

Unclassified

DELSA/HEA/WD/HWP(2009)5

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

23-Nov-2009

English text only

DIRECTORATE FOR EMPLOYMENT, LABOUR AND SOCIAL AFFAIRS
HEALTH COMMITTEE

Health Working Papers

OECD HEALTH WORKING PAPERS No. 47

HEALTH CARE QUALITY INDICATORS PROJECT: PATIENT SAFETY INDICATORS
REPORT 2009

Saskia Drösler, Patrick Romano, Lihan Wei

JEL Classifications: I19

Cross-country comparisons on patient safety are cautioned given that data presented in this working paper require further research and development.

This working paper has a separate annex with the code DELSA/HEA/WD/HWP(2009)5/ANN.

JT03274823

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format



DELSA/HEA/WD/HWP(2009)5
Unclassified

English text only

DIRECTORATE FOR EMPLOYMENT, LABOUR AND SOCIAL AFFAIRS

www.oecd.org/els

OECD HEALTH WORKING PAPERS

<http://www.oecd.org/els/health/workingpapers>

This series is designed to make available to a wider readership health studies prepared for use within the OECD. Authorship is usually collective, but principal writers are named. The papers are generally available only in their original language – English or French – with a summary in the other.

Comment on the series is welcome, and should be sent to the Directorate for Employment, Labour and Social Affairs, 2, rue André-Pascal, 75775 PARIS CEDEX 16, France.

The opinions expressed and arguments employed here are the responsibility of the author(s) and do not necessarily reflect those of the OECD.

**Applications for permission to reproduce or translate
all or part of this material should be made to:**

**Head of Publications Service
OECD
2, rue André-Pascal
75775 Paris, CEDEX 16
France**

Copyright OECD 2009

ACKNOWLEDGEMENTS

This paper presents the work of several years of research and development work by the HCQI Expert Group, particularly the members of the Patient Safety Subgroup and the data experts contributing data from the following countries: Australia, Belgium, Canada, Denmark, Finland, France, Germany, Iceland, Italy, Ireland, Latvia, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, the United Kingdom and the United States. The authors would like to thank Daniel Tancredi (University of California, Davis) for statistical analysis, Mark Pearson, Gaetan Lafortune, Ian Brownwood and Niek Klazinga for their contributions and Isabelle Vallard for editorial assistance.

The HCQI Project receives financial support from a grant provided by the Directorate General for Public Health and Consumer Affairs of the European Commission.

SUMMARY

This paper reports on the progress in the research and development of the set of patient safety indicators developed by the Health Care Quality Indicators project. The indicators presented here have been recommended by an expert group for further consideration in international reporting on the quality of care on the key dimension of safety. The indicators have been selected by expert consensus, undergone validity testing and have been tested for comparability. While concern remains related to differences in coding and reporting from administrative hospital databases, the rigour with which the indicator work has been undertaken has resulted in the improved ability of countries to report on the quality of care. The work on the development of the patient safety indicators highlights the technical progress made in constructing measures and the ongoing need for methodological improvements. The indicators reported here should not be considered as making inferences on the state of patient safety in countries, but are intended to raise questions towards improving understanding of the reported differences.

RÉSUMÉ

Ce document présente l'état d'avancement de la recherche et du développement d'un ensemble d'indicateurs en matière de sécurité des patients dans le cadre du projet sur les indicateurs de la qualité des soins (HCQI). Un groupe d'experts a recommandé l'utilisation des indicateurs présentés ici pour les comparaisons internationales sur une dimension clé de la qualité des soins : la sécurité. Les indicateurs ont été sélectionnés par un consensus d'experts, leur validité et leur comparabilité ont été testées. Bien qu'il reste quelques problèmes quant aux différences de codage et de déclaration venant des bases de données administratives hospitalières, la rigueur du travail sur les indicateurs a permis d'améliorer la capacité des pays à rendre compte de la qualité des soins. Le développement des indicateurs de la sécurité des patients met l'accent sur les progrès techniques réalisés dans la construction de mesures et le besoin récurrent d'améliorer la méthodologie. Les indicateurs présentés ici ne doivent pas donner lieu à des conclusions quant à la situation de la sécurité des patients dans les pays, mais visent plutôt à poser des questions pour une meilleure compréhension des différences observées.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	3
SUMMARY.....	4
RÉSUMÉ.....	4
INTRODUCTION.....	6
Background.....	6
Participating countries.....	7
Methodology.....	8
Discussion of the recommended indicators.....	13
Catheter-related bloodstream infection.....	13
Postoperative pulmonary embolism (PE) or deep vein thrombosis (DVT).....	17
Postoperative sepsis.....	21
Accidental puncture or laceration.....	24
Foreign body left in during procedure.....	28
Obstetric trauma – vaginal delivery with instrument.....	31
Obstetric trauma – vaginal delivery without instrument.....	31
Detailed results.....	34
Administrative reporting.....	35
Year to year reliability.....	37
Conclusions.....	38
REFERENCES.....	40
Boxes	
Box 1. History of the patient safety indicators.....	6

INTRODUCTION

Background

1. The OECD HCQI project identified patient safety as one of the priority areas for internationally comparable quality indicators. Government officials, national experts, academic researchers from participating countries, and the Secretariat have collaborated in the selection, data collection and validation of the recommended indicators. An extended discussion is presented here outlining the research and development for the seven proposed and agreed-upon indicators.

Box 1. History of the patient safety indicators project

In 2004, the OECD Health Care Quality Indicators Project identified patient safety as one of the five core priority areas for the development of quality indicators (along with cardiac care, primary care, mental health care and diabetes care). Using a structured review process, including a comprehensive literature review, clinician panel review, risk adjustment and empirical analysis, an expert panel selected, evaluated and recommended a set of patient safety indicators for comprehensiveness and cohesiveness across the five domains of patient safety: hospital-acquired infections; operative and postoperative complications; sentinel events¹; obstetrics; and other care-related adverse events. From a total of 59 identified possible indicators, a group of 21 indicators was agreed upon based on their importance and scientific soundness (Millar *et al*, 2004).

In 2006, the first patient safety subgroup expert meeting met to discuss and evaluate the current state of patient safety data systems in OECD countries and potential areas of improvement in order to obtain reliable data for the proposed measures. Adverse events that occur rarely were identified as being somewhat outside the standardised data collection systems. National systems of coding patient safety were identified in OECD countries, although differing versions and adaptations employed in hospital administrative databases were recognised as a key area of development for international data comparability. One of the main outcomes of the meeting was the decision to develop an ICD crosswalk across versions of coding for international harmonisation of data comparability.

A comprehensive technical manual was developed to support the calculation of the 15 patient safety indicators to be collected in national hospital administrative databases. For each of the indicators, the manual provided detailed definitions of the indicators, coding and calculation processes, and a crosswalk from ICD 9 to ICD 10 coding (Drösler, 2008).

¹ While all patient safety indicators refer to events that should occur rarely, sentinel events (including foreign body left in during procedure, transfusion reaction, wrong site surgery, etc.) are those that in theory and practice should never happen and thus whose occurrence should stand as a signal for immediate investigation.

Three years of data collection, indicator development and methodology testing (2007-2009) have included a total participation from 21 countries (Table 1), with monthly teleconference meetings to refine definitions and methodology and yearly meetings of the patient safety expert subgroup to evaluate the validity of the methodology, data reliability and development of the data set. The 2009 data collection included a *revised* technical manual for the calculation of the indicators reflecting refinements and improvements in the methodology (see the Annex).

The long-term objective of the Health Care Quality Indicators project remains to develop a set of internationally comparable indicators that can then be used to raise questions for further exploration of the underlying reasons as to why these differences exist across and within countries. The patient safety indicators (as part of the larger set of quality indicators) represent measures selected for their importance - affect on health, policy relevance and susceptibility to influence by the health care system - and scientific soundness - face validity, content validity and reliability (Mattke *et al.*, 2006).

2. Patient Safety Indicators (PSIs) originally published by the US Agency for Healthcare Research and Quality (AHRQ) have been evaluated to explore the potential of international comparison for patient safety and public reporting. Preliminary pilot studies among seven OECD member countries in 2007 (Drösler *et al.*, 2009), and sixteen countries in 2008 showed the feasibility of the method. Countries were able to calculate most of the PSIs using their administrative hospital databases. Pilot results demonstrated that relative PSI rates appeared to be highly correlated across countries, although some countries had much lower rates for all PSIs than others.

3. The objectives of the PSI development work are:

- 1) To investigate the validity of the PSI rates of participating countries.
- 2) To evaluate the potential impacts on country-specific PSI rates of variation in the distribution of age and gender, length of hospital stay and diagnoses coding practices.

Participating countries

4. During the first three years of indicator development, the following countries (Table 1) participated in the data collection, validity testing and evaluation of the patient safety indicators.

Table 1. Countries participating in the development of the patient safety indicators

	2007	2008	2009
Australia	X		
Belgium		X	X
Canada	X	X	X
Denmark		X	X
Finland		X	X
France		X	X
Germany	X	X	X
Iceland			X
Italy		X	X
Ireland			X
Latvia			X
Netherlands		X	
New Zealand		X	X
Norway		X	X
Portugal		X	X
Singapore		X	X
Spain	X	X	X
Sweden	X	X	X
Switzerland			X
United Kingdom	X	X	X
United States	X	X	X

Methodology

5. A set of 15 patient safety indicators was initially selected for development in 2007. Of these, seven were determined to be ready for data collection and analysis, based on the decision of the OECD's Patient Safety Expert Subgroup in October 2008 (Table 2). The other eight indicators either rely on procedure codes that proved difficult to map across countries (e.g., postoperative respiratory failure, iatrogenic pneumothorax), or they were found to have limited validity in North American studies that distinguished conditions present on admission from hospital-acquired complications (e.g., complications of anaesthesia, decubitus ulcer, postoperative hip fracture).

Table 2: List of patient safety indicators evaluated

Area	Indicator name	AHRQ PSI	Recommended for ongoing evaluation
Hospital-acquired infections	Catheter-related bloodstream infection	PSI 7	X
	Decubitus ulcer	PSI 3	
Operative and post-operative complications	Complications of anaesthesia	PSI 1	
	Postoperative hip fracture	PSI 8	
	Postoperative pulmonary embolism (PE) or deep vein thrombosis (DVT)	PSI 12	X
	Postoperative sepsis	PSI 13	X
	Accidental puncture or laceration	PSI 15	X
	Postoperative respiratory failure	PSI 11	
	Iatrogenic pneumothorax	PSI 6	
Sentinel events	Transfusion reaction	PSI 16	
	Foreign body left in during procedure	PSI 5	X
Obstetrics	Birth trauma - injury to neonate	PSI 17	
	Obstetric trauma – vaginal delivery with instrument	PSI 18	X
	Obstetric trauma – vaginal delivery without instrument	PSI 19	X
	Obstetric trauma - caesarean section	PSI 20	

6. An OECD calculation manual was provided (Annex), which adopted the PSI definitions as published by AHRQ (AHRQ, 2007) for most of the indicators. In addition, the manual provided internationally harmonised code lists in ICD-10 (Quan, 2008) as more than half of the participating countries use ICD-10 (Table 3). Countries were asked to complete an accompanying questionnaire to provide information on the databases and related data collection procedures.

7. Patient safety indicators are largely constructed from secondary diagnoses. The calculation of the numerator builds on secondary diagnosis for most of the indicators. The indicator definitions refer to conditions produced post admission.

8. Nineteen countries performed PSI calculations and delivered their results between February and June 2009. Table 3 shows some key country-related information derived from the questionnaires. Most of the participating countries use a hospital reimbursement system based on Diagnosis Related Groups (DRGs). This information is important as DRGs are based on coded health information. The use of coded health data for reimbursement creates incentives for accurate and complete documentation. However, although many countries use a DRG-based reimbursement system, the mean number of secondary diagnoses among eligible cases varied substantially between 0.87 (Italy, postoperative sepsis – PSI 15) and 7.02 (Belgium, accidental puncture or laceration – PSI 13). Finland reported relatively few secondary diagnoses (0.47) in the 2008 PSI calculation pilot but did not provide any information on the number of secondary diagnoses in 2009. Iceland and Latvia provided only obstetric indicators from registry data, so the number of secondary diagnoses is not relevant for these countries.

Table 3. Country-related key information on medical classifications, documentation and reimbursement

Country	Year	Diagnosis classification	Procedure classification	Mean number of secondary diagnoses, derived from PSI 5	DRG reimbursement
Belgium	2006	ICD-9 CM	ICD-9 CM	6.72	Yes
Canada	2007	ICD-10 CA	CCI	3.11	Yes
Denmark	2008	ICD-10	NOMESCO	2.41	Yes
Finland	2007	ICD-10	NOMESCO	NA	Yes
France	2007	ICD-10	CCAM	NA	Yes
Germany	2007	ICD-10 GM (2007)	OPS 2007	5.31	Yes
Iceland	2007	ICD-10	NOMESCO	NA	Yes
Ireland	2007	ICD-10 AM 4th Edition	ACHI	2.57	Yes
Italy	2007	ICD-9 CM (1997)	ICD-9 CM	1.50	Yes
Latvia	2007	NA	NA	NA	NA
New Zealand	2007	ICD-10 AM	ACHI	3.93	Yes
Norway	2007	ICD-10	NOMESCO	1.90	Yes
Portugal	2007	ICD-9 CM	ICD-9 CM	NA	Yes
Singapore	2007	ICD-9-AM	ICD-9-AM	3.89	Yes
Spain	2007	ICD-9 CM V22 (2004)	ICD-9 CM	3.71	No ¹
Sweden ²	2007	NA	NA	2.50	NA
Switzerland	2007	ICD-10 WHO 1.3	CHOP (ICD-9-CM)	3.03	Yes (partially)
United Kingdom	2007	ICD-10	OPCS 4.3	2.72	Yes
United States	2006	ICD-9 CM	ICD-9 CM	6.02	Yes

1. Reported that a DRG hospital payment system was in use for 2008 pilot.

2. From earlier data collections in 2007 and 2008, Sweden uses ICD-10, NOMESCO and a DRG system in some parts of the country.

9. The number of cases included from each participating country can be approximated by the denominator of PSI 5 (foreign body left in during procedure). The eligible population for this indicator includes nearly all adult medical and surgical cases (Table 4). A comparison to the hospital discharge numbers reported by the OECD (2009) suggests that most countries used their complete administrative database for PSI calculation with the observed differences largely due to the exclusion of paediatric and obstetric patients from the denominator of PSI 5. Three countries (*e.g.* Germany, USA, and Sweden) used a sample for calculating their PSI rates. Only Denmark reported a slightly higher count in the PSI denominator than the total number of hospital discharges, presumably because these data come from 2008 and 2007, respectively.

Table 4. Country-related information on discharge numbers and average length of stay

Country	Denominator population PSI 5 (in millions)	Estimated annual hospital discharges (in millions, 2007)	Average length of stay (denominator PSI 5)	Average length of stay (2007)
Belgium	1.52	1.83 ¹	8.55	7.2 ¹
Canada	1.52	2.75 ¹	8.14	7.3
Denmark	1.00	0.93	4.49	3.5 ²
Finland	0.94	1.00	3.60	4.6
Germany	1.27 ³	18.67	8.76	7.8
Ireland	0.46	0.60	7.30	5.9 ¹
Italy	7.25	8.11 ¹	8.34	6.7 ¹
New Zealand	0.43	0.59	5.51	NA
Norway	0.62	0.82	5.30	5.0
Portugal	NA	1.15	NA	6.8
Singapore	0.24	0.43 ⁵	5.91	NA
Spain	3.29	4.78	7.73	6.6 ¹
Sweden	1.26 ³	1.51	5.05	4.5
Switzerland	0.93	1.26	8.25	7.8
United Kingdom	6.68	7.32	NA	7.2
United States	32.81 ⁴	37.74 ²	4.79	5.5 ⁶

1. 2006.

