 25.64 | 17 | 45.95 | 19 | 31.15 | 15 | 21.13 |
| | 3 | 4 | 4.44 | 0 | 0.00 | 3 | 8.11 | 8 | 13.11 | 11 | 15.49 |
| | 4 | 22 | 24.44 | 4 | 10.26 | 4 | 10.81 | 11 | 18.03 | 15 | 21.13 |
| | 5 | 22 | 24.44 | 2 | 5.13 | 2 | 5.41 | 6 | 9.84 | 15 | 21.13 |
| | 6 | 2 | 2.22 | 2 | 5.13 | 1 | 2.70 | 2 | 3.28 | 4 | 5.63 |

**Table 39 Distribution, by team and indicator, of degree of preventability of the adverse event**

| | Prev | Decubitus ulcer Frequency | Decubitus ulcer Percent | DVT/PE Frequency | DVT/PE Percent | VAP Frequency | VAP Percent | Postoperative sepsis Frequency | Postoperative sepsis Percent | PWI Frequency | PWI Percent |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Team1 | 1 | 0 | 0.00 | 22 | 73.33 | 0 | 0.00 | 0 | 0.00 | 1 | 0.79 |
| | 2 | 1 | 0.00 | 4 | 13.33 | 1 | 2.94 | 3 | 4.41 | 5 | 3.97 |
| | 3 | 1 | 0.00 | 2 | 6.67 | 8 | 23.53 | 12 | 17.65 | 11 | 8.73 |
| | 4 | 4 | 3.10 | 1 | 3.33 | 24 | 70.59 | 12 | 17.65 | 24 | 19.05 |
| | 5 | 32 | 24.81 | 1 | 3.33 | 1 | 2.94 | 18 | 26.47 | 56 | 44.44 |
| | 6 | 91 | 70.54 | 0 | 0.00 | 0 | 0.00 | 23 | 33.82 | 29 | 23.02 |
| Team2 | 1 | 15 | 17.24 | 12 | 29.27 | 3 | 8.11 | 8 | 13.33 | 7 | 9.86 |
| | 2 | 12 | 13.79 | 8 | 19.51 | 14 | 37.84 | 21 | 35.00 | 15 | 21.13 |
| | 3 | 4 | 4.60 | 4 | 9.76 | 8 | 21.62 | 10 | 16.67 | 4 | 5.63 |
| | 4 | 19 | 21.84 | 6 | 14.63 | 4 | 10.81 | 13 | 21.67 | 20 | 28.17 |
| | 5 | 31 | 35.63 | 9 | 21.95 | 4 | 10.81 | 5 | 8.33 | 20 | 28.17 |
| | 6 | 6 | 6.90 | 2 | 4.88 | 4 | 10.81 | 3 | 5.00 | 5 | 7.04 |

We suggested that an evident adverse event could be defined as an adverse event with high management causation (more than 3) or with high preventability (more than 3)[14, 41, 42]. When an adverse event was flagged in the B-HDDS and also identified in the medical files as not present on admission, the reviewers estimated the event as evident in more than 3 out of 4 cases (Table 40). However this proportion was very different between the 2 reviewer teams: Team 1 retrieved evident adverse event in most of cases when team 2 defined it in 60% of the cases.

**Table 40 Proportion of events defined as evident (high preventability or high management causation) when present in B-HDDS and in the medical files but not present on admission**

| | Total | | Team 1 | | Team 2 | |
|---|---|---|---|---|---|---|
| Adverse event evident | Frequency | Percent | Frequency | Percent | Frequency | Percent |
| 0 | 99 | 20.71 | 8 | 3.24 | 91 | 39.39 |
| 1 | 379 | 79.29 | 239 | 96.76 | 140 | 60.61 |

Amongst the five indicators selected, the one showing the most evident adverse event was decubitus ulcer with 88% (Table 41). The less represented was Ventilator-associated pneumonia with 55% of evident adverse events. Results were different between the two teams but in the same proportions as in all cases, decubitus ulcer was the most "evident" indicator and ventilator-associated pneumonia was the less evident.

**Table 41 Proportion of events defined as evident (high preventability or high management causation) when present in B-HDDS and in the medical files but not present on admission – by indicator**

| | Total | | Team 1 | | Team 2 | |
|---|---|---|---|---|---|---|
| | Frequency | Percent | Frequency | Percent | Frequency | Percent |
| Decubitus Ulcer | | | | | | |
| 0 | 14 | 12.39 | 0 | 0.00 | 14 | 31.11 |
| 1 | 99 | 87.61 | 68 | 100.00 | 31 | 68.89 |
| Postoperative sepsis | | | | | | |
| 0 | 16 | 30.19 | 1 | 4.55 | 15 | 48.39 |
| 1 | 37 | 69.81 | 21 | 95.45 | 16 | 51.61 |
| Ventilator-associated pneumonia | | | | | | |
| 0 | 24 | 45.28 | 1 | 4.76 | 23 | 71.88 |
| 1 | 29 | 54.72 | 20 | 95.24 | 9 | 28.13 |
| Postoperative wound infection | | | | | | |
| 0 | 22 | 15.83 | 3 | 3.49 | 19 | 35.85 |
| 1 | 117 | 84.17 | 83 | 96.51 | 34 | 64.15 |
| Deep vein thrombosis / Pulmonary embolism | | | | | | |
| 0 | 27 | 40.30 | 7 | 25.00 | 20 | 51.28 |
| 1 | 40 | 59.70 | 21 | 75.00 | 19 | 48.72 |

# 5 DISCUSSION

The aim of the current study was to validate the accuracy of the B-HDDS to detect 5 adverse events by means of 5 indicators: pressure ulcers, deep venous thrombosis or pulmonary embolism (DVT/PE), postoperative sepsis, ventilator associated pneumonia (VAP) and postoperative wound infection. For this project, we used the definition of adverse event from Wilson et al[14]. According to these authors, an adverse event is "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 disease". As near-misses are unlikely to be found in the medical files and to be coded in administrative data, we prefer to focus on adverse events.

## 5.1 SYNTHESIS OF STUDY RESULTS

The international prevalence of adverse events has been evaluated between 2.9% [41, 42] and 16.6%[14]. This emphasises the need to improve processes of care so that the complication rate in acute care hospitals can be reduced. In order to achieve this goal however, one must define a reliable process to find these adverse events. Since medical inpatient diagnoses can be identified in administrative databases through the use of ICD-9-CM codes, these codes provide a very inexpensive and readily accessible source of clinical information. Databases have therefore been extensively studied throughout the years to assess the validity of the use of ICD-9-CM codes for complications in order to compare providers' performances. Given the fact that the Belgian Hospital Discharge Dataset is compulsory for all inpatient in acute hospitals, the system can be considered representative for the care performed at Belgian acute hospitals. The dataset contains patient demographics, information about the hospital stay (date and type of admission/discharge, referral data, admitting department, destination after discharge,…), as well as clinical information (primary and secondary diagnoses and therapeutic procedures as described in the ICD-9-CM).