2. 2005.

3. PSI calculations are based on a 10% representative sample of the entire population of hospital discharges.

4. PSI calculations are based on a 20% representative sample of US hospital discharges, extrapolated to the total population.

5. www.moh.gov.sg

6. When the same data set (Nationwide Inpatient Sample) from which PSI rates were derived was also used to estimate mean length of stay in the USA, the value was 4.6 days.

Source: OECD Health Data 2009 (for estimated annual hospital discharges and average length of stay)

10. A comparison of mean length of stay between the PSI calculation data and annually reported OECD data on health care activities shows concordance for most countries. It is expected that the average length of stay is longer among the denominator cases of PSI 5 than the corresponding average extracted from OECD Health Data, as children are excluded from the denominator population of PSI 5. Only Finland and USA report a shorter average length of stay in their PSI 5 data than the data they have submitted for OECD Health Data. This discrepancy for the USA is likely due to the use of a different data source for annual reporting to OECD on average length of stay (*i.e.* the American Hospital Associations voluntary annual survey of hospitals).

Validity

11. Applying general concepts of validation to the field of patient safety measurement, several potential domains of validity can be identified. Some of these domains are easier to evaluate than others; unfortunately, the most challenging domains to evaluate are often the most useful for those who wish to understand the meaning of the data.

12. Content validity addresses the extent to which the content of a measure is consistent with professional knowledge about health care quality and the outcomes of high-quality care. Consensual validation is the most rigorous approach for assessing the content validity of health care quality indicators, because it requires agreement or near-consensus among professionals from different disciplines, different regions, and different practice environments. Ideally, the expert panels convened for consensual validation represent all of the disciplines involved in treating the condition(s) of interest, include at least 8-10 members, and discuss all of the relevant evidence supporting use of the quality indicator.

13. Construct validity addresses the extent to which one purported measure of quality is correlated with other measures with which a high correlation would be expected, according to the conceptual framework underlying quality improvement research. The most common application of this approach, known as convergent validity, is to estimate correlations between measures of the process of care and measures of the outcomes of that care. Process measures include both implicit assessments, in which health professionals review available documents or other evidence to formulate a global assessment of quality, and explicit assessments, which focus on specific evidence-based diagnostic tests or treatments. Explicit process measures are typically preferred, because they are often (but not always) based on randomised controlled trials, which are relatively immune to bias from unmeasured confounders. Another approach to construct validation is to study associations between outcome measures and structural indicators, such as nurse staffing levels and skill mix, which have previously been shown to represent markers of quality. Finally, some authors test the construct that any meaningful adverse outcome should be associated with other adverse outcomes. Applied to patient safety measurement, this construct suggests that in-hospital adverse events should be associated with subsequent mortality, re-admissions, prolonged length of stay, and long-term disability.

14. Criterion validity addresses the extent to which one purported measure of quality is correlated with other, better measures of the same phenomenon. It implies the existence of a “gold standard” that can be used to evaluate less costly – and presumably less accurate – measurement methods. Applied to patient safety measurement, this approach typically involves comparing indicators based on routinely collected administrative data with indicators of the same outcomes based on more complex linked data, in-depth medical record review, physician/nurse interview, patient interview, or even direct observation. Criterion validity may represent the strongest validation approach, but its applicability is often limited by the lack of an accepted “gold standard”.

Discussion of the recommended indicators

15. Charts in the following sections do not contain all data displayed in Table 6 (detailed results section below) due to validity concerns. Analysis of data quality revealed documentation patterns with the data that may indicate that some rates from some countries may not have been appropriate for PSI calculation. Even for rare events, zero rates are considered implausible for patient safety events (except sentinel events).

Catheter-related bloodstream infection

Validity

16. *Rationale and concerns.* As nosocomial infections are often preventable (Eggimann *et al.*, 2000), the occurrence of infections in the course of medical care is an important measure of the quality of care. The indicator catheter-related bloodstream infection is supposed to flag cases with a hospital-acquired infection caused by intravenous lines or catheters. Infections related to medical care can be a very serious problem, leading to death in some cases. Often patients experience pain and others discomfort. Currently, WHO Patient Safety promotes the campaign “Clean Care is Safer Care” as it is proven that proper hand hygiene reduces the incidence of catheter-related bloodstream infections (WHO, 2009).

17. *Content (consensual) validity.* Content validity was addressed in the technical report accompanying the original release of the AHRQ PSIs (AHRQ, 2007). Although panellists’ ratings of the “usefulness” of each candidate indicator were used to select the final PSI set, panellists were also asked to rate each indicator on its preventability and its likelihood of being due to medical error. The median rating of this indicator was 7, with indeterminate agreement, on the former dimension, leading to a classification of “acceptable”. The median rating was 6, with indeterminate agreement, on the latter dimension, leading to a classification of “unclear”. Through similar processes, this PSI was endorsed by the OECD Patient Safety Panel (Millar *et al.*, 2004), but rejected by the SimPatIE (Safety Improvement for Patients in Europe) project as “not suitable for implementation” due to potential casemix bias (Kristensen *et al.*, 2009). A 47-member Delphi panel convened by RAND rated this indicator “low” in importance, although an otherwise identical indicator based on clinical data was rated “moderate” in importance and “close to ready for use” (Farley *et al.*, 2008).

18. *Construct (convergent, predictive) validity.* This indicator rates very highly on predictive validity. Cases from the US Nationwide Inpatient Sample (NIS) that were flagged by this PSI in 2000 had 4.3% excess mortality, 9.6 days of excess hospitalisation, and \$38 700 in excess hospital charges, relative to carefully matched controls that were not flagged (Zhan and Miller, 2003). This finding was confirmed in the Veterans Affairs (VA) hospital system, where cases that were flagged by this PSI in 2001 had 2.7% excess mortality, 4.5-9.5 days of excess hospitalisation, and \$7 292-13 816 in excess hospital costs, relative to carefully matched controls that were not flagged (Rivard *et al.*, 2008). A more recent replication using 2007 data, corrected for infections that were reported as “present on admission”, estimated 16.1 attributable hospital days and \$33 118 in attributable hospital costs for the average case (Foster *et al.*, 2009). In a commercial claims database from 45 large employers in the USA, each event (aggregating this PSI with postoperative sepsis) was associated not just with 3.1% excess mortality, but also with 7.7% excess readmissions, which added \$2 594 to the total attributable cost per event (Encinosa and Hellinger, 2008). A case control analysis from England estimated excess mortality of 5.7% and 11.4 days of excess hospitalisation (Raleigh *et al.*, 2008). Finally, the largest reported estimates of these impacts came from a study of children at 38 freestanding paediatric hospitals in the USA in 2006 (i.e., 22.4 hospital days, \$172 484 in hospital charges) (Kronman *et al.*, 2008).

19. In a study testing construct validity using an implicit process measure of quality (Miller *et al.*, 2005), smoothed rates of this PSI among 2 116 hospitals surveyed by the Joint Commission in 1997-1999 were not associated with summary evaluation scores (although a later study of 115 hospitals surveyed in 2002 found a significant association with one patient safety practice subscore on “assessing patient needs”) (Thornlow and Merwin, 2009). Indeed, the occurrence of one or more events flagged by this indicator, among 3 594 hospitals that treated Medicare patients in the USA in 2003, was associated with *better* performance on process-of-care measures for three medical conditions and *lower* risk-adjusted mortality for five of six high-risk categories of patients (Isaac and Jha, 2008). Rates of this PSI were inversely associated with adoption of strategic information technology applications among 98 Florida hospitals (Menachemi *et al.*, 2007), and with adoption of electronic medical record (EMR) systems among Medicare patients (Parente and McCullough 2009), but not among 66 Georgia hospitals (Culler *et al.*, 2007).

20. A correlational study based on the 1997-2002 Nationwide Inpatient Sample (USA) labelled this indicator as a “canary measure” because it was significantly and consistently associated with at least nine other AHRQ Patient Safety Indicators at the hospital level (Yao *et al.*, 2009). For both Medicare and Veterans Health Administration patients, this indicator loaded strongly with two other PSIs (iatrogenic pneumothorax and postoperative DVT/PE) on a common factor (Rosen *et al.*, 2009). Finally, this indicator was significantly associated with re-admission within three months (risk ratio= 1.29), but not within one month (risk ratio=1.00), after adjusting for patient characteristics using 2004 surgical data from seven US states (Friedman *et al.*, 2009). Unadjusted data from England confirm the association between this PSI and re-admission (Bottle and Aylin, 2009).

21. Although children aged less than 15 years were excluded from OECD analyses, data on the validity of this indicator in paediatric populations may still be relevant. Physicians participating in the National Association of Children’s Hospitals and Related Institutions’ (NACHRI) Pediatric PSI Collaborative reviewed 145 flagged events from 14 hospitals in 2003, using an online tool to assess implicit process of care, and judged 39% to be preventable and only 31% to be clearly non-preventable (Sedman *et al.*, 2005; Scanlon *et al.*, 2006). In a follow-up study of 2003-2005 data from 28 children’s hospitals (N=285 events), 11% were incorrectly coded and 43% of the remainder were present on admission; only 20-41% of all flagged events were judged to be preventable (Scanlon *et al.*, 2008).

22. *Criterion validity.* No evidence about the criterion validity of this indicator was available before its original release as an AHRQ indicator. It is now known that this indicator has a minor problem due to missing data about timing. Some US data sets now include a “flag” variable denoting whether each diagnosis was present at admission. The percentage of cases flagged by this PSI for whom the event was reported to be a complication of the hospital stay was 65% in California, 65% in New York, 60% in the Rochester, Minnesota area, and 36-44% at the University of Michigan (Houchens *et al.*, 2008; Naessens *et al.*, 2007; Bahl *et al.*, 2008). Hospital-specific rates including infections reported as present on admission were moderately correlated with hospital-specific rates excluding such infections ($r=0.91$ in California, $r=0.88$ in New York), especially among coronary artery bypass surgery patients ($r=0.99$ in California) (Glance *et al.*, 2008).

23. The best recent evidence about the positive predictive value (PPV) of this indicator comes from the 47 hospitals participating in the AHRQ PSI Validation Pilot Project (N=191). In this study, 20% of the flagged events were present at admission, 21% lacked clear documentation of an eligible infection, and 4% had an unreported disqualifying comorbid condition (i.e., cancer, severe malnutrition, immunodeficiencies), leaving 55% that were confirmed as iatrogenic complications (Zrelak *et al.*, 2009). All of the confirmed events were attributable to a vascular device, including central venous catheters (74%), peripheral venous catheters, and arterial catheters.

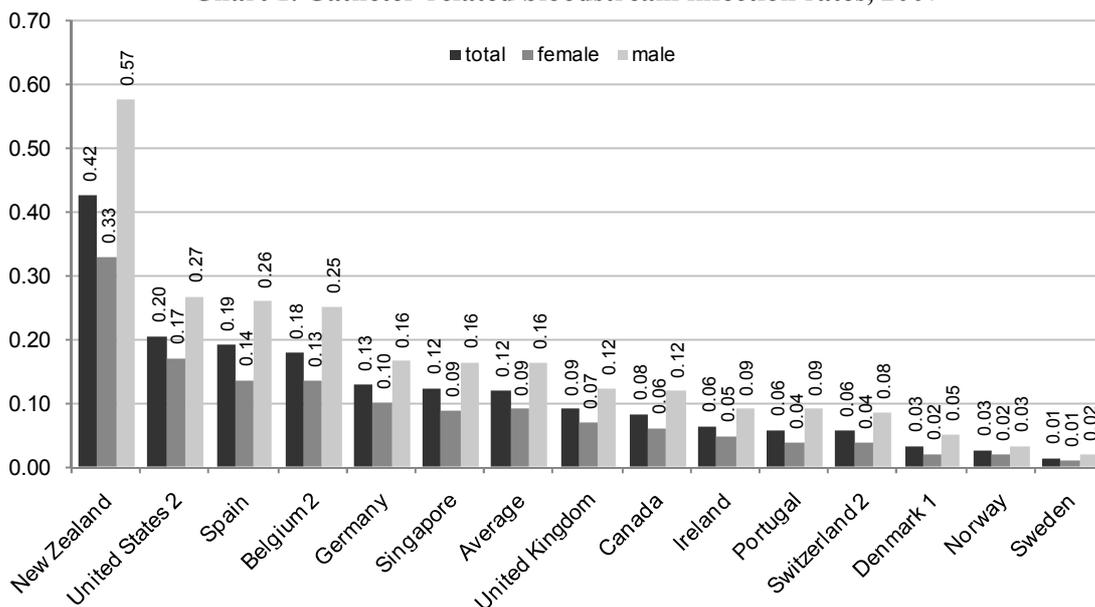
24. A similar review of medical records of 168 cases from 18 English NHS (National Health Service) trusts found that 6% of the flagged events were present at admission and 12% were miscoded, leaving 79% that were confirmed (Bottle and Aylin 2008). Finally, evidence from New York and New Zealand suggests that a significant number of true events may not be ascertained because they occur after hospital discharge; linking 30-day re-admissions in New York increased the overall rate of this PSI from 2.02 to 2.52 per 1 000 eligible discharges; 56% of the post-discharge events were complications of haemodialysis (Gallagher *et al.*, 2005a; Stevanovic 2009). One study from 24 US hospitals participating in a patient safety collaborative reported the sensitivity of this PSI as only 9% relative to case ascertainment using National Healthcare Safety Network protocols (N=89); PPV could not be evaluated due to the study design (Stone *et al.*, 2007).

25. In summary, recent evidence on construct validity and criterion validity is moderately supportive of this indicator. A recent change to the ICD-9-CM coding of catheter-associated infections (999.31 = “infection due to central venous catheter”) should enhance its criterion validity in countries that use ICD-9-CM, by excluding most of the false positive cases captured by the previous definition. Based on limited information, underreporting is a serious concern. Risk adjustment is recommended for inter-provider comparisons (AHRQ, 2007) to ensure that variation due to different patient populations across institutions is removed.

Findings

26. Reported catheter-related bloodstream infection rates vary across participating OECD countries between 0.01 and 0.4% (Chart 1). Hospital-acquired infections occur far more often in male patients than in females in all countries. Direct adjustment for 5-year age-gender strata did not affect indicator rates or rankings across countries. A correlation between indicator rate and amount of documentation can be found for the indicator (Chart 2, Spearman $r=0.912$, $p<0.0001$). In Belgium and USA the mean number of secondary diagnoses is more than two times greater than in Denmark and Switzerland, which might explain elevated infection rates reported by these countries. Underreporting is likely for countries with low infection rates. The rates represented in the charts below have not been age-sex standardised or adjusted for secondary diagnoses or length of stay.

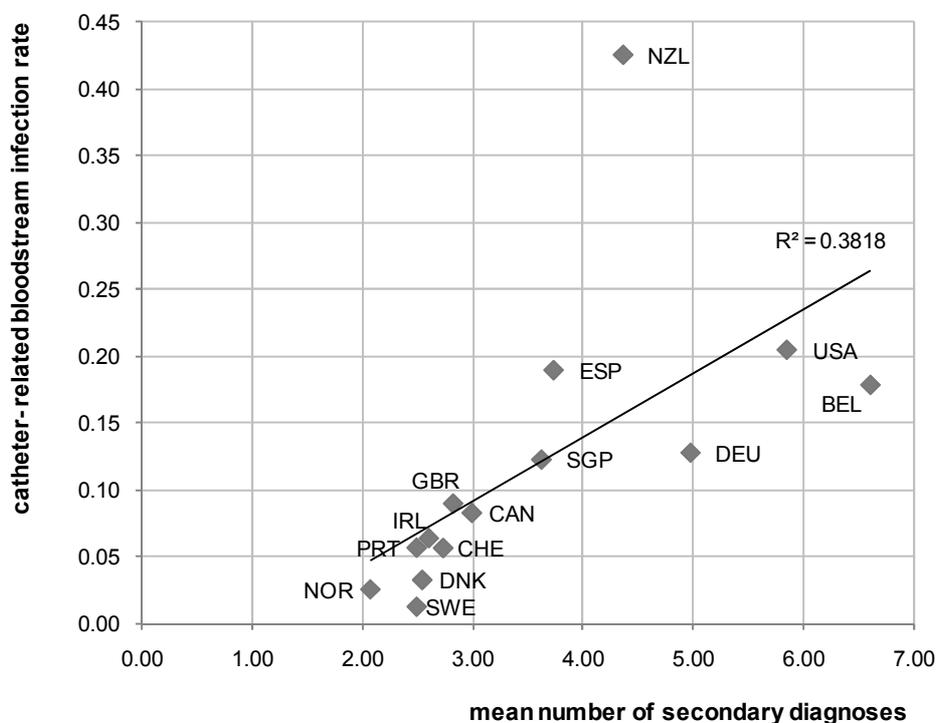
Chart 1. Catheter-related bloodstream infection rates, 2007



1. 2008. 2. 2006

Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project.

Chart 2. Catheter-related bloodstream infection rates and mean number of secondary diagnoses, 2007



Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project.

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

27. *Rationale and concerns.* The occurrence of postoperative PE/DVT can range from mild symptoms to devastating clinical consequences including pain, respiratory distress, and death. Regarding pulmonary embolism, the mortality rate is less than 8% when detected and correctly treated but about 30% when the condition is unrecognised and not treated (Olin, 2002). Because PE/DVT can cause unnecessary prolongation of hospital stays as well as unnecessary pain, suffering and death, this indicator has important financial and quality improvement implications. This adverse event can be prevented through the appropriate use of anticoagulants and other preventive measures; evidence based guidelines are available (Geerts *et al.*, 2004). This indicator based on administrative hospital data was proposed in 1994 (Iezzoni *et al.*, 1994) and subsequently further investigated (Miller *et al.*, 2001).

28. *Content (consensual) validity.* Content validity was addressed in the technical report accompanying the original release of the AHRQ PSIs. Although panellists' ratings of the "usefulness" of each candidate indicator were used to select the final PSI set, panellists were also asked to rate each indicator on its preventability and its likelihood of being due to medical error. The median ratings of this indicator on the former dimension from two independent panels were 7 and 6, with indeterminate agreement and disagreement, respectively, leading to classifications of "acceptable" and "unclear", respectively. The median ratings of this indicator on the latter dimension were 6 and 3, with indeterminate agreement, leading to classifications of "unclear" and "unacceptable", respectively.

29. Through similar processes, this PSI was endorsed by the OECD Patient Safety Panel (Millar *et al.*, 2004) and the National Quality Forum (time-limited due to ICD-9-CM changes), but rejected by the SimPatIE (Safety Improvement for Patients in Europe) project as "not suitable for implementation" due to potential casemix bias (Kristensen *et al.*, 2009). A 47-member Delphi panel convened by RAND rated this indicator "moderate" in importance, but lower on validity (i.e., median ratings 8 and 6, respectively) (Farley *et al.*, 2008). Over diagnosis through increased ultrasound screening of high-risk but asymptomatic postoperative patients is an emerging concern (Haut *et al.*, 2007).

30. *Construct (convergent, predictive) validity.* This indicator rates very highly on predictive validity. Cases from the NIS that were flagged by this PSI had 6.6% excess mortality, 5.4 days of excess hospitalisation, and \$21 700 in excess hospital charges, relative to carefully matched controls that were not flagged (Zhan and Miller, 2003). This finding was confirmed in the Veterans Affairs hospital system, where cases that were flagged by this PSI had 6.1% excess mortality, 4.5-5.5 days of excess hospitalisation, and \$7 205-9 064 in excess hospital costs, relative to carefully matched controls that were not flagged (Rivard *et al.*, 2008). A more recent replication using 2007 data, corrected for thromboses that were reported as "present on admission", estimated 7.8 attributable hospital days and \$18 331 in attributable hospital costs for the average case (Foster *et al.*, 2009).

31. In a study testing construct validity using an implicit process measure of quality (Miller *et al.*, 2005), smoothed rates of this PSI among 2 116 hospitals surveyed by the Joint Commission in 1997-1999 were marginally ($p=0.06$) associated with summary evaluation scores, in the expected direction. In addition, hospitals with high smoothed rates of this PSI were less likely to receive favourable accreditation decisions than hospitals with lower rates. For both Medicare and Veterans Health Administration patients, this indicator loaded strongly with two other PSIs (iatrogenic pneumothorax and selected infections due to medical care) on a common factor (Rosen *et al.*, 2009). Finally, this indicator was significantly associated with re-admission within either three months (risk ratio=1.28) or one month (risk ratio=1.25), after adjusting for patient characteristics using 2004 surgical data from seven US states (Friedman *et al.*, 2009).

32. Although children aged less than 15 years were excluded from OECD analyses, data on the validity of this indicator in paediatric populations may still be relevant. Physicians participating in the NACHRI Pediatric PSI Collaborative reviewed 120 flagged events from 14 hospitals, using an online tool to assess implicit process of care, and judged only 28% to be preventable and 49% to be clearly non-preventable (Sedman *et al.*, 2005; Scanlon *et al.*, 2006). As a result, this indicator was dropped from the AHRQ Pediatric Quality Indicator module.

33. Historically, this indicator evolved from one of the “flags” in Iezzoni’s Complications Screening Program (CSP). Explicit process of care failures in the CSP validation study were relatively frequent among cases flagged on this indicator (72% of major surgery patients, 69% of medical patients), after excluding patients who had DVT/PE at admission, but unflagged controls were not evaluated on the same criteria (Iezzoni *et al.*, 1999). Major surgical cases flagged on this indicator and unflagged controls differed marginally (11% versus 4%, $p=0.09$) on a composite of 17 generic process criteria. Physician reviewers identified potential quality problems in 50% of major surgery patients and 20% of medical patients flagged on this indicator, versus 2% of unflagged controls for each risk group (Weingart *et al.*, 2000).

34. At least two older studies assessed the construct validity of the ICD-9-CM codes mapped to this PSI through correlation with structural measures of nurse staffing. Needleman and Buerhaus (Needleman *et al.*, 2002) found that nurse staffing was independent of the occurrence of DVT/PE among both major surgical and medical patients from 799 hospitals in 11 states in 1997. However, Kovner and Gergen reported that among 506 community hospitals in the 1993 NIS, having more registered nurse hours and non-RN hours per adjusted patient day were both associated with a lower rate of DVT/PE after major surgery (Kovner and Gergen, 1998). Nurse staffing was not associated with the rate of DVT/PE after invasive vascular procedures. Rates of this PSI were marginally ($p=0.06$) associated with adoption of clinical information technology applications among 98 Florida hospitals (Menachemi *et al.*, 2007), but were not associated with adoption of EMR systems among Medicare patients (Parente and McCullough, 2009) or among 66 Georgia hospitals (Culler *et al.*, 2007).

35. *Criterion validity.* The original CSP definition of this PSI, which differed slightly from the current AHRQ definition, had an adequate confirmation rate among major surgical cases sampled from FY1994 Medicare inpatient claims from California and Connecticut (i.e., 59% according to coders, 70% according to physicians, 68% according to nurses who relied on physician documentation) (Lawthers *et al.*, 2000; Weingart *et al.*, 2000; McCarthy *et al.*, 2000). Several smaller, older studies also suggested adequate sensitivity and PPV of PE codes among surgical patients, although the sensitivity of DVT codes was notably poorer (Keeler *et al.*, 1992; Romano *et al.*, 2002; Hawker *et al.*, 1997; Best *et al.*, 2002). Based on these findings, AHRQ limited this PSI to surgical cases (defined using DRGs).

36. One weakness of this indicator is its inability to distinguish thromboses that were present at admission from thromboses that developed during a hospital stay. Some US data sets now include a “flag” variable denoting whether each diagnosis was present at admission. The percentage of cases flagged by this PSI for whom the event was reported to be a complication of the hospital stay was only 46% in California, 43% in New York, 40% in the Rochester, Minnesota area, and 51-67% at the University of Michigan (Houchens *et al.*, 2008; Naessens *et al.*, 2007; Bahl *et al.*, 2008). Hospital-specific rates including thromboses reported as present on admission were variably correlated with hospital-specific rates excluding such thromboses ($r=0.80$ in California, $r=0.41$ in New York), even among coronary artery bypass surgery patients ($r=0.63$ in California) (Glance *et al.*, 2008).

37. The best recent evidence about the PPV of this indicator comes from the 47 hospitals participating in the AHRQ PSI Validation Pilot project (N=155) and the 33 hospitals participating in a parallel benchmarking initiative by the University Health System Consortium (UHC, N=505). In this study, only 13% of the flagged events were present at admission and 8% lacked clear documentation of an

acute venous thrombosis, leaving 79% that were confirmed as iatrogenic complications (White *et al.*, 2009). However, 29% of the confirmed events involved upper extremity, thoracic, or superficial veins (which are not the target for prevention), reducing the overall PPV for clinical purposes to 56% (95% CI, 52-60%). In the 33 teaching hospitals, abstractors also reviewed 517 cases that were not flagged as having this PSI, and found zero false-negatives (sensitivity=100%; 95% CI, 53-100%). A single US teaching hospital (not in the UHC study) separately reported 87% sensitivity and 50% PPV, using the same clinical definition (Henderson *et al.*, 2009), and eight Belgian hospitals reported 54-59% PPV with only one false negative case among 1,392 records reviewed (Gillet *et al.*, 2008). These estimates significantly exceed the sensitivity and PPV estimates of 68% and 29%, respectively, based on 2002-2004 data from the Medicare Patient Safety Monitoring System (Zhan *et al.*, 2007), suggesting that the validity of this indicator may depend on the number of available diagnosis fields (as Medicare data are limited to nine secondary diagnoses).

38. Comparing hospital administrative data from the Department of Veterans Affairs against the National Surgical Quality Improvement Program's clinically abstracted data from 2001, Romano *et al.*, (2009) reported a sensitivity of 56% and a PPV of 22%, although the PPV in a more recent review of 112 randomly selected cases from 2004-2007 was 56% (Borzecki *et al.*, 2009). As in the AHRQ PSI Validation Pilot, most of the false positives in the VA were attributable to chronic thromboses that were present at admission, upper extremity thromboses, or superficial lower extremity thromboses that did not require anticoagulation.

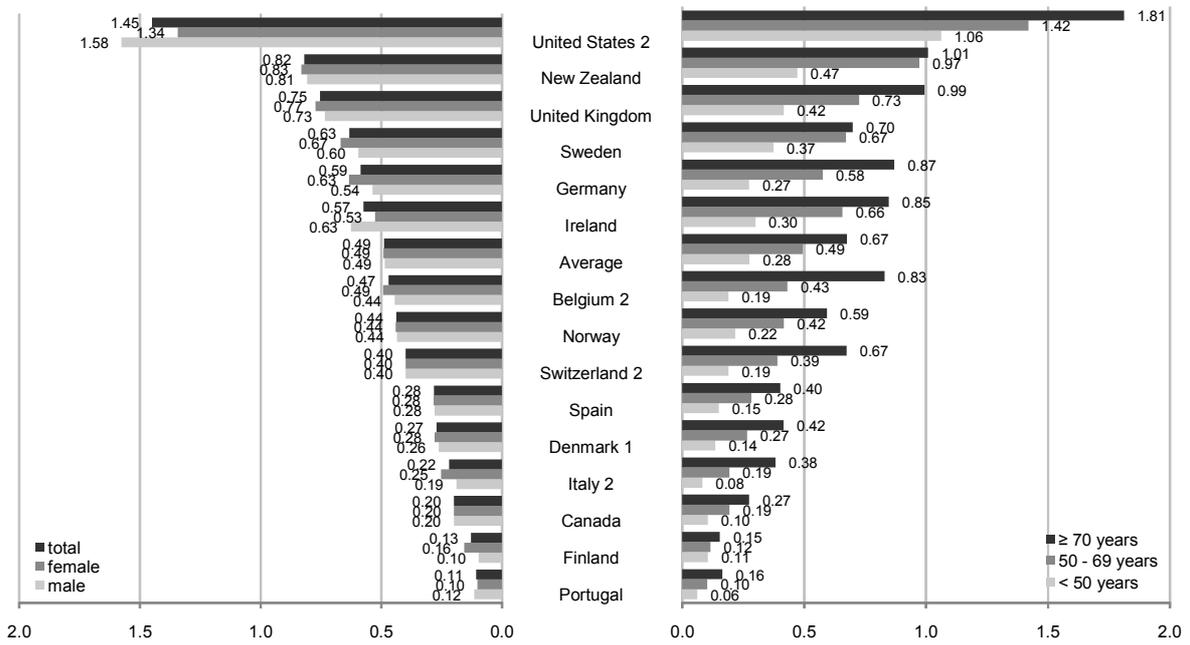
39. Finally, evidence from New York, Denmark, and New Zealand suggests that a significant number of true events may not be ascertained because they occur after hospital discharge; linking 30-day re-admissions in New York increased the overall rate of this PSI from 9.3 to 11.3 per 1,000; 45% of the post-discharge events were pulmonary emboli (Weller *et al.*, 2004; Stevanovic 2009).

40. In summary, recent evidence suggests that this indicator should be used very cautiously for comparing hospital performance unless validated information is available about the timing of the diagnosis and/or the specific veins involved. Recent changes to the ICD-9-CM coding of deep vein thromboses, including new codes specifying superficial, upper extremity, and thoracic thromboses, should enhance criterion validity in countries that use ICD-9-CM, by excluding most of the false positive cases captured by the previous definition. Underreporting does NOT appear to be a problem, based on studies from the USA and Belgium. Over diagnosis through screening of high-risk but asymptomatic postoperative patients is an emerging concern, and may explain the markedly elevated risk of this event at major teaching hospitals and large hospitals (Vartak *et al.*, 2008). Risk adjustment is recommended for inter-provider comparisons (AHRQ, 2007) to ensure that variation due to different patient populations across institutions is removed.

Findings

41. Reported postoperative PE/DVT rates vary across participating OECD countries between 0.1 and 1.4% (Chart 3). Most countries report slightly higher rates in females, but the U.S. and Ireland have higher rates in men and a larger gender discrepancy. As expected, rates are higher in patients aged 70 years and older than in younger age groups in all countries (Chart 3). Direct adjustment for 5-year age-gender strata did not materially affect indicator rates across countries, although two countries moved up one rank (and two others moved down one rank to compensate). A statistically significant dependency between indicator rates and amount of administrative documentation, expressed as the mean number of secondary diagnoses, was found and is shown in Chart 4 (Spearman $r=0.670$, $p=0.009$). Countries with more coded diagnoses reported higher rates. The rates represented in the charts below have not been age-sex standardised or adjusted for secondary diagnoses or length of stay.