The aim of the current study was to validate the accuracy of the B-HDDS to detect a group of 5 adverse events indicators: pressure ulcers, deep venous thrombosis or pulmonary embolism (DVT/PE), postoperative sepsis, ventilator associated pneumonia (VAP) and postoperative wound infection. Since 'near misses' are unlikely to be found in the medical record, we used the definition of an adverse event by Wilson et al[14]. According to this definition, an adverse event is 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 disease.

Using retrospective medical record review with predefined strict clinical criteria as the gold standard, we searched for evidence supporting the presence of any of the 5 adverse events which were then compared with the ICD-9-CM discharge diagnoses. The overall positive predictive value was 64.91%. When correcting for events not present on admission, it decreased to 61.37%. Considering all indicators, the positive predictive values were relatively low, ranging from 29.94% for ventilator associated pneumonia up to 75.52% for pressure ulcer. These values were reconfirmed when excluding events present on admission. Values then ranged from 29.94% for ventilator associated pneumonia up to 68.07% for pressure ulcer. Even though AHRQ suggests defining the limit of positive predictive values to 75%[56], we have observed that this limit couldn't be met. The closest rate was reflected in the 2/3rd of the pressure ulcer cases flagged in the B-HDDS where the event occurred during hospital stay. These results can be put into perspective by taking into consideration the fact that the stringency of the clinical criteria defined for the medical record review was, as described in the next paragraph, too strict in some instances (e.g. ventilator associated pneumonia).

**Table 42 Positive Predictive Value by indicator present (or not) on admission (POA)**

| Indicator | PPV (%) |
|---|---|
| Overall | 64.91 |
| Pressure ulcer | 74.52 |
| Postoperative DVT/PE | 58.52 |
| Postoperative sepsis | 45.00 |
| Ventilator associated pneumonia | 29.94 |
| Postoperative wound infection | 69.06 |
| Overall, not POA | 61.37 |
| Pressure ulcer, not POA | 68.07 |
| Postoperative DVT/PE, not POA | 54.47 |
| Postoperative sepsis, not POA | 44.54 |
| Ventilator associated pneumonia, not POA | 29.94 |
| Postoperative wound infection, not POA | 66.83 |

## 5.2 MISMATCH MEDICAL RECORD REVIEW VERSUS ADMINISTRATIVE DATA

In general, previous work notified against the general use of ICD-9-CM codes to measure adverse event occurrence. Explanation for the weakness of adverse event data in administrative files is probably of multi-factorial origin. Possible causes include the lack of incentives for coding them, the vague clinical content of ICD-9-CM definitions, and failure or reluctance of clinicians to list complication diagnoses at discharge. Furthermore, the administrative database often fails to distinguish a condition present on admission from an adverse event occurring during the hospital stay[71].

In order to understand the mismatches between medical record review and administrative data, the current study also provided a more detailed analysis on all false negative and false positive cases by re-evaluating the medical records involved.

Pressure ulcers were especially prone to be underreported with almost 35% of all decubitus ulcer found during medical review not reported in the B-HDDS. However, nearly 42% of these cases concerned grade one pressure ulcers with no disability for the patient. Correcting for this low grade of pressure ulcer will probably result in a more accurate detection of the indicator through administrative data.

Only one false negative was found for DVT/PE confirming an accurate detection of this adverse event through administrative data. This can probably be explained by the rarity of this illness which is less likely to be randomly found during the medical record review. Moreover, its severity makes it more prone to being reported by practitioners in the patient's medical file.

Several explanations could be provided for false positive cases. Given the much larger portion of true negative cases compared to true positive cases, the percentage of false positive results was relatively low, ranging from 4.7% for DVT/PE to up to 8.6% for VAP.

Over reporting was the only explanation for the false positive results on pressure ulcers and was also a clarification for DVT/PE (60%), postoperative wound infection (45.4%) and postoperative sepsis (47.8%) in a significant amount of cases. The indicator appeared to be present on admission in 17.7% for DVT/PE and 29.8% for postoperative wound infection.

The ICD-9-CM code for postoperative sepsis was clearly not specific enough, since in nearly 57% of cases no distinction was made between postoperative sepsis, postoperative haemorrhagic shock and postoperative cardiogenic shock. Also for ventilator associated pneumonia (19%) and DVT/PE (11.8%) the ICD-9-CM codes lacked specificity.

The main problem for ventilator associated pneumonia however was the stringency of used clinical criteria. In 58.6% of cases a clear diagnosis of respiratory infection was found during medical record screening, but an insufficient amount of retained clinical criteria prohibited concluding a VAP. This finding is consistent with recent work by Klompas and colleagues[90] in which it appears that the clinical diagnosis of VAP is notoriously inaccurate. The physicians' ability to diagnose VAP is poor because many pulmonary complications of intensive care present with similar clinical signs. Therefore, the author conclude that due to the difficulty in rendering an accurate diagnosis of VAP and the subjective nature of the Centers for Disease Control and Prevention (CDC) criteria make VAP an unreliable basis for either internal quality control or interhospital benchmarking of quality care. Clinical criteria also appeared to be too strict in a less explicit way for postoperative wound infection (27.3%) and for postoperative sepsis (16.2%).

## 5.3 PRIOR RESULTS

The work of Iezzoni and colleagues in the early 1990s was the first systematic exploration of the value of administrative data in quality and patient safety research. They developed the Complications Screening Program (CSP), a computerized algorithm to screen for potential problems with the quality of hospital care using readily available discharge abstract data. The goal of most of the 27 screens was to identify facilities with higher-than-expected rates of complications, following adjustment for a number of patient characteristics, including chronic illnesses. Further study by the same author revealed that cases with complications had significant higher hospital mortality, a longer hospital stay and a higher total hospital cost. Despite these relationships, models aiming to predict complications using a variety of patient and hospital variables produced modest results, suggesting that the complications identified by the CSP were not easily explained using information available from administrative data[91]. In later work, the CSP comprised 28 complications screens in which it was validated as a quality indicator by using explicit process of care criteria to determine whether hospital discharges flagged by the CSP experienced more process problems than unflagged discharges. No statistically significant differences were found in the absolute number of process problems across flagged and unflagged cases[77]. Lawthers and colleagues[76] examined the validity of the CSP on 1,298 cases from California and Connecticut, of which 813 were surgical cases and 485 were medical cases. For the purpose of identifying in-hospital events, the surgical screens validated much better than the medical screens. Six surgical screens validated particularly well in terms of code corroboration and timing assumptions, with overall confirmation rates > 80%. Especially for medical screens, the difficulty was determining whether a code represented a pre-existing condition present on admission or a condition arising from the hospitalization. Our work support this problem, since especially 'pressure ulcers', as a medical screen, was particularly sensitive for the present on admission coding.

The CSP aimed to identify cases for in-depth review, not to make absolute judgments. Ten surgical screens had VPP of ~88% or higher and would be good-to-excellent candidates as screens for complications[76]. Only one medical screen – postprocedural haemorrhage or haematoma – appeared useful as a screening tool. The authors concluded not to recommend using the other medical screens to screen for complications. They suggested that the addition of an indicator of the timing of secondary diagnosis would significantly increase the utility of claims data for identifying in-hospital complications[76]. Naessens[92] and Glance[93-95] agreed that adding a 'condition present at admission' modifier to administrative data would significantly increase the value of clinical information in these data. Weingart et al[75] examined the accuracy of the CSP. Using physician judgments as the gold standard, they found that the flagged complication was present in 68.4% of surgical and 27.2% of medical cases.

McCarthy et al[74] examined, by using explicit criteria, medical records for objective clinical evidence supporting hospital-assigned ICD-9-CM discharge diagnoses and procedures codes that could represent complications. Overall, 30% of medical and 19% of surgical patients lacked any documented evidence in the medical record, even physicians' notes. These findings raised questions about whether the clinical conditions represented by ICD-9-CM codes used by some CSP were in fact present. Fry et al[23] demonstrated that the addition of a present-on-admission code, numerical laboratory

data, and vital signs to standard administrative data could greatly improve the accuracy of predictions of adverse outcomes associated with two selected surgical procedures (mortality rates following craniotomy and rates of postoperative sepsis after elective surgical procedures).

Geraci et al[71] tested the ability of ICD-9-CM codes in discharge abstracts to identify medical inpatients that experienced an in-hospital complication, using complications identified through chart review as the gold standard. Two sets of ICD-9-CM codes were used : an inclusive set including many medical diagnoses that may also be coexistent complicating conditions on admission rather than complications and an exclusive set consisting primarily of ICD-9-CM-specified complication and adverse drug events codes. Neither set appeared to perform well as a diagnostic test for complication occurrence. Positive predictive values were 32% in the inclusive set and 37% in the exclusive set. The authors concluded that administrative data significantly underestimated complication frequency in the selected medical patient population. Likewise, Moro and colleagues[96] stated that the Italian hospital discharge database could not be used to monitor postoperative infections developed during hospital stay. Romano et al[70] studied how accurately postoperative complications from elective lumbar diskectomies were reported in administrative data and found a markedly underreporting. They concluded that most complications were defined so vaguely in ICD-9-CM, or were so dependent on physician documentation and coder interpretation, that hospitals had difficulty reporting them consistently.

In contrast, Arnason et al[81] measured the accuracy of ICD-9-CM codes for bleeding and thrombo-embolic diagnoses in one Canadian university hospital. Compared to the gold standard chart review, the ICD-9-CM codes for identifying definite bleeding and major bleeding had a positive predictive value greater than 87%. However, the ICD-9-CM codes for identifying acute thrombo-embolism have a positive predictive value of only 63%. The authors concluded that it was valid to use bleeding codes as indicators of bleeding events as opposed to acute thrombo-embolism.

## 5.4 CHARACTERISTICS OF THE CLASSIFICATION OF ADVERSE EVENTS

When an adverse event was identified in the patient's file, reviewers were requested to qualify the event in terms of causes, health care management causation and degree of preventability. We have observed that both teams have interpreted these questions in different manners, re-questioning the reproducibility of this part of the abstraction tool. One possible hypothesis suggests that team 1 might have misinterpreted some questions. Indeed they defined some adverse events present on admission as being highly preventable and caused by health care management while team 2 didn't qualify such cases. Moreover, whilst team 1 would almost systematically put the responsibility on health care management, team 2 would put more focus on the morbidity of the patient: responsibility was taken off health care management as from the point where medication was provided to the patient, with no further check that posology was adapted to the situation. We can only assume that the true explanation lies somewhere in the middle.

Despite these differences, some trends have been identified. Both teams qualified the DVT/PE as being the event less likely to be preventable or caused by health care management (all: 60%, team 1: 75%, team 2: 49%). Inversely, the pressure ulcers are events that are the most preventable or most likely to be caused by health care management (all: 80%, team 1: 100%, team 2: 69%). This leads to observe that both teams tended to evaluate these two specific events similarly.

For the VAP however, the approaches of the teams diverged completely. Team 1 estimated that 95% of VAP were highly preventable or caused by health care management while the proportion went down to 28% as far as team 2 was concerned. In this specific case, an explanation could be that both teams not only had very different specialties (internal medicine vs. surgery) but also different points in safety culture between two generations.

Results of this study should however be toned downed as, when considering the files audited and flagged by RCM as a whole, only 60% of them can be considered as true positive adverse events occurring during hospital stay. Within these true positive

events, and based upon the review team considered, 40 to 5% of cases are non preventable or are not considered as related to the health care management. As shown in a previous study the rate of preventable adverse events may vary widely based on the analysis method used or, in our case, based on the review team assessing the medical records.

## 5.5 POTENTIALS AND CONDITIONS

Taking the current and previous international results into account, administrative data do have potential to accurately screen for adverse events. In our opinion, some conditions have to be met in order to achieve this goal:

- Refinement of administrative data by adding a 'present on admission' code. This code was recently added to the B-HDDS with a new version to begin the second semester of 2008.
- Increasing the specificity of the ICD-9-CM codes. Transition to the ICD-10-codes might already partially resolve this problem.
- At present, coding is based on the medical record only in some hospitals. Looking at the whole patient file, which includes nursing notes, would probably increase the accuracy of administrative data of those hospitals.
- Attention should be paid to the algorithm used for screening administrative data.
- Transparency in results on adverse events based on administrative data is essential. An approach of benchmarking have to be considered carefully but it will attract the attention of hospitals and will probably increase the quality of coding.

## 5.6 STUDY LIMITATIONS

Our study had several limitations. For individual indicators, the number of cases examined was relatively small, although it was sufficient to make reasonable assessments about validity. The design we had to use to attain a sufficient number of cases, however, unabled us to calculate sensitivity and specificity of the indicators.

The explicit clinical criteria were developed by one physician based on the international literature which was sometimes spares for some indicators and controversial for others. In order to partially resolve this problem, clinical criteria were presented to a panel of experts and adjustments were made whenever required.

The retrospective medical record review of the selected cases may not represent a true gold standard. First, no inter-rater variability test was performed by means of a random re-abstraction sample and exchange in between the two observers. Second, since the medical record screening itself was time consuming, no time was left for comparison between the findings on medical record screening and administrative data and subsequently for re-evaluation of the medical record on possibly false positive or negative results. Third, a medical record was judged to be accurate whenever two out of three items – nursing progress notes, documentation procedures and discharge note including medication prescription – were present. However, missing nursing progress notes could have had a significant impact on the finding of an indicator since multiple clinical criteria depended on this notes. Concerning these remarks, Weingart et al[75] already suggested that physician review is at best a 'bronze standard' for evaluating quality.