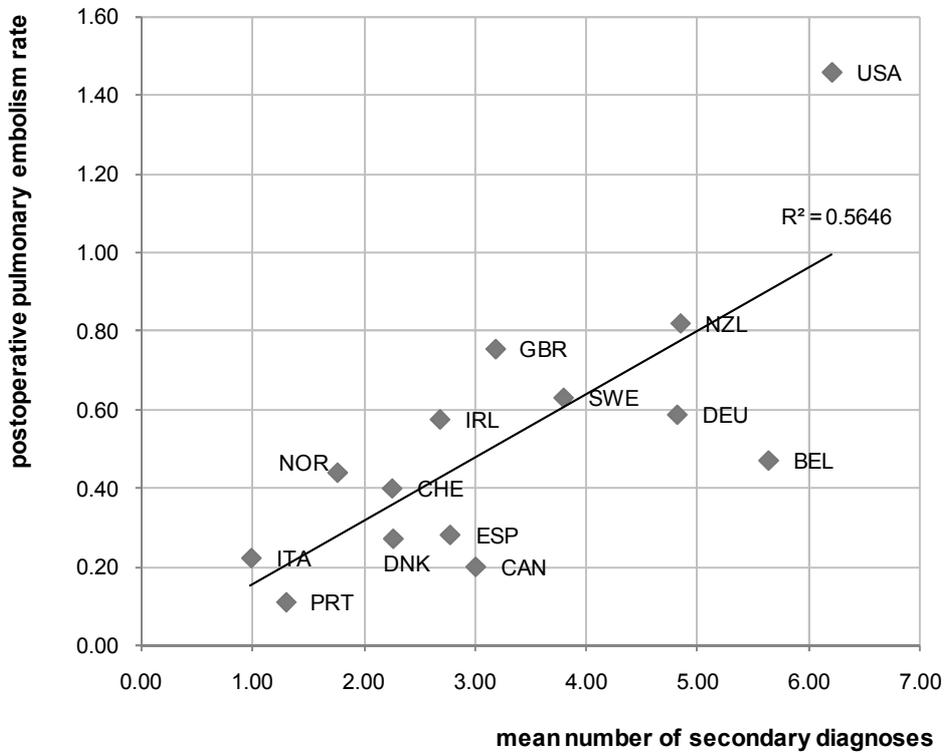
Chart 3. Postoperative pulmonary embolism or deep vein thrombosis, 2007



1. 2008. 2. 2006.

Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project.

Chart 4. Postoperative pulmonary embolism and mean number of secondary diagnoses



Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project.

Postoperative sepsis

Validity

42. *Rationale and concerns.* The occurrence of sepsis following surgery is a severe complication with a mortality rate of up to 30%. 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 good postoperative care. 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. Consequently, this indicator is a plausible patient safety measure. Given the dramatic nature of this complication, it is likely to be reliably coded in administrative data sources, relative to less serious complications (Miller *et al.*, 2004).

43. *Content (consensual) validity.* Content validity was addressed in a technical report accompanying the original release of the AHRQ PSIs. Although panellists' ratings of the "usefulness" of each candidate indicator were used to select the final PSI set, panellists were also asked to rate each indicator on its preventability and its likelihood of being due to medical error. The median rating of this indicator was 6.5 (with agreement) on the former dimension, and 6 (with indeterminate agreement) on the latter dimension, leading to a classification of "unclear" on both dimensions. Through similar processes, this PSI was endorsed by both the OECD Patient Safety Panel (Millar *et al.*, 2004) and the SimPatIE (Safety Improvement for Patients in Europe) project (Kristensen *et al.*, 2009). A 47-member Delphi panel convened by RAND rated this indicator "moderate" in importance, but lower on validity (i.e., median ratings 8 and 5, respectively) (Farley *et al.*, 2008).

44. *Construct (convergent, predictive) validity.* This indicator rates very highly on predictive validity. Cases from the NIS that were flagged by this PSI had 21.9% excess mortality, 10.9 days of excess hospitalisation, and \$57 700 in excess hospital charges, relative to carefully matched controls that were not flagged (Zhan and Miller, 2003). This finding was confirmed in the Veterans Affairs hospital system, where cases that were flagged by this PSI had 30.2% excess mortality, 5.7-18.8 days of excess hospitalisation, and \$13 395-31 262 in excess hospital costs, relative to carefully matched controls that were not flagged (Rivard *et al.*, 2008). A more recent replication using 2007 data, corrected for sepsis that was reported as "present on admission", estimated 13.7 attributable hospital days and \$39 117 in attributable hospital costs for the average case (Foster *et al.*, 2009). In a commercial claims database from 45 large employers in the USA, each event (aggregating this PSI with catheter-related bloodstream infections) was associated not just with 3.1% excess mortality, but also with 7.7% excess readmissions, which added \$2 594 to the total attributable cost per event (Encinosa and Hellinger, 2008). A case control analysis from England estimated excess mortality of 27.1% and 15.9 days of excess hospitalisation (Raleigh *et al.*, 2008). Finally, the largest reported estimates of these impacts came from a study of children at 38 freestanding paediatric hospitals in the USA in 2006 (i.e., 23.5 hospital days, \$261 173 in hospital charges) (Kronman *et al.*, 2008). In a study testing construct validity using an implicit process measure of quality (Miller *et al.*, 2005), smoothed rates of this PSI among 2 116 hospitals surveyed by the Joint Commission in 1997-1999 were marginally ($p=0.10$) associated with summary evaluation scores, in the expected direction. For both Medicare and Veterans Health Administration patients, this indicator loaded strongly with two other PSIs (postoperative sepsis and respiratory failure) on a common factor representing perioperative continuity of care (Rosen *et al.*, 2009). Finally, this indicator was significantly associated with re-admission within three months (risk ratio=1.26), but not within one month (risk ratio=0.99), after adjusting for patient characteristics using 2004 surgical data from seven US states (Friedman *et al.*, 2009). Unadjusted data from England confirm the association between this PSI and re-admission (Bottle and Aylin, 2009).

Although children aged less than 15 years were excluded from OECD analyses, data on the validity of this indicators in paediatric populations may still be relevant. In a study of 2003-2005 data from 28 children's hospitals (N=279 events), 20% were incorrectly coded and 40% of the remainder were present on admission; only 12-32% of all flagged events were judged to be preventable (Scanlon *et al.*, 2008).

45. At least one older study assessed the construct validity of the ICD-9-CM codes mapped to this PSI through correlation with structural measures of nurse staffing. Needleman and Buerhaus (Needleman *et al.*, 2002) found that nurse staffing was independent of the occurrence of sepsis among both major surgical and medical patients from 799 hospitals in 11 states in 1997. Rates of this PSI were inversely associated with adoption of clinical information technology applications among 98 Florida hospitals (Menachemi *et al.*, 2007), but not among 66 Georgia hospitals (Culler *et al.*, 2007).

46. *Criterion validity.* Several small studies provided limited data about the criterion validity of the ICD-9-CM codes mapped to this indicator before its release. Unfortunately, several of these studies either did not clearly document their ICD-9 definitions (Massanari *et al.*, 1987; Belio-Blasco *et al.*, 2000) or did not stratify the subgroup of patients with a secondary diagnosis of DVT/PE (Barbour, 1993). In comparison with the VA's National Surgical Quality Improvement Program database from 123 hospitals in 1994-95, in which "systemic sepsis" was defined by a positive blood culture with systemic manifestations of sepsis within 30 days after surgery, ICD-9-CM diagnoses had a sensitivity of 37% and a PPV of 30% (Best *et al.*, 2002).

47. This indicator has a minor problem due to missing data about timing. Some US data sets now include a "flag" variable denoting whether each diagnosis was present at admission. The percentage of cases flagged by this PSI for whom the event was reported to be a complication of the hospital stay was 73% in California, 70% in New York, 76% in the Rochester, Minnesota area, and 59-60% at the University of Michigan (Houchens *et al.*, 2008; Naessens *et al.*, 2007; Bahl *et al.*, 2008). Hospital-specific rates including sepsis reported as present on admission were moderately correlated with hospital-specific rates excluding such infections ($r=0.72$ in California, $r=0.82$ in New York), especially among coronary artery bypass surgery patients ($r=0.93$ in California) (Glance *et al.*, 2008).

48. The best recent evidence about the PPV of this indicator comes from the 47 hospitals participating in the AHRQ PSI Validation Pilot Project (N=164). In this study, 17% of the flagged events (or precursor infections) were present at admission, and 17% lacked clear documentation of sepsis, bacteraemia, or SIRS (systemic inflammatory response syndrome) with infection, leaving 66% that were confirmed as complications. (However, an additional 25% of flagged cases may have been ineligible because the reviewer perceived the "index" surgery as being non-elective.) The primary site of infection was catheter-related in 24%, lungs in 39%, surgical site in 9%, and urinary tract in 19%. Comparing hospital administrative data from the Department of Veterans Affairs against the National Surgical Quality Improvement Program's clinically abstracted data from 2001, Romano *et al.*, (2009) reported a sensitivity of 37%, PPV of 45%, and positive likelihood ratio of 131. Most of the "false positives" were patients with clinical evidence of sepsis, who were treated for presumptive sepsis, but lacked "definitive evidence of infection". A similar review of medical records of 53 cases from 18 English NHS trusts found that 6% of the flagged events were present at admission and 21% were miscoded, leaving 70% that were confirmed (Bottle and Aylin, 2008). Eight Belgian hospitals reported 45% PPV (largely due to overreporting or failure to satisfy clinical criteria for sepsis), with 25 false negative cases among 1 396 records reviewed (98.2% negative predictive value) (Gillet *et al.*, 2008).

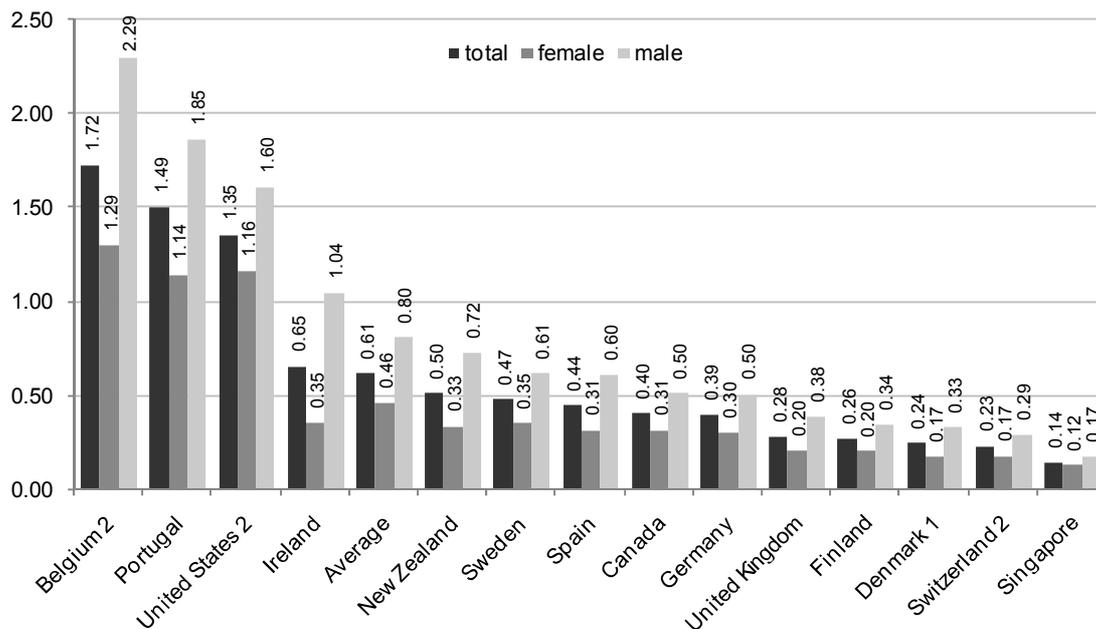
49. In summary, recent evidence on construct validity and criterion validity is somewhat supportive of this indicator, but raises questions about the ability to accurately identify patients for elective surgery. Very little evidence is available on potential underreporting. Risk adjustment is recommended for inter-

provider comparisons (AHRQ, 2007) to ensure that variation due to different patient populations across institutions is removed.

Findings

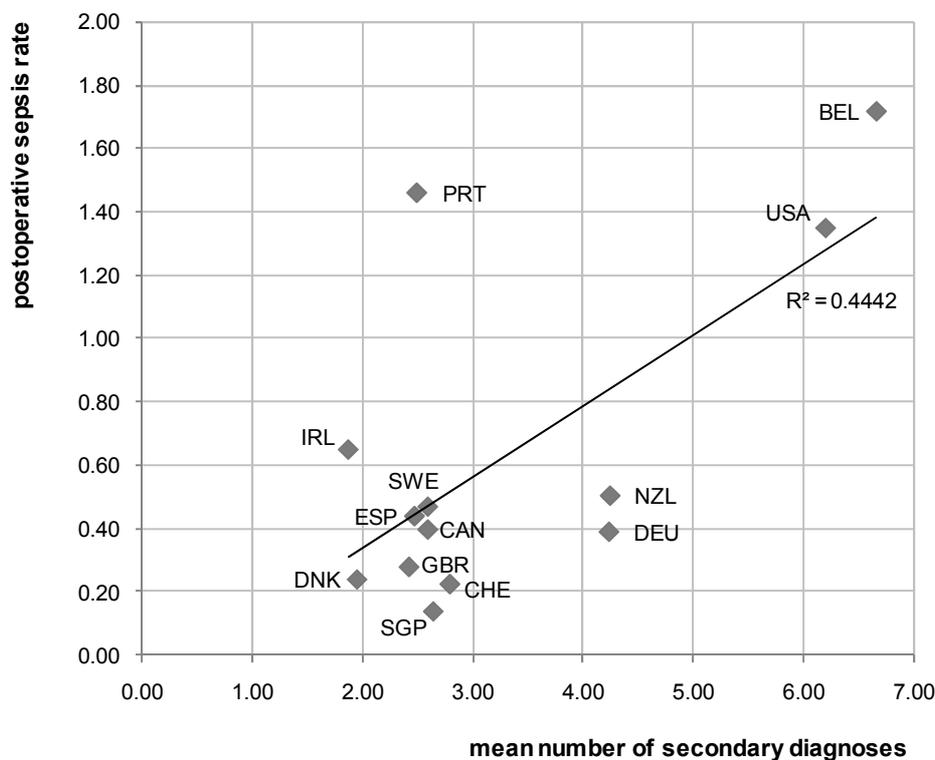
50. Reported postoperative sepsis rates vary across participating countries between 0.2 and 1.7% (Chart 5); the average is 0.61%. Three countries (Belgium, Portugal, and USA) reported markedly higher sepsis rates than the others. Postoperative sepsis occurs far more often in male patients than in females in all countries. This finding corresponds to the literature (Angus *et al.*, 2001). Direct adjustment for 5-year age-gender strata did not materially affect indicator rates across countries, although three countries moved up one rank (and three others moved down one rank to compensate). A non-significant correlation between indicator rates and amount of documentation was found (Chart 6, Spearman $r=0.414$, $p=0.142$), based on the mean number of secondary diagnoses in denominator cases. In Belgium and USA, the mean number of secondary diagnoses is more than two times greater than in Singapore, which might explain elevated infection rates in the first two countries. Underreporting is likely for countries with low infection rates. The rates represented in the charts below have not been age-sex standardised or adjusted for secondary diagnoses or length of stay.

Chart 5. Postoperative sepsis rates, 2007



1. 2008. 2. 2006.

Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project.

Chart 6. Postoperative sepsis rates and mean number of secondary diagnoses, 2007

Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project.

Accidental puncture or laceration

Validity

51. *Rationale and concerns.* This indicator captures events related to technical and process limitations of hospital care. While accidental cut, puncture, perforation or laceration during a surgical procedure is a recognised risk, for example of abdominal surgery, elevated rates of such complications may indicate systems problems, such as inadequate training or fatigued health staff. Experts involved in the selection of this indicator stated that some incidents might not be preventable (AHRQ, 2007). The indicator captures surgical as well as medical discharges based on administrative hospital data. They were proposed in 1994 (Iezzoni *et al.*, 1994) and subsequently further investigated (Miller *et al.*, 2001).