In order to overcome a few of the mentioned problems, persons who performed the screening were experienced in clinical practice and well-regarded coders. Both team members consulted each other whenever there was doubt about an indicator. Furthermore, difficult files were proposed to a clinical expert panel in order to obtain agreement whenever necessary.

Finally, the eight selected hospitals accounted for just 6.9% of all acute care hospitals in Belgium. Since each acute hospital was invited to take part on this study with an agreement of only 17 hospitals (14.7%), the responding hospitals might be more active in the field of quality of care and therefore serving as a potential hospital selection bias.

# 6 CONCLUSION

Despite general consensus on the need to monitor the quality and outcome of hospital care, questions persist about how to operate in a clinically credible and efficient fashion.

The most critical obstacle in the patient safety campaign remains the lack of system that can reliably identify and report adverse events.

Medical records have so far been the primary source for researching adverse events. This system contains rich clinical details that allow identification of various medical injuries or near misses, and analysis of circumstances and causes of errors. A significant limitation of this system as a broad-based monitoring is its associated costs, access restrictions, privacy and the requirement of a specific clinical contextual knowledge.

Alternative systems for safety research include mandatory and voluntary reports of medical errors, drug safety surveillance, nosocomial infection surveillance, combined modalities and medical malpractice data.

- Incident reporting is the voluntary reporting of a medical event by a health care provider. Studies that have compared the rate of adverse event detection through incident reports versus those detected by chart review have found that only 1.5% of adverse events are detected through them.
- The use of trained observers for the detection of adverse events can be performed prospectively and has been described as the most sensitive detection method for identifying adverse drug events. However, its cost is a limitation of direct observation.
- Combined modalities are methods for detecting adverse events that rely on both electronic and manual review processes. In general, these systems identify an electronically stored 'signal' as screening criteria to identify charts for further review. These systems generally require less reviewer time and thus are less expensive to operate than manual systems. However, as a result of low specificity, much of these approaches still require some form of manual review[24].
- Finally, administrative data might be a viable source, and their potential in patient safety research is increasingly recognized. They are readily available, inexpensive, computer readable, typically continuous, and cover large populations. However, indicators using this source still need to overcome the limitations discussed previously.

Since near misses are difficult to find in the medical record and unlikely to be coded in administrative data, this project focused on five adverse events. We used the definition of Wilson et al[14] of an adverse event 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 disease".

Knowing that there are many concerns regarding ICD-9-CM coding and patient safety research in the international literature, this study aimed at validating the Belgian Hospital Discharge Dataset for a select group of five adverse events.

Our results lead to some comments concerning the usefulness of identifying adverse events through administrative data. First, some adverse events are subject to under reporting in the B-HDDS, e.g. pressure ulcer. On the other hand, some pathologies are over reported. The cause of these errors, whether it be coding error, misinterpretation of the coding rules, maximization of the reimbursements, etc… could not be identified based on the results of this study. Moreover, ICD-9-CM codes sometimes lack sufficient specificity for describing an adverse event leading to false positive results. Finally, the definition, in administrative data such as the B-HDDS, of an accurate algorithm targeting the true adverse event occurring during hospital stay remains an extremely difficult exercise.

We strongly believe that the addition, in the B-HDDS release to be issued in the course of the 2nd semester 2008, of the "Present On Admission" item will allow for a more targeted identification of adverse events. This study also illustrates that some ICD-9-CM diagnostic codes can be confusing as they do not always allow for a precise identification of the event, e.g. the same code is used to identify postoperative sepsis, postoperative haemorrhagic shock and/or postoperative cardiogenic shock. Having clear defined clinical criteria for the chosen adverse event is likewise a critical step in the selection of indicators. Since nearly 60% of positive results on ventilator associated pneumonia were due to the controversial and stringent clinical criteria, this indicator doesn't seem to be valid for the purpose of searching administrative data.

Our study highlights an indicator of adverse event – pressure ulcer – which appears to be relevant to their identification. Indeed, this indicator reveals a positive predictive value excluding POA of 68% suggesting thereby that the event occurred in over 2/3rd of the flagged cases. Moreover, it has been evaluated as the most preventable indicator or the one most likely to be caused by health management.

In conclusion, the B-HDDS can probably detect a select group of adverse events on the condition that the limitations discussed in this study are resolved. Since administrative data provide a very inexpensive and readily accessible source of clinical information, some indicators (such as the pressure ulcer) could be used as a first step, amongst several tools, to review a selection of files towards an approach to improve the quality of care. In this approach, the choice of the indicator of the adverse event depends on :(i) the specificity of the ICD-9-CM code (ii) the perimeter defined by the algorithm used and (iii) the definition of clear clinical criteria without an excessive stringency.

# 7 REFERENCES


1. Kohn LT, Corrigan JM, Donaldson MS. To Err is Human : building a safer health system. Washington, DC Institute of Medicine Report.; 1999.
2. Leape LL, Berwick DM. Five years after To Err Is Human: what have we learned? JAMA. 2005;293(19):2384-90.
3. Longo DR, Hewett JE, Ge B, Schubert S. The long road to patient safety: a status report on patient safety systems. JAMA. 2005;294(22):2858-65.
4. Wachter RM, Shojania KG, Markowitz AJ, Smith M, Saint S. Quality grand rounds: the case for patient safety. Ann Intern Med. 2006;145(8):629-30.
5. Brennan TA, Gawande A, Thomas E, Studdert D. Accidental deaths, saved lives, and improved quality. N Engl J Med. 2005;353(13):1405-9.
6. Amalberti R, Auroy Y, Berwick D, Barach P. Five system barriers to achieving ultrasafe health care. Ann Intern Med. 2005;142(9):756-64.
7. Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. J Gen Intern Med. 2003;18(1):61-7.
8. Michel P, Quenon JL, de Sarasqueta AM, Scemama O. Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. BMJ. 2004;328(7433):199.
9. Lilford RJ, Mohammed MA, Braunholtz D, Hofer TP. The measurement of active errors: methodological issues. Qual Saf Health Care. 2003;12 Suppl 2:ii8-12.
10. DesHarnais SI, Forthman MT, Homa-Lowry JM, Wooster LD. Risk-adjusted clinical quality indicators: indices for measuring and monitoring rates of mortality, complications, and readmissions.[erratum appears in Qual Manag Health Care. 2001 Winter;9(2):ix]. Quality Management in Health Care. 2000;9(1):14-22.
11. Zhan C, Miller MR. Administrative data based patient safety research: a critical review. . Quality & Safety in Health Care. 2003;12 Suppl 2:ii58-ii63.
12. Van Den Heede K, Sermeus W, Diya L, Lesaffre E, Vleugels A. Adverse outcomes in Belgian acute hospitals: Retrospective analysis of the national hospital discharge dataset. International Journal for Quality in Health Care. 2006;18(3):211-9.
13. Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. N Engl J Med. 1991;324(6):370-6.
14. Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L, Hamilton JD. The Quality in Australian Health Care Study. Med J Aust. 1995;163(9):458-71.
15. Gawande AA, Thomas EJ, Zinner MJ, Brennan TA. The incidence and nature of surgical adverse events in Colorado and Utah in 1992. Surgery. 1999;126(1):66-75.
16. Schioler T, Lipczak H, Pedersen BL, Mogensen TS, Bech KB, Stockmarr A, et al. [Incidence of adverse events in hospitals. A retrospective study of medical records]. Ugeskr Laeger. 2001;163(39):5370-8.
17. Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. BMJ. 2001;322(7285):517-9.
18. Davis P, Lay-Yee R, Briant R, Ali W, Scott A, Schug S. Adverse events in New Zealand public hospitals I: occurrence and impact. N Z Med J. 2002;115(1167):U271.
19. Baker GR, Norton PG, Flintoft V, Blais R, Brown A, Cox J, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. CMAJ. 2004;170(11):1678-86.
20. Wagner C, De Bruijne M. Onbedoelde Schade in Nederlandse Ziekenhuizen. Utrecht: EMGO Instituut en NIVEL.