52. *Content (consensual) validity.* Content validity was addressed in the technical report accompanying the original release of the AHRQ PSIs. Although panellists' ratings of the "usefulness" of each candidate indicator were used to select the final PSI set, panellists were also asked to rate each indicator on its preventability and its likelihood of being due to medical error. The median rating of this indicator was 7, with agreement, on the former dimension, leading to a classification of "acceptable". The median rating was 6, with indeterminate agreement, on the latter dimension, leading to a classification of "unclear". Through similar processes, this PSI was endorsed by the OECD Patient Safety Panel (Millar *et al.*, 2004) and the US National Quality Forum, but rejected by the SimPatIE (Safety Improvement for Patients in Europe) project as "not workable for implementation in Europe". A 47-member Delphi panel convened by RAND rated this indicator "low" in importance, although the median rating of 7 would have qualified for endorsement in the AHRQ panel process (Farley *et al.*, 2008).

53. *Construct (convergent, predictive) validity.* This indicator rates highly on predictive validity. Cases from the NIS that were flagged by this PSI had 2.2% excess mortality, 1.3 days of excess hospitalisation, and \$8 300 in excess hospital charges, relative to carefully matched controls that were not flagged (Zhan and Miller, 2003). This finding was confirmed in the Veterans Affairs hospital system, where cases that were flagged by this PSI had 3.2% excess mortality, 1.4-3.1 days of excess hospitalisation, and \$3 359-6 880 in excess hospital costs, relative to carefully matched controls that were not flagged (Rivard *et al.*, 2008). A more recent replication using 2007 data, corrected for injuries that were reported as “present on admission”, estimated 3.7 attributable hospital days and \$12 087 in attributable hospital costs for the average case (Foster *et al.*, 2009). In a commercial claims database from 45 large employers in the USA, these events (aggregated with four rarer PSIs) were not associated with either excess mortality or excess readmissions (Encinosa and Hellinger, 2008). Finally, the largest reported estimates of these impacts came from a study of children at 38 freestanding paediatric hospitals in the USA in 2006 (i.e., 2.8 hospital days, \$34 884 in hospital charges) (Kronman *et al.*, 2008).

54. In a study testing construct validity using an implicit process measure of quality (Miller *et al.*, 2005), smoothed rates of this PSI among 2,116 hospitals surveyed by the Joint Commission in 1997-1999 were significantly ($p < 0.01$) associated with summary evaluation scores, in the expected direction. For both Medicare and Veterans Health Administration patients, this indicator loaded strongly with three other PSIs (foreign body left in, postoperative haemorrhage or hematoma, and wound dehiscence) on a common factor representing technical complications of care (Rosen *et al.*, 2009). Finally, this indicator was significantly associated with re-admission within either three months (risk ratio=1.29) or one month (risk ratio=1.00), after adjusting for patient characteristics using 2004 surgical data from seven US states (Friedman *et al.*, 2009). Unadjusted data from England confirm the association between this PSI and re-admission (Bottle and Aylin, 2009).

55. Rates of this PSI were inversely associated with adoption of strategic information technology applications among 98 Florida hospitals (Menachemi *et al.*, 2007), but not among 66 Georgia hospitals (Culler *et al.*, 2007).

56. Although children aged less than 15 years were excluded from OECD analyses, data on the validity of this indicator in paediatric populations may still be relevant. Physicians participating in the NACHRI Pediatric PSI Collaborative reviewed 119 flagged events from 14 hospitals, using an online tool to assess implicit process of care, and judged 64% to be preventable and only 14% to be clearly non-preventable (Sedman *et al.*, 2005; Scanlon *et al.*, 2006). In a follow-up study of 2003-2005 data from 28 children’s hospitals (N=285 events), 9% were incorrectly coded and 7% of the remainder were present on admission; 27-57% of all flagged events were judged to be preventable (Scanlon *et al.*, 2008).

57. *Criterion validity.* Several studies that were published before the release of this indicator offered conflicting conclusions about the criterion validity of the underlying ICD-9 codes. For example, a study of laparoscopic cholecystectomy in 18 Ontario hospitals in 1991-95 (Taylor, 1998) found that 95% (99/104) of patients with an ICD-9 code of 998.2 or E870.0 had a confirmed injury to the bile duct or gallbladder (although only 27% were “clinically significant”). A similar study of all cholecystectomies performed in Western Australia between 1988 and 1994 reported that these two codes had a sensitivity of 40% (19/48) and a PPV of 23% (19/84) in identifying bile duct injuries (Valinsky *et al.*, 1999). Among 185 total knee replacement patients from 5 Ontario hospitals in 1984-90, Hawker *et al.*, (1997) found that the sensitivity and PPV of codes describing “miscellaneous mishaps during or as a direct result of surgery” were 86% (6/7) and 55% (6/11), respectively. Romano *et al.*, (2002) identified 19 of 45 chart-confirmed episodes of accidental puncture or laceration using discharge abstracts of discectomy patients at 30 California hospitals in 1990-91, with only one false positive.

58. This indicator does not appear to have a significant problem due to missing data about timing. Some US data sets now include a “flag” variable denoting whether each diagnosis was present at admission. The percentage of cases flagged by this PSI for whom the event was reported to be a complication of the hospital stay was 87% in California, 87% in New York, 85% in the Rochester, Minnesota area, and 84-91% at the University of Michigan (Houchens *et al.*, 2008; Naessens *et al.*, 2007; Bahl *et al.*, 2008). Hospital-specific rates including injuries reported as present on admission were highly correlated with hospital-specific rates excluding such injuries ($r=0.97$ in California, $r=0.96$ in New York, $r=0.95$ for CABG patients).

59. The best recent evidence about the PPV of this indicator comes from the 47 hospitals participating in the AHRQ PSI Validation Pilot Project (N=249). In this study, 2% of the flagged events were present at admission, and 7% lacked clear documentation of an accidental puncture or laceration, leaving about 91% (95% CI, 88-94%) that were confirmed as complications. Of these events, 71% occurred in the abdomen or pelvis, 10% in the chest, and 16% in the spine (Utter *et al.*, 2009). About 75% of the confirmed injuries were categorised as “potentially consequential”, meaning that they would generally require surgical repair. Similarly, the PPV in a recent review of 112 randomly selected cases from the Department of Veterans Affairs in 2004-2007 was 86% (95% CI, 78-92%) (Borzecki *et al.*, 2009).

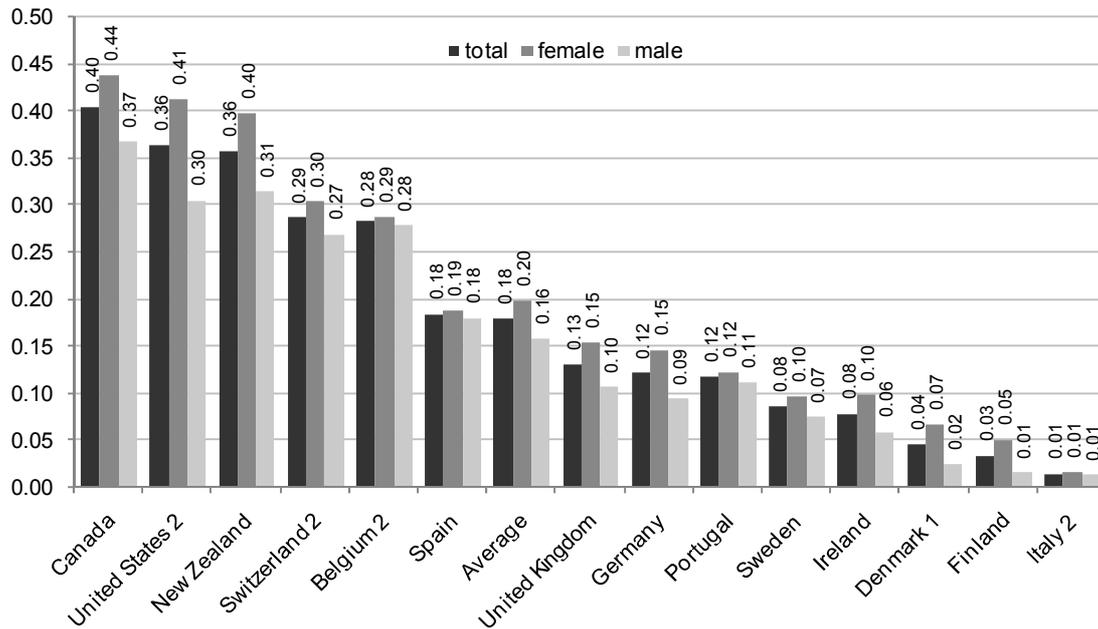
60. There is very limited evidence about the sensitivity of this indicator. Investigators in New York systematically searched their hospital administrative data for procedure codes suggesting repair of iatrogenic injuries, and reported that this PSI may have missed 27% of bladder injuries from hysterectomy, 21% of bowel injuries from cholecystectomy, 47% of abdominal injuries from lysis of adhesions, 54% of abdominal injuries from nephroureterectomy, and 20% of spinal injuries from lumbar surgery (Gallagher *et al.*, 2005b). AHRQ is currently evaluating whether these procedure codes can be added to the PSI definition to improve its sensitivity, without compromising its PPV.

61. In summary, recent evidence on construct validity and criterion validity is moderately supportive of this indicator. Based on limited information, underreporting is a valid concern. Risk adjustment is recommended for inter-provider comparisons (AHRQ, 2007) to ensure that variation due to different patient populations across institutions is removed.

Findings

62. Reported accidental puncture or laceration rates vary across participating OECD countries between 0.01 and 0.4% (Chart 7). Gender subgroups show a slight female predominance of rates in all countries. Direct adjustment for 5-year age-gender strata did not materially affect indicator rates across countries, although two countries switched ranks. Underreporting is assumed for countries reporting low rates. Chart 8 supports this assumption and demonstrates how rates of this indicator are correlated with the amount of administrative documentation (Spearman $r=0.714$, $p=0.006$); however, not all participating countries provided this additional information. The rates represented in the charts below have not been age-sex standardised or adjusted for secondary diagnoses or length of stay.

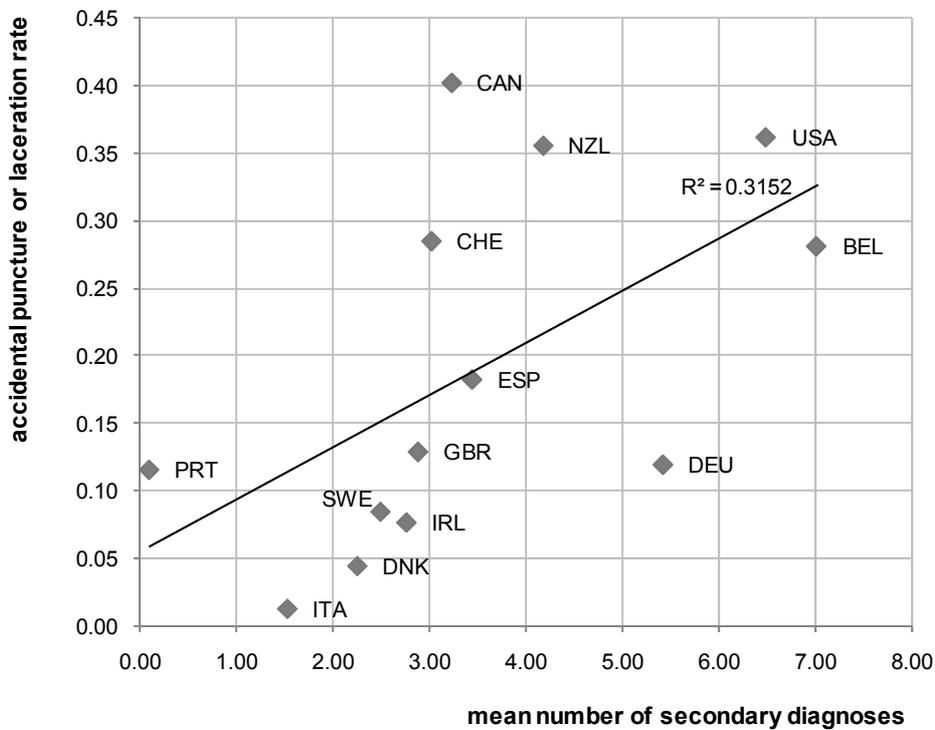
Chart 7. Accidental puncture or laceration rates, 2007



1. 2008. 2. 2006.

Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project.

Chart 8. Accidental puncture or laceration rates and mean number of secondary diagnoses



Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project.

Foreign body left in during procedure*Validity*

63. *Rationale and concerns.* This indicator captures events related to technical and process limitations of hospital care. Errors relating to the failure to remove surgical instruments at the end of a procedure (*i.e.* needles, knife blades, electrosurgical adaptors, safety pins or sponges) are clinically significant in about 50% of all patients with a 10 % mortality following intra-abdominal surgery (Gonzalez-Ojeda *et al.*, 1999). The true frequency of this adverse event remains unclear as underreporting is assumed. Although this indicator captures a rather infrequent complication, it should be addressed after happening. So called sentinel events might reflect serious process problems. This complication is susceptible to be influenced by the health care system: one study identified risk factors for retained instruments and sponges after surgery, namely emergencies, unplanned changes in procedure and obesity (Gawande *et al.*, 2003). The Joint Commission assesses these events as “reviewable” even when the outcome is not death or major permanent loss of function.

64. The indicator captures surgical as well as medical discharges, based on administrative hospital data. It was proposed in 1994 (Iezzoni *et al.*, 1994) and subsequently further investigated (Miller *et al.*, 2001).

65. *Content (consensual) validity.* Content validity was addressed in the technical report accompanying the original release of the AHRQ PSIs. Although panellists’ ratings of the “usefulness” of each candidate indicator were used to select the final PSI set, panellists were also asked to rate each indicator on its preventability and its likelihood of being due to medical error. The median ratings of this indicator on the former dimension from two independent panels were 8 and 7.5, with agreement, leading to classifications of “acceptable” from both panels. The median ratings of this indicator on the latter dimension were 8 and 7, with agreement and indeterminate agreement, respectively, leading to classifications of “acceptable” from both panels. Through similar processes, this PSI was endorsed by the OECD Patient Safety Panel (Millar *et al.* 2004) and the US National Quality Forum, but rejected by the SimPatIE (Safety Improvement for Patients in Europe) project as “not suitable for implementation” due to potential casemix bias (Kristensen *et al.*, 2009). A 47-member Delphi panel convened by RAND rated this indicator “high” in importance, but lower on validity (*i.e.*, median ratings 8 and 6, respectively) (Farley *et al.*, 2008).

66. *Construct (convergent, predictive) validity.* This indicator rates highly on predictive validity. Cases from the NIS that were flagged by this PSI had 2.1% excess mortality, 2.1 days of excess hospitalisation, and \$13 300 in excess hospital charges, relative to carefully matched controls that were not flagged (Zhan and Miller, 2003). A more recent replication using 2007 data, corrected for foreign bodies that were reported as “present on admission”, estimated 4.5 attributable hospital days and \$13 202 in attributable hospital costs for the average case (Foster *et al.*, 2009). Finally, the largest reported estimates of these impacts came from a study of children at 38 freestanding paediatric hospitals in the USA in 2006 (*i.e.*, 14.3 hospital days, \$144 889 in hospital charges) (Kronman *et al.*, 2008).

67. In a study testing construct validity using an implicit process measure of quality (Miller *et al.*, 2005), smoothed rates of this PSI among 2 116 hospitals surveyed by the Joint Commission in 1997-1999 were not associated with summary evaluation scores. For both Medicare and Veterans Health Administration patients, this indicator loaded strongly with three other PSIs (accidental puncture or laceration, postoperative haemorrhage or hematoma, and wound dehiscence) on a common factor representing technical complications of care (Rosen *et al.*, 2009). However, unadjusted data from England showed no association between this PSI and readmission (Bottle and Aylin, 2009).

68. Although children aged less than 15 years were excluded from OECD analyses, data on the validity of this indicators in paediatric populations may still be relevant. Physicians participating in the NACHRI Pediatric PSI Collaborative reviewed 45 flagged events from 14 hospitals, using an online tool to assess implicit process of care, and judged 51% to be preventable and only 27% to be clearly non-preventable (Sedman *et al.*, 2005; Scanlon *et al.*, 2006). In a follow-up study of 2003-2005 data from 28 children's hospitals (N=72 events), 22% were incorrectly coded and 20% of the remainder were present on admission; 28-50% of all flagged events were judged to be preventable (Scanlon *et al.*, 2008).

69. *Criterion validity.* No evidence about the criterion validity of this indicator was available before its original release as an AHRQ indicator. This indicator may have a problem due to missing data about timing. Some US data sets now include a "flag" variable denoting whether each diagnosis was present at admission. The percentage of cases flagged by this PSI for whom the event was reported to be a complication of the hospital stay was 64% in California, 76% in New York, 54% in the Rochester, Minnesota area, and 33-80% at the University of Michigan (Houchens *et al.*, 2008; Naessens *et al.*, 2007; Bahl *et al.*, 2008). Hospital-specific rates including foreign bodies reported as present on admission were highly correlated with hospital-specific rates excluding such foreign bodies ($r=0.89$ in California, $r=0.94$ in New York).

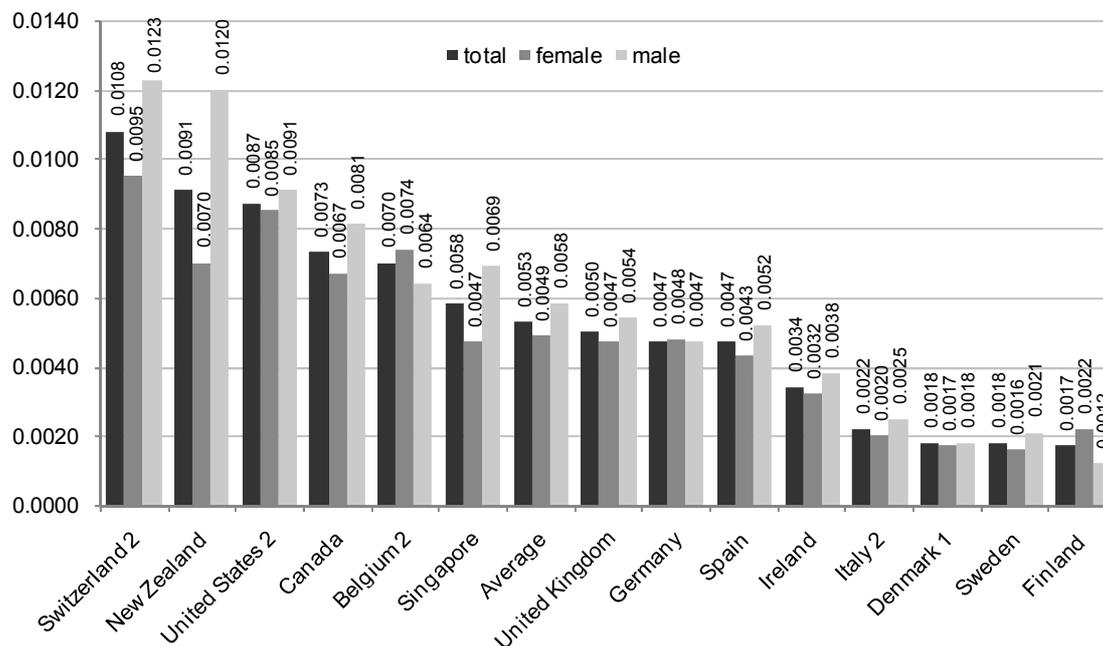
70. A review of medical records of 29 cases from 18 English NHS (National Health Service) trusts found that 10% of the flagged events were present at admission and 38% were miscoded, leaving 52% that were confirmed (Bottle and Aylin, 2008). A validation study in the USA is currently underway. However, recent evidence from New Zealand suggests that a significant number of true events may not be ascertained because they occur after hospital discharge (Stevanovic, 2009).

71. In summary, recent evidence on construct validity and criterion validity is moderately supportive of this indicator. No information on underreporting is currently available. Risk adjustment is not recommended, based on the rarity and sentinel nature of this outcome (AHRQ, 2007).

Findings

72. Reported rates of foreign body left in during procedure vary across participating OECD countries from 2 to 11 cases per 100 000 hospital admissions (Chart 9). Regarding inter-country variability, this indicator varies least across countries among all patient safety indicators. Rather low rates in general reflect the nature of a sentinel event indicator and make scientifically sound interpretations of age or gender subgroup data difficult. Direct adjustment for 5-year age-gender strata did not materially affect indicator rates across countries, although one country moved up two ranks (and two others moved down one rank to compensate). Underreporting is suspected for countries reporting low rates. Chart 10 supports this hypothesis and demonstrates how indicator rates are dependent on the amount of administrative documentation (Spearman $r=0.621$, $p=0.024$); however, the relationship is less consistent than for other PSIs. The rates represented in the charts below have not been age-sex standardised or adjusted for secondary diagnoses or length of stay.

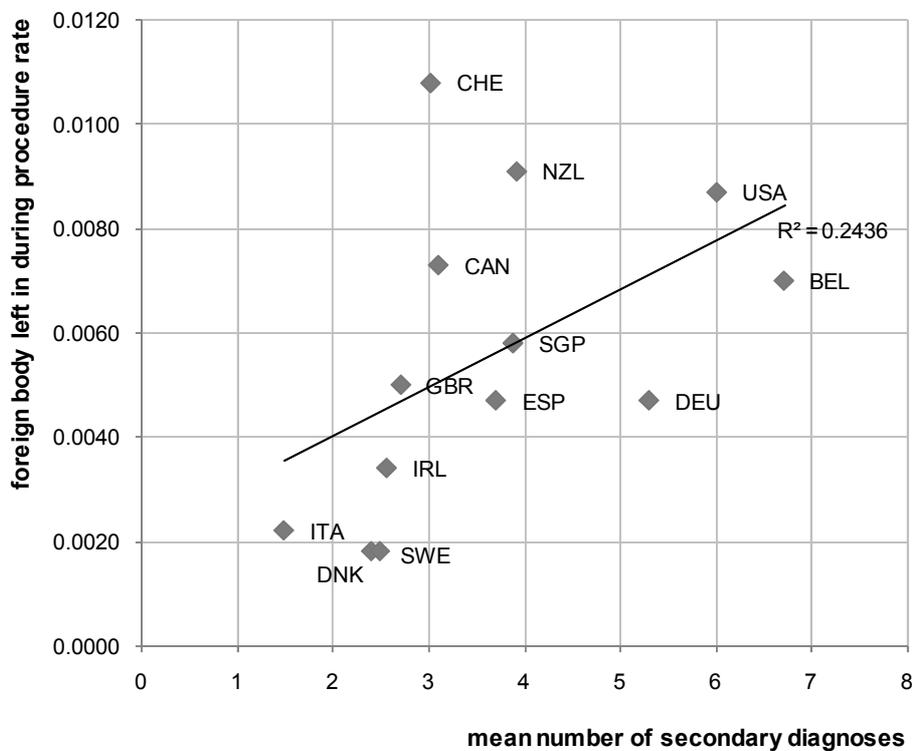
Chart 9. Foreign body left in during procedure rates, 2007



1. 2008. 2. 2006.

Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project.

Chart 10. Foreign body left in during procedure and mean number of secondary diagnoses



Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project.

Obstetric trauma – vaginal delivery with instrument***Obstetric trauma – vaginal delivery without instrument****Validity*

73. *Rationale and concerns.* These indicators are intended to flag cases of potentially preventable trauma (3rd and 4th degree perineal tears) during vaginal delivery. It is estimated that in about 11% of all deliveries perineal trauma associated with subsequent faecal incontinence occurs (Dudding *et al.*, 2008). Although risk factors as prolonged labour and no previous deliveries have been identified, precise recommendations for labour management in order to prevent obstetric lacerations remain unclear (Wheeler and Richter, 2007) resulting in an unclear preventability of these indicators. However, the percentage of deliveries involving higher degree lacerations is a useful monitor for the quality of obstetrical care and can assist in reducing the morbidity. Obstetric trauma indicators have been used by the US Joint Commission as well as by different international quality initiatives analysing obstetric data such as “BQS” the German statutory external quality assurance programme. As the risk of a perineal laceration is significantly increased in instrument-assisted labour (vacuum, forceps), rates for this patient population are reported separately.

74. *Content (consensual) validity.* Content validity was addressed in the technical report accompanying the original release of the AHRQ PSIs. Although panellists’ ratings of the “usefulness” of each candidate indicator were used to select the final PSI set, panellists were also asked to rate each indicator on its preventability and its likelihood of being due to medical error. The median rating of this indicator was 7, with agreement, on the former dimension, leading to a classification of “acceptable”. The median rating was 5, with disagreement, on the latter dimension, leading to a classification of “unclear”. Through similar processes, this PSI was endorsed by both the OECD Patient Safety Panel (Millar *et al.*, 2004) and the SimPatIE (Safety Improvement for Patients in Europe) project (Kristensen *et al.*, 2009). A 47-member Delphi panel convened by RAND rated this indicator “moderate” in importance, but lower on validity (i.e., median ratings 8 and 5, respectively) (Farley *et al.*, 2008). A similar indicator was initially implemented by the Joint Commission, which accredits health care organisations in the US, but subsequently withdrawn. The obstetric indicator was withdrawn because of clinical controversy about how often these lacerations are preventable, and whether use of the indicator could inadvertently promote caesarean delivery for questionable clinical indications. Empirically, we found no association between risk-adjusted laceration rates and caesarean rates at the hospital level (unpublished data) in California.

75. *Construct (convergent, predictive) validity.* This indicator rates moderately on predictive validity. Cases from the NIS that were flagged by this PSI had no excess mortality, but they did have 0.05-0.07 excess hospital days and up to \$220 in excess hospital charges, relative to carefully matched controls that were not flagged (Zhan and Miller, 2003). The reported differences in hospital length-of-stay and total charges were small, but statistically significant given the large number of events. A more recent replication using 2007 data, corrected for lacerations that were reported as “present on admission”, estimated 0.13-0.14 attributable hospital days and \$210-243 in attributable hospital costs for the average case (Foster *et al.*, 2009). Finally, a case control analysis from England estimated 0.48-0.56 days of excess hospitalisation (Raleigh *et al.*, 2008).

76. In a study testing construct validity using an implicit process measure of quality (Miller *et al.*, 2005), smoothed rates of this PSI among 2,116 hospitals surveyed by the Joint Commission in 1997-1999 were positively (counter intuitively) associated ($p=0.04$) with summary evaluation scores, but only in the subset of women with forceps or vacuum deliveries. Similarly, unadjusted data from England showed no association between this PSI and readmission (Bottle and Aylin, 2009).

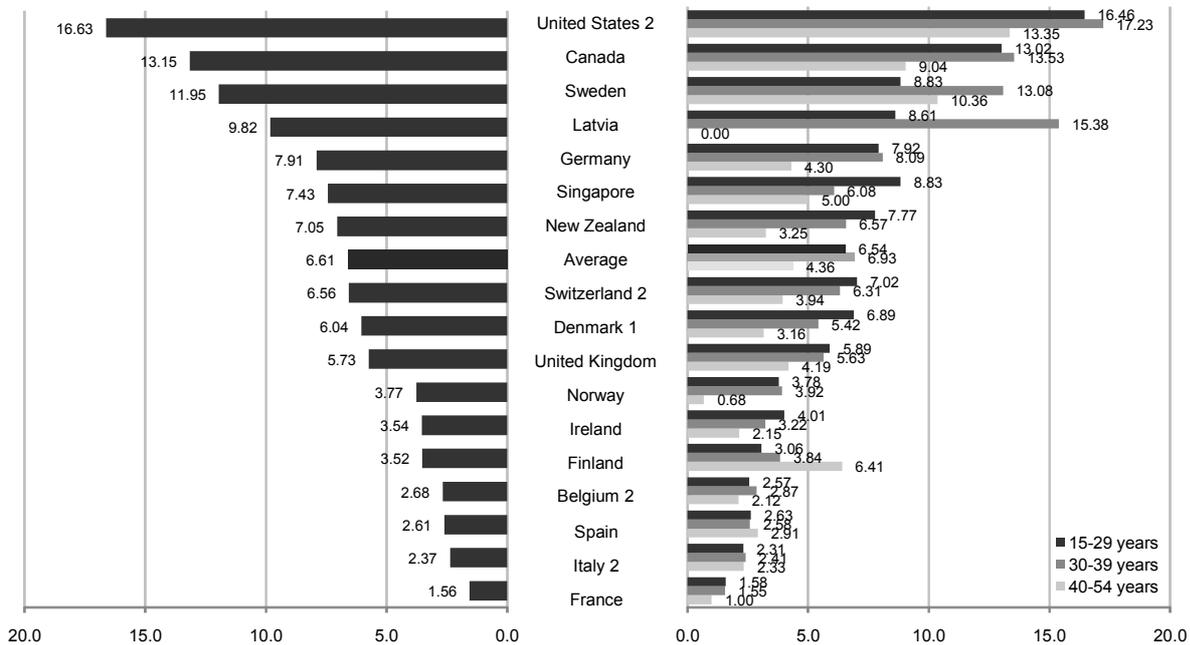
77. *Criterion validity.* No evidence about the criterion validity of this indicator was available before its release. This indicator is not likely to have a significant problem due to missing data about timing, because the indicator is inherently limited to women who have an in-hospital delivery. The best data about criterion validity come from the California Obstetric Validation Study (Romano *et al.*, 2005), which involved a stratified random cluster sample of 1 662 records from 52 hospitals (51% vaginal), of which over 97% were reviewed by an “expert” coder and obstetric nurse abstractor. This PSI demonstrated a sensitivity of 90% (95% CI, 82-96%) and a PPV of 90-95%; adjusting for the complex stratified sampling design increased the sensitivity to 93% (95% CI, 82-97%) but decreased the PPV to 73%. A subsequent study based on a clinical research data set with 393 indicator-positive (3rd/4th degree tears) and 383 indicator-negative vaginal deliveries (Brubaker *et al.*, 2007) reported a sensitivity of 77% (95% CI, 72-81%) and a specificity of 99.7% (95% CI, 98.5-99.4%). PPV could not be estimated due to the sampling design, but should be approximately 93% given a typical prevalence of 5%. A similar review of medical records of 955 cases from 18 English NHS (National Health Service) trusts found that none of the flagged events was present at admission and 15% were miscoded, leaving 85% that were confirmed (Bottle and Aylin, 2008).

78. In summary, recent evidence on construct validity of this indicator is inconclusive, but the evidence on criterion validity is quite supportive. Risk adjustment for maternal age and comorbidity is recommended for these indicators (AHRQ, 2007; Grobman *et al.*, 2006). To take the individual birth weight of the newborn into account when calculating these indicators would be a desirable risk adjustment approach. As maternal and neonatal data are stored separately in many administrative hospital databases, this methodological improvement is not feasible in most countries.

Findings

79. The rate of obstetric trauma after vaginal delivery with instrument shows rather high variability among countries (Chart 11). Reported rates vary from below 3% (France, Italy, Spain, Belgium) to more than 10% (Sweden, Canada, USA), and the rate of the USA is ten-fold higher than the rate of France. Rates of obstetric trauma after vaginal delivery without instrument (Chart 12) range from 0.2% to 3.8 %. Charts 11 and 12 display corresponding rates by age categories. In most countries, obstetric trauma rates are lower at age 40 years and older, probably due to the fact that in higher age groups, the risk factor of a first delivery is not as frequent as in younger age groups. As two countries (Finland and Sweden) use probably more reliable register data instead of administrative hospital data to report these indicators, validity concerns regarding the obstetric indicators are not as marked as compared to the other patient safety indicators. Direct adjustment for 5-year age-gender strata did not materially affect indicator rates across countries, although two countries moved up one rank (and two others moved down one rank to compensate). Rates of these obstetric indicators were not associated with the mean number of secondary diagnoses at the country level ($R^2 < 2\%$, $p > 0.5$). Ongoing projects on comparisons of obstetric trauma rates drawn from different data sources will provide information on the validity of these indicators in the near future. The rates represented in the charts below have not been age standardised or adjusted for secondary diagnoses or length of stay.