: 2007.

21. McDonald K. Measures of patient safety based on administrative data - The Patient Safety Indicators. prepared for the Agency for Healthcare Research and Quality, contractnumber 290-97-0013: University of California San Francisco - Standford Evidence-based Practice Center; 2002.
22. Best WR, Khuri SF, Phelan M, Hur K, Henderson WG, Demakis JG, et al. Identifying patient preoperative risk factors and postoperative adverse events in administrative databases: results

from the Department of Veterans Affairs National Surgical Quality Improvement Program.[see comment]. Journal of the American College of Surgeons. 2002;194(3):257-66.

23. Fry DE, Pine MB, Jordan HS, Hoaglin DC, Jones B, Meimban R, et al. The hazards of using administrative data to measure surgical quality. American Surgeon. 2006;72(11):1031-7.
24. Murff HJ, Patel VL, Hripcsak G, Bates DW. Detecting adverse events for patient safety research: a review of current methodologies. J Biomed Inform. 2003;36(1-2):131-43.
25. Tamuz M, Thomas EJ, Franchois KE. Defining and classifying medical error: lessons for patient safety reporting systems 404. Quality & Safety in Health Care. 2004;13(1):13-20.
26. Joyce P, Boaden R, Esmail A. Managing risk: a taxonomy of error in health policy. Health Care Anal. 2005;13(4):337-46.
27. Zhang J, Patel VL, Johnson TR, Shortliffe EH. Toward a cognitive taxonomy of medical errors. Proc AMIA Symp. 2002:934-8.
28. Reason J. Human error: models and management. BMJ. 2000;320(7237):768-70.
29. Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, et al. The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. N Engl J Med. 1991;324(6):377-84.
30. Murff HJ, Forster AJ, Peterson JF, Fiskio JM, Heiman HL, Bates DW. Electronically screening discharge summaries for adverse medical events. J Am Med Inform Assoc. 2003;10(4):339-50.
31. Rothschild JM, Landrigan CP, Cronin JW, Kaushal R, Lockley SW, Burdick E, et al. The Critical Care Safety Study: The incidence and nature of adverse events and serious medical errors in intensive care. Crit Care Med. 2005;33(8):1694-700.
32. Kalra J. Medical errors: an introduction to concepts. Clin Biochem. 2004;37(12):1043-51.
33. Handler JA, Gillam M, Sanders AB, Klasco R. Defining, identifying, and measuring error in emergency medicine. Acad Emerg Med. 2000;7(11):1183-8.
34. McNutt RA, Abrams RI. A model of medical error based on a model of disease: interactions between adverse events, failures, and their errors. Qual Manag Health Care. 2002;10(2):23-8.
35. Considine J, Botti M. Who, when and where? Identification of patients at risk of an in-hospital adverse event: implications for nursing practice. Int J Nurs Pract. 2004;10(1):21-31.
36. Kellogg VA, Havens DS. Adverse events in acute care: an integrative literature review. Res Nurs Health. 2003;26(5):398-408.
37. Etchells E, O'Neill C, Bernstein M. Patient safety in surgery: error detection and prevention. World J Surg. 2003;27(8):936-41; discussion 41-2.
38. Grober ED, Bohnen JM. Defining medical error. Can J Surg. 2005;48(1):39-44.
39. Guse CE, Yang H, Layde PM. Identifying risk factors for medical injury. Int J Qual Health Care. 2006;18(3):203-10.
40. Johnstone MJ, Kanitsaki O. The ethics and practical importance of defining, distinguishing and disclosing nursing errors: a discussion paper. Int J Nurs Stud. 2006;43(3):367-76.
41. Thomas EJ, Brennan TA. Incidence and types of preventable adverse events in elderly patients: population based review of medical records. BMJ. 2000;320(7237):741-4.
42. Thomas EJ, Studdert DM, Burstin HR, Orav EJ, Zeena T, Williams EJ, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. Med Care. 2000;38(3):261-71.
43. Donabedian A. Explorations in quality assessment and monitoring. The definition of quality and approaches to its assessment. Chicago: Health Administration Pr,

: 1980.

44. Iezzoni LI. Assessing quality using administrative data. Ann Intern Med. 1997;127(8 Pt 2):666-74.
45. Zhan C, Kelley E, Yang HP, Keyes M, Battles J, Borotkanics RJ, et al. Assessing patient safety in the United States: challenges and opportunities. Med Care. 2005;43(3 Suppl):I42-7.
46. Rubin HR, Pronovost P, Diette GB. From a process of care to a measure: the development and testing of a quality indicator. Int J Qual Health Care. 2001;13(6):489-96.
47. Rubin HR, Pronovost P, Diette GB. The advantages and disadvantages of process-based measures of health care quality. Int J Qual Health Care. 2001;13(6):469-74.
48. Mant J. Process versus outcome indicators in the assessment of quality of health care. Int J Qual Health Care. 2001;13(6):475-80.
49. Reason J. Human Errer. Cambridge University Press; 1990.

50. Chang A, Schyve PM, Croteau RJ, O'Leary DS, Loeb JM. The JCAHO patient safety event taxonomy: a standardized terminology and classification schema for near misses and adverse events. Int J Qual Health Care. 2005;17(2):95-105.
51. National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP). NCC MERP Taxonomy of Medication Errors. In; 1998.
52. 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(3):194-200.
53. Brixey J, Johnson TR, Zhang J. Evaluating a medical error taxonomy. Proc AMIA Symp. 2002:71-5.
54. CHU Liege, SPF Santé Publique Securite de la chaîne alimentaire et Environnement. Feedback des Patient Safety Indicators. La sécurité des patients dans les hôpitaux belges. 2008 Avril 2008.
55. Millar J, Mattke S, and the Members of the OECD Patient Safety Panel. Selecting indicators for patient safety at the health systems level in OECD countries. 2004. OECD Health Technical Papers No.18 Available from: http://www.oecd.org/dataoecd/53/26/33878001.pdf
56. AHRQ Quality Indicators. Guide to Patient Safety Indicators. Rockeville, MD: Agency for Healthcare Research and Quality, Version 3.1 (March 12, 2007). AHRQ Pub.03-R203,

: 2003.