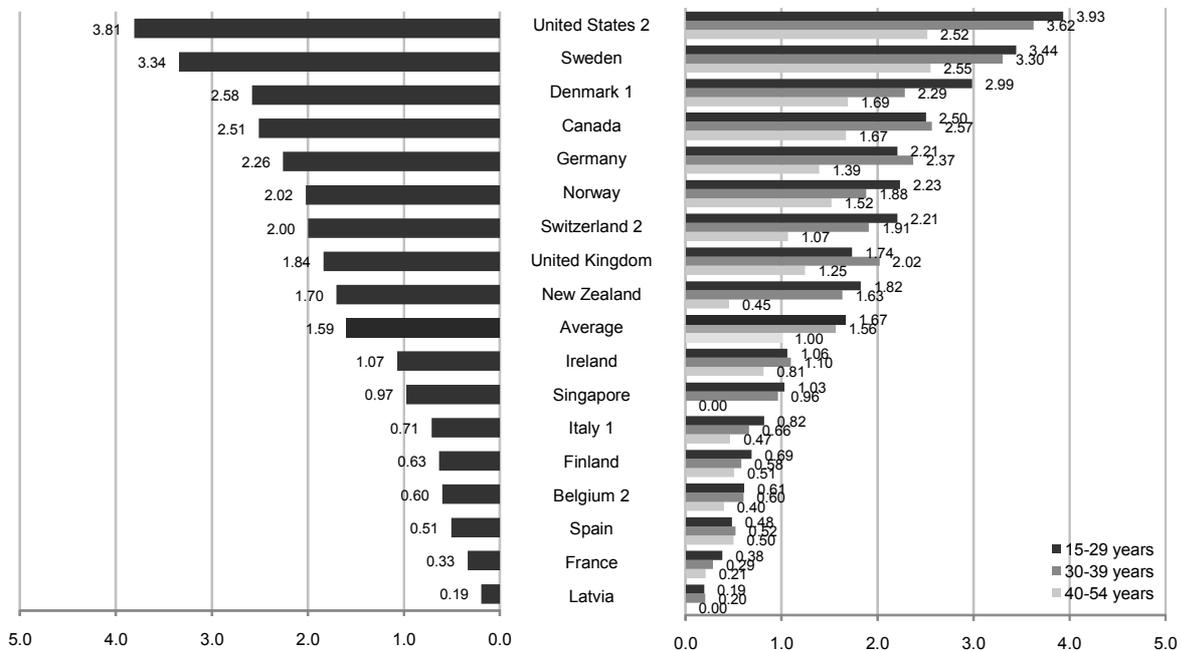
Chart 11. Obstetric trauma after vaginal delivery with instrument



1. 2008. 2. 2006.

Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project.

Chart 12. Obstetric trauma after vaginal delivery without instrument



1. 2008. 2. 2006.

Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project.

Detailed results

80. Table 6 depicts all reported PSI rates from each country. Empty cells represent missing rates. In one country (Norway), rates of zero (no numerator cases) occurred. PSI rates of zero are suspicious and might reflect under-coding. Table 7 shows minimal and maximal rates for each PSI across all reporting countries, replacing zero rates with the lowest nonzero rate of another country in Table 6 to permit calculation of ratios between the highest and lowest PSI rate. Portugal and Norway contributed most of the minimal values. The ratios between the highest and the lowest rates, representing inter-country variability, were rather high for some PSIs; for example, the ratio was 106 for PSI 7 (catheter-related bloodstream infection). The least variability was found for PSI 5 (foreign body left in during procedure) and obstetric trauma - vaginal delivery with instrument (9.8). Regarding PSI 13 (postoperative sepsis), the reported rates varied between 0.1 and 8.8%.

Table 6. PSI rates (%) reported by 19 countries in 2009

	Foreign body left in during procedure (PSI 5)	Catheter-related bloodstream infection (PSI 7)	Post operative pulmonary embolism or deep vein thrombosis (PSI 12)	Post operative sepsis (PSI 13)	Accidental Puncture or Laceration (PSI 15)	Obstetric trauma – vaginal delivery with instrument (PSI 18)	Obstetric trauma – vaginal delivery without instrument (PSI 19)
Belgium	0.007	0.178	0.470	1.717	0.281	2.682	0.602
Canada	0.007	0.083	0.199	0.398	0.402	13.147	2.511
Denmark	0.002	0.033	0.271	0.240	0.045	6.043	2.582
Finland	0.002	0.004	0.128	0.264	0.033	3.521	0.633
France						1.556	0.338
Germany	0.005	0.128	0.587	0.391	0.120	7.905	2.257
Iceland							6.681
Ireland	0.003	0.064	0.575	0.650	0.077	3.537	1.070
Italy	0.002	0.006	0.220	0.144	0.013	2.370	0.710
Latvia						9.821	0.193
New Zealand	0.009	0.425	0.819	0.505	0.356	7.055	1.704
Norway	0.000	0.026	0.438	8.081	0.000	3.768	2.021
Portugal		0.057	0.108	1.493	0.116	1.698	0.632
Singapore	0.006	0.123		0.140		7.434	0.975
Spain	0.005	0.189	0.281	0.440	0.183	2.611	0.505
Sweden	0.002	0.013	0.170	0.474	0.085	11.950	3.342
Switzerland	0.011	0.056	0.399	0.226	0.285	6.564	1.997
United Kingdom	0.005	0.090	0.754	0.276	0.129	5.731	1.837
United States	0.009	0.205	1.450	1.349	0.362	16.626	3.811

Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project. Minimal and maximal rates printed in bold.

Table 7. Minimum and maximum PSI rates (%) reported by all 18 countries

	N (countries)	Min Rate (%)	Max Rate (%)	Ratio between highest and lowest PSI rate, 16 countries
Foreign body left in during procedure (PSI 5)	15	0.002	0.011	6.0
Catheter-related bloodstream infection (PSI 7)	16	0.004	0.425	106.25
Postoperative pulmonary embolism (PE) or deep vein thrombosis (DVT) (PSI 12)	15	0.108	1.450	13.50
Postoperative sepsis (PSI 13)	16	0.140	8.081	57.73
Accidental Puncture or Laceration (PSI 15)	15	0.013	0.402	31.15
Obstetric trauma - vaginal delivery with instrument (PSI 18)	17	1.556	16.626	10.69
Obstetric trauma - vaginal delivery without instrument (PSI 19)	18	0.193	6.681	34.62

Note: Rates based on hospital administrative databases as reported by countries for the research and development work of the HCQI project.

Administrative reporting

Age and sex adjustment

81. Direct adjustment for age and gender did not materially affect the rates across countries for most indicators.

Secondary diagnoses

82. Thirteen countries reported the average number of secondary diagnoses calculated from the denominator-eligible cases for each PSI; Table 8 shows the corresponding data.

Table 8. Reported averages of secondary diagnoses per PSI

	Foreign body left in during procedure (PSI 5)	Catheter-related bloodstream infection (PSI 7)	Post operative pulmonary embolism (PE) or deep vein thrombosis (DVT) (PSI 12)	Post operative sepsis (PSI 13)	Accidental Puncture or Laceration (PSI 15)	Obstetric trauma – vaginal delivery with instrument (PSI 18)	Obstetric trauma – vaginal delivery w/o instrument (PSI 19)
Belgium	6.72	6.61	5.63	6.66	7.02	4.89	3.75
Canada	3.11	3.00	3.00	2.60	3.24	3.59	2.00
Denmark	2.41	2.55	2.26	1.96	2.26	6.88	6.62
Germany	5.31	4.98	4.81	4.24	5.43	5.30	3.32
Ireland	2.57	2.61	2.68	1.88	2.77	2.67	1.83
Italy	1.50		0.99	0.87	1.54	1.79	1.50
New Zealand	3.93	4.37	4.84	4.25	4.19	3.77	2.20
Norway	1.90	2.08	1.76	7.00	1.90	2.30	2.00
Portugal		2.50	1.30	2.50	0.10		
Singapore	3.89	3.63		2.65		3.85	2.21
Spain	3.71	3.74	2.77	2.48	3.45	3.32	2.79
Sweden	2.50	2.50	3.79	2.60	2.50	2.60	1.40
Switzerland	3.03	2.74	2.25	2.80	3.03	4.07	3.09
United Kingdom ¹	2.72	2.83	3.18	2.43	2.89	2.83	1.90
United States	6.02	5.85	6.20	6.20	6.50	3.77	2.83

1. Data for the United Kingdom refers to England only.

83. There is currently no valid information on the mean number of secondary diagnoses from Finland, Portugal and Sweden. The United Kingdom reported this information only for England, which contributed most of the cases to the denominator. Age-sex standardised PSI rates were significantly correlated with the mean number of secondary diagnoses across countries for all of the non-obstetric indicators. Among 12 countries that reported the mean number of secondary diagnoses for PSI 15 (accidental puncture or laceration), the Spearman rank correlation coefficient was 0.706 ($p=0.010$). Among 13 countries that reported the mean number of secondary diagnoses for PSI 5 (foreign body left in during procedure), the rank correlation was 0.643 ($p=0.018$). Among 12 countries that reported the mean number of secondary diagnoses for PSI 12 (postoperative PE or DVT), the rank correlation was 0.657 ($p=0.020$). Among 13 countries that reported the mean number of secondary diagnoses for PSI 13 (postoperative sepsis), the rank correlation was 0.643 ($p=0.018$). Among 13 countries that reported the mean number of secondary diagnoses for PSI 7 (catheter-related bloodstream infection), the rank correlation was 0.918 ($p<0.0001$). The corresponding R-squared statistics (proportion of variance explained) ranged from 28% to 52%. Italy, followed by Norway (except PSI 13), reported the lowest mean number of secondary diagnoses among eligible cases, which explains some rather low PSI rates reported by these countries.

84. Based on the consistent country-level association between the mean number of secondary diagnoses (among denominator cases) and rates of all non-obstetric PSIs, it is possible to adjust country-

specific PSI rates for variation in coding intensity. Specifically, an ordinary least squares unweighted regression model was estimated for each PSI using data from all countries that reported both the mean number of secondary diagnoses and the corresponding rate for that PSI. Data from Norway were excluded from the analyses of accidental puncture or laceration, foreign body, and postoperative sepsis, because Norway reported implausibly low (zero) or high rates on these three indicators. The outcome variable in these regression models was a country's age-sex standardised PSI rate; the predictor variable was the mean number of secondary diagnoses among denominator cases. Parameter estimates from these models were used to estimate country-specific residuals, which were then linearly transformed into adjusted PSI rates with the same mean value as the unadjusted but standardised rates. The resulting adjusted PSI rates demonstrate far less variation (measured by the range, maximum-minimum ratio, standard deviation, and coefficient of variation) than unadjusted rates, for all non-obstetric PSIs. Based on the R^2 statistic, 28% to 52% of the observed variation at the country level in PSI rates is attributable to variation in diagnostic coding. Removing this portion of the variation appears to yield more valid country-specific PSI rates.

Length of stay

85. Mean length of stay data among denominator-eligible cases for each indicator were reported by 14 countries. In no case was there a statistically significant association between mean length of stay and the age-sex standardised PSI rate. This finding suggests that international variation in length of stay does not meaningfully contribute to the variation in PSI rates, probably because all participating countries attempted to remove same-day hospital cases in the 2009 cycle of data submission

Year to year reliability

86. To estimate divergence across years, PSI rates reported in the current calculation cycle were compared to rates reported in the 2008 OECD cycle. Most indicators did not significantly change in definition and were used for this analysis. Thirteen of seventeen countries participating in 2008 also participated in 2009 and completely recalculated their PSI rates: Belgium, Canada, Denmark, Finland, Germany, Italy, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, the United Kingdom, and the United States.

87. To assess the reliability of country-specific PSI rates across years, correlation analyses were performed using the 2008 and 2009 rates reported by each country for each PSI. The correlation analyses show high reliability both within countries (Table 9) and within PSIs (Table 10).

88. The year-to-year reliability of PSI rates, based on Pearson's correlation coefficient, was greater than $r=0.94$ ($p \leq 0.01$) for all countries except Portugal. However, the correlation coefficient between Portuguese and US data improved, suggesting that year-to-year reliability was poor simply because of an improvement of the Portuguese data or calculation methods. The correlation coefficient is slightly lower for the United Kingdom than for other countries, probably due to the fact that only data from England were included in last year's calculation.

Table 9. Pearson's correlation coefficients of PSI rates between 2008 and 2009, by participating country

	n (number of PSI rates [%])	Pearson's correlation coefficient	Year of data collection (2008 calculation)	Year of data collection (2009 calculation)
Belgium	7	0.999	2004	2006
Canada	7	1	2006-2007	2007
Denmark	7	1	2007	2008
Finland	7	0.997	2006	2007
Germany	7	0.999	2006	2007
Italy	7	1	2005	2007
New Zealand	7	0.998	2006-2007	2007
Norway	7	0.941	2006	2007
Portugal	6	0.546	2006	2007
Singapore	5	1 ²	2007	2007
Spain	7	0.996	2006	2007
Sweden	7	0.999	2006	2007
United Kingdom	7	0.994	2006-2007	2007
United States	7	1 ¹	2006	2006

1. The year-to-year correlation for the USA is close to one, because 2007 data were not yet available from all participating states at the time of this data submission.

2. Singapore used data from the same year of data collection in both calculations and shows a year-to-year correlation close to one.

Table 10. Pearson's correlation coefficients of PSI rates between 2008 and 2009 calculations, by PSI

Indicator name	n (countries)	Pearson's correlation coefficient (p≤0.01)
Foreign body left in during procedure (PSI 5)	13	0.920
Catheter-related bloodstream infection (PSI 7)	14	0.994
Postoperative pulmonary embolism (PE) or deep vein thrombosis (DVT) (PSI 12)	13	0.821
Postoperative sepsis (PSI 13)	14	0.970
Accidental Puncture or Laceration (PSI 15)	13	0.916
Obstetric trauma – vaginal delivery with instrument (PSI 18)	14	0.996
Obstetric trauma – vaginal delivery without instrument (PSI 19)	14	0.981

89. The year-to-year reliability of PSI 12 (postoperative pulmonary embolism or deep vein thrombosis) was 0.892 ($p \leq 0.01$), which was less than that of the other indicators listed in Table 10.

Conclusions

90. All participating countries performed PSI calculations and contributed data; however, validity concerns arose in the analysis of data from several countries, as discussed above. Criteria to ensure data quality proposed in the last year's project are partially applied: rates of indicators of zero are regarded as invalid (except for the sentinel event indicator transfusion reaction which was not investigated this year) and should not be publicly reported.

91. As several countries did not report average counts of secondary diagnoses, it was not feasible to implement a quality criterion based on the “mean number of secondary diagnoses among denominator cases”. As the PSI methodology is based on analyses of secondary diagnoses reported in administrative hospital databases, thorough reporting of the average number of secondary diagnoses together with the corresponding PSI rate will be required in future projects envisaging public reporting.

92. The test-retest analyses suggest excellent reliability for eleven of twelve countries participating in 2008 and 2009 and using data from different collection periods. One country (Portugal) showed poorer year-to-year reliability, probably due to reported difficulties in the data collection and analysis during 2008, with subsequent improvement this year. Similarly, there was good test-to-retest reliability for all tested PSIs except for the postoperative pulmonary embolism or deep vein thrombosis indicator.