57. Romano PS, Geppert JJ, Davies S, Miller MR, Elixhauser A, McDonald KM. A national profile of patient safety in US hospitals. Health Affairs. 2003;22(2):154-66.
58. Coffey RM, Andrews RM, Moy E, Coffey RM, Andrews RM, Moy E. Racial, ethnic, and socioeconomic disparities in estimates of AHRQ patient safety indicators. Medical Care. 2005;43(3 Suppl):I48-I57.
59. Rosen AK, Rivard P, Zhao S, Loveland S, Tsilimingras D, Christiansen CL, et al. Evaluating the patient safety indicators: how well do they perform on Veterans Health Administration data? Medical Care. 2005;43(9):873-84.
60. Needleman J, Buerhaus P, Mattke S, Stewart M, Zelevinsky K, Needleman J, et al. Nurse-staffing levels and the quality of care in hospitals.[see comment]. New England Journal of Medicine. 2002;346(22):1715-22.
61. Smith MA, Frytak JR, Liou JI, Finch MD, Smith MA, Frytak JR, et al. Rehospitalization and survival for stroke patients in managed care and traditional Medicare plans. Medical Care. 2005;43(9):902-10.
62. Blegen MA, Goode CJ, Reed L. Nurse staffing and patient outcomes. Nurs Res. 1998;47(1):43-50.
63. Lichtig LK, Knauf RA, Milholland DK. Some impacts of nursing on acute care hospital outcomes. J Nurs Adm. 1999;29(2):25-33.
64. Rothschild JM, Bates DW, Leape LL. Preventable medical injuries in older patients. Arch Intern Med. 2000;160(18):2717-28.
65. Iezzoni LI, Foley SM, Heeren T, Daley J, Duncan CC, Fisher ES, et al. A method for screening the quality of hospital care using administrative data: preliminary validation results. Quality Review Bulletin. 1992;18(11):361-71.
66. Arah OA, Klazinga NS. How safe is the safety paradigm? Quality and Safety in Health Care. 2004;13(3):226-32.
67. McCloskey BA, Diers DK, McCloskey BA, Diers DK. Effects of New Zealand's health reengineering on nursing and patient outcomes. Medical Care. 2005;43(11):1140-6.
68. Murff HJ, France DJ, Blackford J, Grogan EL, Yu C, Speroff T, et al. Relationship between patient complaints and surgical complications. Quality and Safety in Health Care. 2006;15(1):13-6.
69. Mattke S, Needleman J, Buerhaus P, Stewart M, Zelevinsky K, Mattke S, et al. Evaluating the role of patient sample definitions for quality indicators sensitive to nurse staffing patterns. Medical Care. 2004;42(2 Suppl):II21-II33.
70. Romano PS, Chan BK, Schembri ME, Rainwater JA, Romano PS, Chan BK, et al. Can administrative data be used to compare postoperative complication rates across hospitals?[see comment]. Medical Care. 2002;40(10):856-67.

71. Geraci JM, Ashton CM, Kuykendall DH, Johnson ML, Wu L. International Classification of Diseases, 9th Revision, Clinical Modification codes in discharge abstracts are poor measures of complication occurrence in medical inpatients. Medical Care. 1997;35(6):589-602.
72. Berlowitz DR, Brand HK, Perkins C. Geriatric syndromes as outcome measures of hospital care: can administrative data be used? Journal of the American Geriatrics Society. 1999;47(6):692-6.
73. White RH, Zhou H, Romano PS. Incidence of symptomatic venous thromboembolism after different elective or urgent surgical procedures. Thrombosis and Haemostasis. 2003;90(3):446-55.
74. McCarthy EP, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, et al. Does clinical evidence support ICD-9-CM diagnosis coding of complications? Medical Care. 2000;38(8):868-76.
75. Weingart SN, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, et al. Use of administrative data to find substandard care: validation of the complications screening program. Medical Care. 2000;38(8):796-806.
76. Lawthers AG, McCarthy EP, Davis RB, Peterson LE, Palmer RH, Iezzoni LI. Identification of in-hospital complications from claims data. Is it valid? Medical Care. 2000;38(8):785-95.
77. Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, et al. Does the Complications Screening Program flag cases with process of care problems? Using explicit criteria to judge processes. Int.J Qual Health Care. 1999;11(2):107-18.
78. Kovner C, Jones C, Zhan C, Gergen PJ, Basu J, Kovner C, et al. Nurse staffing and postsurgical adverse events: an analysis of administrative data from a sample of U.S. hospitals, 1990-1996. Health Services Research. 2002;37(3):611-29.
79. Kearon C, Kearon C. Epidemiology of postoperative venous thromboembolism: lessons from an administrative data base.[comment]. Thrombosis & Haemostasis. 2003;90(3):367-8.
80. White RH, Brickner LA, Scannell KA, White RH, Brickner LA, Scannell KA. ICD-9-CM codes poorly indentified venous thromboembolism during pregnancy. Journal of Clinical Epidemiology. 2004;57(9):985-8.
81. Arnason T, Wells PS, van Walraven C, Forster AJ. Accuracy of coding for possible warfarin complications in hospital discharge abstracts. Thrombosis Research. 2006;118(2):253-62.
82. Sermeus W, Van den Heede K, Vleugels A. Database Quality of Nursing Care: scientific report. . Leuven: Federaal Wetenschapsbeleid; 2007.
83. Gastmeier P, Geffers C. Prevention of ventilator-associated pneumonia: analysis of studies published since 2004. J Hosp Infect. 2007;67(1):1-8.
84. National Quality Forum. National Voluntary Consensus Standards for Nursing-Sensitive Care: An Initial Performance Measure Set. http://www.qualityforum.org (Last accessed on November 2005 ). In.
85. Collard HR, Saint S. Prevention of ventilator-associated pneumonia. In : AHRQ Evidence Report/Technical Assessment No 43 Making Health Safer,.
86. Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH, Aiken LH, et al. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction.[see comment]. JAMA : the journal of the American Medical Association. 2002;288(16):1987-93.
87. Quan H, Parsons GA, Ghali WA, Quan H, Parsons GA, Ghali WA. Assessing accuracy of diagnosis-type indicators for flagging complications in administrative data. Journal of Clinical Epidemiology. 2004;57(4):366-72.
88. Vlayen J, Van De Water G, Camberlin C, Paulus D, Leys M, Ramaekers D, et al. Indicateurs de qualité cliniques. Objective Elements - Communication (OEC). Bruxelles: Centre fédéral d'expertise des soins de santé (KCE); 2006. KCE reports 41B (D/2006/10.273/44).
89. Terryn N, Aelvoet W, Windey F, Loosen F, Mertens I, Cellule Audit RCM SD, et Cellule Statistique, Service Datamanagement. ANALYSE DE LA QUALITE DES DONNÉES RCM Basée sur le financement de 2004 à 2007. December 2007. Available from: https://portal.health.fgov.be/pls/portal/docs/PAGE/INTERNET_PG/HOMEPAGE_MENU/GEZONDHEIDZORG1_MENU/ZORGINSTELLINGEN1_MENU/REGISTRATIESYSTEMEN1_MENU/MKGMINIMALEKLINISCHEGEGEVENS1_MENU/PUBLICATIES17_HIDE/PUBLICATIES17_DOCS/20080114%20ANALYSE%20DE%20LA%20QUALIT%C3%89%20DES%20DONN%C3%89ES%20RCM.PDF

90. Klompas M, Platt R. Ventilator-associated pneumonia--the wrong quality measure for benchmarking. Ann Intern Med. 2007;147(11):803-5.

91. Iezzoni LI, Daley J, Heeren T, Foley SM, Hughes JS, Fisher ES, et al. Using administrative data to screen hospitals for high complication rates. Inquiry. 1994;31(1):40-55.

92. Naessens JM, Scott CG, Huschka TR, Schutt DC. Do complication screening programs detect complications present at admission? Joint Commission journal on quality and safety. 2004;30(3):133-42.

93. Glance LG, Dick AW, Osler TM, Mukamel DB, Glance LG, Dick AW, et al. Accuracy of hospital report cards based on administrative data. Health Services Research. 2006;41(4 Pt 1):1413-37.

94. Glance LG, Dick AW, Osler TM, Mukamel DB, Glance LG, Dick AW, et al. Does date stamping ICD-9-CM codes increase the value of clinical information in administrative data? Health Services Research. 2006;41(1):231-51.

95. Glance LG, Osler TM, Mukamel DB, Dick AW. Impact of the present-on-admission indicator on hospital quality measurement: experience with the Agency for Healthcare Research and Quality (AHRQ) Inpatient Quality Indicators. Med Care. 2008;46(2):112-9.

96. Moro ML, Morsillo F. Can hospital discharge diagnoses be used for surveillance of surgical-site infections? Journal of Hospital Infection. 2004;56(3):239-41.

*This page is left intentionally blank.*

Wettelijk depot : D/2008/10.273/73

## KCE reports

1. Effectiviteit en kosten-effectiviteit van behandelingen voor rookstop. D/2004/10.273/1.
2. Studie naar de mogelijke kosten van een eventuele wijziging van de rechtsregels inzake medische aansprakelijkheid (fase 1). D/2004/10.273/2.
3. Antibioticagebruik in ziekenhuizen bij acute pyelonefritis. D/2004/10.273/5.
4. Leukoreductie. Een mogelijke maatregel in het kader van een nationaal beleid voor bloedtransfusieveiligheid. D/2004/10.273/7.
5. Het preoperatief onderzoek. D/2004/10.273/9.
6. Validatie van het rapport van de Onderzoekscommissie over de onderfinanciering van de ziekenhuizen. D/2004/10.273/11.
7. Nationale richtlijn prenatale zorg. Een basis voor een klinisch pad voor de opvolging van zwangerschappen. D/2004/10.273/13.
8. Financieringssystemen van ziekenhuisgeneesmiddelen: een beschrijvende studie van een aantal Europese landen en Canada. D/2004/10.273/15.
9. Feedback: onderzoek naar de impact en barrières bij implementatie – Onderzoeksrapport: deel 1. D/2005/10.273/01.
10. De kost van tandprothesen. D/2005/10.273/03.
11. Borstkankerscreening. D/2005/10.273/05.
12. Studie naar een alternatieve financiering van bloed en labiele bloedderivaten in de ziekenhuizen. D/2005/10.273/07.
13. Endovasculaire behandeling van Carotisstenose. D/2005/10.273/09.
14. Variaties in de ziekenhuispraktijk bij acuut myocardinfarct in België. D/2005/10.273/11.
15. Evolutie van de uitgaven voor gezondheidszorg. D/2005/10.273/13.
16. Studie naar de mogelijke kosten van een eventuele wijziging van de rechtsregels inzake medische aansprakelijkheid. Fase II : ontwikkeling van een actuarieel model en eerste schattingen. D/2005/10.273/15.
17. Evaluatie van de referentiebedragen. D/2005/10.273/17.
18. Prospectief bepalen van de honoraria van ziekenhuisartsen op basis van klinische paden en guidelines: makkelijker gezegd dan gedaan.. D/2005/10.273/19.
19. Evaluatie van forfaitaire persoonlijk bijdrage op het gebruik van spoedgevallendienst. D/2005/10.273/21.
20. HTA Moleculaire Diagnostiek in België. D/2005/10.273/23, D/2005/10.273/25.
21. HTA Stomamateriaal in België. D/2005/10.273/27.
22. HTA Positronen Emissie Tomografie in België. D/2005/10.273/29.
23. HTA De electieve endovasculaire behandeling van het abdominale aorta aneurysma (AAA). D/2005/10.273/32.
24. Het gebruik van natriuretische peptides in de diagnostische aanpak van patiënten met vermoeden van hartfalen. D/2005/10.273/34.
25. Capsule endoscopie. D/2006/10.273/01.
26. Medico–legale aspecten van klinische praktijkrichtlijnen. D2006/10.273/05.
27. De kwaliteit en de organisatie van type 2 diabeteszorg. D2006/10.273/07.
28. Voorlopige richtlijnen voor farmaco-economisch onderzoek in België. D2006/10.273/10.
29. Nationale Richtlijnen College voor Oncologie: A. algemeen kader oncologisch kwaliteitshandboek B. wetenschappelijke basis voor klinische paden voor diagnose en behandeling colorectale kanker en testiskanker. D2006/10.273/12.
30. Inventaris van databanken gezondheidszorg. D2006/10.273/14.
31. Health Technology Assessment prostate-specific-antigen (PSA) voor prostaatkankerscreening. D2006/10.273/17.
32. Feedback : onderzoek naar de impact en barrières bij implementatie – Onderzoeksrapport : deel II. D/2006/10.273/19.
33. Effecten en kosten van de vaccinatie van Belgische kinderen met geconjugeerd pneumokokkenvaccin. D/2006/10.273/21.
34. Trastuzumab bij vroegtijdige stadia van borstkanker. D/2006/10.273/23.
35. Studie naar de mogelijke kosten van een eventuele wijziging van de rechtsregels inzake medische aansprakelijkheid (fase III)- precisering van de kostenraming. D/2006/10.273/26.
36. Farmacologische en chirurgische behandeling van obesitas. Residentiële zorg voor ernstig obese kinderen in België. D/2006/10.273/28.
37. HTA Magnetische Resonantie Beeldvorming. D/2006/10.273/32.

38. Baarmoederhalskankerscreening en testen op Human Papillomavirus (HPV). D/2006/10.273/35
39. Rapid assessment van nieuwe wervelzuil technologieën : totale discusprothese en vertebro/ballon kyfoplastie. D/2006/10.273/38.
40. Functioneel bilan van de patiënt als mogelijke basis voor nomenclatuur van kinesitherapie in België? D/2006/10.273/40.
41. Klinische kwaliteitsindicatoren. D/2006/10.273/43.
42. Studie naar praktijkverschillen bij electieve chirurgische ingrepen in België. D/2006/10.273/45.
43. Herziening bestaande praktijkrichtlijnen. D/2006/10.273/48.
44. Een procedure voor de beoordeling van nieuwe medische hulpmiddelen. D/2006/10.273/50.
45. HTA Colorectale Kankerscreening: wetenschappelijke stand van zaken en budgetimpact voor België. D/2006/10.273/53.
46. Health Technology Assessment. Polysomnografie en thuismonitoring van zuigelingen voor de preventie van wiegendood. D/2006/10.273/59.
47. Geneesmiddelengebruik in de belgische rusthuizen en rust- en verzorgingstehuizen. D/2006/10.273/61
48. Chronische lage rugpijn. D/2006/10.273/63.
49. Antivirale middelen bij seizoensgriep en grieppandemie. Literatuurstudie en ontwikkeling van praktijkrichtlijnen. D/2006/10.273/65.
50. Eigen betalingen in de Belgische gezondheidszorg. De impact van supplementen. D/2006/10.273/68.
51. Chronische zorgbehoeften bij personen met een niet- aangeboren hersenletsel (NAH) tussen 18 en 65 jaar. D/2007/10.273/01.
52. Rapid Assessment: Cardiovasculaire Primaire Preventie in de Belgische Huisartspraktijk. D/2007/10.273/03.
53. Financiering van verpleegkundige zorg in ziekenhuizen. D/2007/10 273/06
54. Kosten-effectiviteitsanalyse van rotavirus vaccinatie van zuigelingen in België
55. Evidence-based inhoud van geschreven informatie vanuit de farmaceutische industrie aan huisartsen. D/2007/10.273/12.
56. Orthopedisch Materiaal in België: Health Technology Assessment. D/2007/10.273/14.
57. Organisatie en Financiering van Musculoskeletale en Neurologische Revalidatie in België. D/2007/10.273/18.
58. De Implanteerbare Defibrillator: een Health Technology Assessment. D/2007/10.273/21.
59. Laboratoriumtesten in de huisartsgeneeskunde. D2007/10.273/24.
60. Longfunctie testen bij volwassenen. D/2007/10.273/27.
61. Vacuümgeassisteerde Wondbehandeling: een Rapid Assessment. D/2007/10.273/30
62. Intensiteitsgemoduleerde Radiotherapie (IMRT). D/2007/10.273/32.
63. Wetenschappelijke ondersteuning van het College voor Oncologie: een nationale praktijkrichtlijn voor de aanpak van borstkanker. D/2007/10.273/35.
64. HPV Vaccinatie ter Preventie van Baarmoederhalskanker in België: Health Technology Assessment. D/2007/10.273/41.
65. Organisatie en financiering van genetische diagnostiek in België. D/2007/10.273/44.
66. Health Technology Assessment: Drug-Eluting Stents in België. D/2007/10.273/47
67. Hadrontherapie. D/2007/10.273/50.
68. Vergoeding van schade als gevolg van gezondheidszorg – Fase IV : Verdeelsleutel tussen het Fonds en de verzekeraars. D/2007/10.273/52.
69. Kwaliteit van rectale kankerzorg – Fase 1: een praktijkrichtlijn voor rectale kanker D/2007/10.273/54.
70. Vergelijkende studie van ziekenhuisaccrediterings-programma's in Europa D/2008/10.273/57.
71. Aanbevelingen voor het gebruik van vijf oftalmologische testen in de klinische praktijk .D/2008/10.273/04
72. Het aanbod van artsen in België. Huidige toestand en toekomstige uitdagingen. D/2008/10.273/07
73. Financiering van het zorgprogramma voor de geriatrische patiënt in algemene ziekenhuizen: definitie en evaluatie van een geriatrische patiënt, definitie van de interne liaisongeriatrie en evaluatie van de middelen voor een goede financiering. D/2008/10.273/11
74. Hyperbare Zuurstoftherapie: Rapid Assessment. D/2008/10.273/13.
75. Wetenschappelijke ondersteuning van het College voor Oncologie: een nationale praktijkrichtlijn voor de aanpak van slokdarm- en maagkanker. D/2008/10.273/16.

76. Kwaliteitsbevordering in de huisartsenpraktijk in België: status quo of quo vadis? D/2008/10.273/18.
77. Orthodontie bij kinderen en adolescenten. D/2008/10.273/20.
78. Richtlijnen voor farmaco-economische evaluaties in België. D/2008/10.273/23.
79. Terugbetaling van radioisotopen in België. D/2008/10.273/26
80. Evaluatie van de effecten van de maximumfactuur op de consumptie en financiële toegankelijkheid van gezondheidszorg. D/2008/10.273/35.
81. Kwaliteit van rectale kankerzorg – phase 2: ontwikkeling en test van een set van kwaliteitsindicatoren. D/2008/10.273/38
82. 64-Slice computertomografie van de kransslagaders bij patiënten met vermoeden van coronaire hartziekte. D/2008/10.273/40
83. Internationale vergelijking van terugbetalingsregels en juridische aspecten van plastische heelkunde. D/200810.273/43
84. Langverblijvende psychiatrische patiënten in T-bedden. D/2008/10.273/46
85. Vergelijking van twee financieringssystemen voor de eerstelijnszorg in België. D/2008/10.273/49.
86. Functiedifferentiatie in de verpleegkundige zorg: mogelijkheden en beperkingen. D/2008/10.273/52.
87. Het gebruik van kinesitherapie en van fysische geneeskunde en revalidatie in België. D/2008/10.273/54.
88. Chronisch Vermoeidheidssyndroom: diagnose, behandeling en zorgorganisatie. D/2008/10.273/58.
89. Rapid assessment van enkele nieuwe behandelingen voor prostaatkanker en goedaardige prostaathypertrofie. D/2008/10.273/61
90. Huisartsgeneeskunde: aantrekkingskracht en beroepstrouw bevorderen. D/2008/10.273/63
91. Hoorapparaten in België: health technology assessment. D/2008/10.273/67
92. Nosocomiale infecties in België, deel 1: nationale prevalentiestudie. D/2008/10.273/70.
93. Detectie van adverse events in administratieve databanken. D/2008/10.273/73.