93. Standardisation of PSI rates using relatively narrow, five-year age strata proved that international variation in hospitalisation practices and population age structure explains very little of the observed variation in PSI rates. For all of the non obstetric PSIs, the consistent relationship between a country’s standardised PSI rate and its mean number of secondary diagnoses is a cause for concern. This finding suggests that some countries have more thorough reporting of secondary diagnoses than others, and that variation in reporting may account for 28-52% of the observed variation in PSI rates. Further analyses are underway to model the impact of underreporting on PSI rates, and to impute corrected PSI rates after adjusting for underreporting. Underreporting does not appear to be an issue for postoperative DVT/PE and obstetric trauma, based on both published literature and empirical analyses of OECD data, although overdiagnosis due to widespread ultrasound screening in some hospitals or areas is an emerging concern for the former PSI.

94. Research and development work on patient safety indicators over the past four years has provided encouraging evidence that system-level patient safety indicators can be calculated from administrative databases by a majority of OECD countries. Although significant advancement in methodology and international comparability has been achieved during this period the requirement remains for ongoing research and development. During 2010-2011, work in this domain will focus on the technical specifications of the existing indicators, impact of different coding practices and classification systems on indicator rates and the potential use of data adjustment methods.

REFERENCES

- AHRQ (Agency for Health Care Research and Quality) (2007), PSI Technical Specifications. Version 3.1, Revision 3 http://www.qualityindicators.ahrq.gov/archives/psi/psi_technical_specs_v30a.pdf.
- Angus DC, Linde-Zwirble WT, Lidicker J, Clermont G, Carcillo J, Pinsky MR. "Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care." *Crit Care Med.* 2001 Jul;29(7):1303-10.
- Bahl V, Thompson MA, Kau TY, Hu HM, Campbell DA. (2008), "Do the AHRQ Patient Safety Indicators flag conditions that are present at the time of hospital admission?" *Medical Care*, Vol. 46, No. 5, pp. 516-522.
- Barbour GL. (1993), "Usefulness of a discharge diagnosis of sepsis in detecting iatrogenic infection and quality of care problems", *American Journal of Medical Quality*, Vol. 8, No. 1, pp. 2-5.
- Belío-Blasco C, Torres-Fernández-Gil MA, Echeverría-Echarri JL, Gómez-López LI. (2000), "Evaluation of two retrospective active surveillance methods for the detection of nosocomial infection in surgical patients", *Infection Control and Hospital Epidemiology*, Vol. 21, No. 1, pp 24-27.
- Best WR, Khuri SF, Phelan M, Hur K, Henderson WG, Demakis JG, Daley J. (2002), "Identifying patient preoperative risk factors and postoperative adverse events in administrative databases: results from the Department of Veterans Affairs National Surgical Quality Improvement Program", *Journal of the American College of Surgeons*, Vol. 194, No. 3, pp.257-266.
- Borzecki A, Kaafarani H, Hanchate A, Loveland S, Mull H, Shin M, Rosen A. (2009), "Validating the Patient Safety Indicators (PSI) in the Veterans Health Administration", abstract presented at AcademyHealth Annual Research Meeting, <http://www.academyhealth.org/files/2009/tuesday/borzecki.pdf>
- Bottle A, Aylin P. (2008), "How NHS trusts could use patient safety indicators to help improve care", *HealthCareRiskReport*, May 2008, pp. 12-14.
- Bottle A, Aylin P. (2009), "Application of AHRQ Patient Safety Indicators to English hospital data", *Quality and Safety in Health Care*, Vol. 18, pp. 303-308.
- Brubaker L, Bradley CS, Handa VL, Richter HE, Visco A, Brown MB, Weber AM. (2007), "Anal sphincter laceration at vaginal delivery: is this event coded accurately?", *Obstetrics and Gynecology*, Vol. 109, No.5, pp.1141-1145.
- Culler SD, Hawley JN, Naylor V, Rask KJ. (2007), "Is the availability of hospital IT applications associated with a hospital's risk adjusted incidence rate for Patient Safety Indicators: Results from 66 Georgia hospitals", *Journal of Medical Systems*, Vol. 31, pp. 319-327.
- Drösler SE, (2008), "Facilitating Cross-National Comparisons of Indicators for Patient Safety at the Health-System Level in OECD Countries", *OECD Health Technical Papers*, No. 19, OECD, Paris.

- Drösler SE, Klazinga NS, Romano PS, Tancredi DJ, Gogorcena Aoiz MA, Hewitt MC, Scobie S, Soop M, Wen E, Quan H, Ghali WA, Mattke S, Kelley E., (2009), "Application of patient safety indicators internationally: a pilot study among seven countries", *International Journal for Quality in Health Care*, Vol.21, No.4, pp.272-278.
- Dudding TC, Vaizey CJ, Kamm MA. (2008), "Obstetric anal sphincter injury: incidence, risk factors, and management", *Annals of Surgery*, Vol. 247, No.2, pp.224-237.
- Eggimann P, Harbarth S, Constantin MN, Touveneau S, Chevrolet JC, Pittet D. (2000), "Impact of a prevention strategy targeted at vascular-access care on incidence of infections acquired in intensive care", *Lancet*, Vol. 355, No.9218, pp.1864-1868.
- Encinosa WE, Hellinger FJ. (2008), "The impact of medical errors on ninety-day costs and outcomes: An examination of surgical patients", *Health Services Research*, Vol. 43, No. 6, pp. 2067-2085.
- Farley DO, Greenberg MD, Haviland AM, Lovejoy S. (2008), "Prioritizing Patient Safety Outcomes Measures: Results of an Expert Consensus Process", prepared for the Agency for Healthcare Research and Quality, http://192.5.14.110/pubs/working_papers/2008/RAND_WR601.pdf.
- Foster D, Young J, Heller S. (2009), "US national estimates of mortality, length of stay, and costs attributable to inpatient complications of care", abstract presented at AcademyHealth 2009 Annual Research Meeting (<http://www.academyhealth.org/files/arm/ARM-2009-Posters.pdf>).
- Friedman B, Encinosa W, Jiang HJ, Mutter R. (2009), "Do patient safety events increase readmissions?" *Medical Care*, Vol. 47, No. 5, pp. 583-590.
- Gallagher B, Cen L, Hannan EL. (2005a), "Readmissions for Selected Infections Due to Medical Care: Expanding the definition of a Patient Safety Indicator", In "Advances in Patient Safety: from Research to Implementation", Agency for Health Care Research and Quality 2005.
- Gallagher B, Cen L, Hannan EL. (2005b), "Validation of AHRQ's Patient Safety Indicator for Accidental Puncture or Laceration." In "Advances in Patient Safety: from Research to Implementation", Agency for Health Care Research and Quality 2005.
- Gawande AA, Studdert DM, Orav EJ, Brennan TA, Zinner MJ. (2003), "Risk factors for retained instruments and sponges after surgery", *New England Journal of Medicine*, Vol. 348, No.3, pp.229-235.
- Gonzalez-Ojeda A, Rodriguez-Alcantar DA, Arenas-Marquez H, Sanchez Perez-Verdia E, Chavez-Perez R, Alvarez-Quintero R, Perea-Sanchez A. (1999), "Retained foreign bodies following intra-abdominal surgery", *Hepatogastroenterology*, Vol.46, No.26, pp.808-812.
- Geerts WH, Pineo GF, Heit JA, Bergqvist D, Lassen MR, Colwell CW, Ray JG. (2004), "Prevention of venous thromboembolism: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy", *Chest*, Vol.126, No.3 Suppl., pp.338S-400S.
- Gillet P, Kolh P, Sermeus W, Vleugels A, Jacques J, Van Den Heede K, Devriese S, Vrijens F, Verelst S. (2008), "Détection des événements indésirables dans les bases de données administratives", KCE reports 93B, http://www.kce.fgov.be/index_fr.aspx?SGREF=3439&CREF=11889.

- Glance LG, Li Y, Osler TM, Mukamel DB, Dick AW. (2008), "Impact of date stamping on patient safety measurement in patients undergoing CABG: Experience with the AHRQ Patient Safety Indicators", *BMC Health Services Research*, Vol. 8, pp. 176.
- Grobman WA, Feinglass J, Murthy S. (2006), "Are the Agency for Healthcare Research and Quality obstetric trauma indicators valid measures of hospital safety?" *American Journal of Obstetrics and Gynecology*, Vol. 195, pp. 868-874.
- Haut ER, Noll K, Efron DT, Berenholz SM, Haider A, Cornwell EE, Pronovost PJ. (2007), "Can increased incidence of deep vein thrombosis (DVT) be used as a marker of quality of care in the absence of standardized screening? The potential effect of surveillance bias on reported DVT rates after trauma", *Journal of Trauma*, Vol. 63, No. 5, pp. 1132-1137.
- Hawker GA, Coyte PC, Wright JG, Paul JE, Bombardier C. (1997), "Accuracy of administrative data for assessing outcomes after knee replacement surgery", *Journal of Clinical Epidemiology*, Vol.50, No.3, pp.265-273.
- Henderson KE, Recktenwald AJ, Reichley RM, Bailey TC, Waterman BM, Diekemper RL, Storey PE, Ireland BK, Dunagan WC. (2009), "Clinical validation of the AHRQ Postoperative Venous Thromboembolism Patient Safety Indicator", *Joint Commission Journal on Quality and Patient Safety*, Vol. 35, No. 7, pp. 370-376.
- Houchens RL, Elixhauser A, Romano PS. (2008), "How often are potential patient safety events present on admission?" *Joint Commission Journal on Quality and Patient Safety*, Vol.34, No.3, pp.154-163.
- Iezzoni LI, Daley J, Heeren T, Foley SM, Hughes JS, Fisher ES, Duncan CC, Coffman GA. (1994a), "Using administrative data to screen hospitals for high complication rates", *Inquiry*, Vol.31, No.1, pp.40-55.
- Iezzoni LI, Shwartz M, Ash AS, Mackiernan Y, Hotchkin EK. (1994b), "Risk adjustment methods can affect perceptions of outcomes", *American Journal of Medical Quality*, Vol.9, No.2, pp.43-48.
- Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, Phillips RS, Banks NJ, Davis DT Jr. (1999), "Does the Complications Screening Program flag cases with process of care problems? Using explicit criteria to judge processes", *International Journal of Quality in Health Care*, Vol.11, No.2, pp.107-118.
- Isaac T, Jha AK. (2008), "Are Patient Safety Indicators related to widely used measures of hospital quality?" *Journal of General Internal Medicine*, Vol. 23, No. 9, pp. 1373-1378.
- Keeler EB, Kahn KL, Bentow SS. (1992), *Assessing quality of care for hospitalized Medicare patients with hip fracture using coded diagnoses from the Medicare Provider Analysis and Review File (Prepared for the Health Care Financing Administration, US Department HHS) RAND 1992*
- Kovner C, Gergen PJ. (1998), "Nurse staffing levels and adverse events following surgery in U.S. hospitals", *Image - the Journal of Nursing Scholarship*, Vol.30, No.4, pp.315-321.
- Kristensen S, Mainz J, Bartels P. (2009), "Selection of indicators for continuous monitoring of patient safety: recommendations of the project 'Safety Improvement for Patients in Europe'", *International Journal for Quality in Health Care*, Vol. 21, No. 3, pp. 169-175.

- Kronman MP, Hall M, Slonim AD, Shah SS. (2008), "Charges and length of stay attributable to adverse patient-care events using pediatric-specific quality indicators: A multicenter study of freestanding children's hospitals", *Pediatrics*, Vol. 121, No. 6, pp. e1653-e1659.
- Lawthers AG, McCarthy EP, Davis RB, Peterson LE, Palmer RH, Iezzoni LI. (2000), "Identification of in-hospital complications from claims data. Is it valid?", *Medical Care*, Vol.38, No.8, pp785-795.
- Massanari RM, Wilkerson K, Streed SA, Hierholzer WJ Jr. (1987), "Reliability of reporting nosocomial infections in the discharge abstract and implications for receipt of revenues under prospective reimbursement", *American Journal of Public Health*, Vol.77, No.5, pp.561-564.
- McCarthy EP, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, Phillips RS, Davies DT Jr. (2000), "Does clinical evidence support ICD-9-CM diagnosis coding of complications?", *Medical Care*, Vol.38, No.8, pp.868-876.
- Mattke S, et al. (2006), "Health Care Quality Indicators Project Initial Indicators Report", OECD Health Working Papers, No. 22, OECD, Paris
- Menachemi N, Saunders C, Chukmaitov A, Matthews MC, Brooks RG. (2007), "Hospital adoption of information technologies and improved patient safety: A study of 98 hospitals in Florida", *Journal of Healthcare Management*, Vol. 52, No. 6, pp. 398-409.
- Millar J, Mattke S and the members of the OECD Patient Safety Panel (2004), "Selecting Indicators for Patient Safety at the Health Systems Level in OECD Countries", OECD Health Technical Papers, No. 18, OECD, Paris.
- Miller MR, Elixhauser A, Zhan C, Meyer GS. (2001), "Patient Safety Indicators: using administrative data to identify potential patient safety concerns", *Health Services Research*, Vol.36, No.6 (part 2), pp.110-132.
- Miller MR, Pronovost P, Donithan M, Zeger S, Zhan C, Morlock L, Meyer GS. (2005), "Relationship between performance measurement and accreditation: implications for quality of care and patient safety", *American Journal of Medical Quality*, Vol.20, No.5, pp.239-252.
- Naessens JM, Campbell CR, Berg B, Williams AR, Culbertson R. (2007), "Impact of diagnosis-timing indicators on measures of safety, comorbidity, and case mix groupings from administrative data sources", *Medical Care*, Vol.45, No.12, pp.1234.
- Needleman J, Buerhaus P, Mattke S, Stewart M, Zelevinsky K. (2002), "Nurse-staffing levels and the quality of care in hospitals", *New England Journal of Medicine*, Vol.346, No.22, pp.1715-1722.
- OECD Health Data 2008, June 2008, Frequently Requested Data.
- Olin JW. (2002), "Pulmonary embolism", *Reviews in Cardiovascular Medicine*, Vol.3, No.Suppl 2, pp.S68-75.
- Parente ST, McCullough JS. (2009), "Health information technology and patient safety: Evidence from panel data", *Health Affairs*, Vol. 28, No. 2, pp. 357-360.
- Quan, H et al.(2008) for the IMECCHI Investigators. Adaptation of AHRQ Patient Safety Indicators for Use in ICD-10 Administrative Data by an International Consortium. 2008.
http://www.ahrq.gov/downloads/pub/advances2/vol1/advances-quan_52.pdf.

- Raleigh VS, Cooper J, Bremner SA, Scobie S. (2008), "Patient safety indicators for England from hospital administrative data: case-control analysis and comparison with US data", *BMJ*, Vol.337, No.a, pp.1702.
- Rivard PE, Luther SL, Christiansen CL, Shibe Zhao, Loveland S, Elixhauser A, Romano PS, Rosen AK. (2008), "Using patient safety indicators to estimate the impact of potential adverse events on outcomes", *Medical Care Research and Review*, Vol.65, No.1, pp.67-87.
- Romano PS, Schembri ME, Rainwater JA. (2002), "Can administrative data be used to ascertain clinically significant postoperative complications?", *American Journal of Medical Quality*, Vol. 17, No.4, pp.145-154.
- Romano PS, Yasmeen S, Schembri ME, Keyzer JM, Gilbert WM. (2005), "Coding of perineal lacerations and other complications of obstetric care in hospital discharge data", *Obstetrics and Gynecology*, Vol.106, No.4, pp.717-725.
- Romano PS, Mull HJ, Rivard PE, Zhao S, Henderson WG, Loveland S, Tsilimingras D, Christiansen CL, Rosen AK. (2009), "Validity of selected AHRQ patient safety indicators based on VA National Surgical Quality Improvement Program data", *Health Services Research*, Vol.44, No.1, pp.182-204.
- Rosen AK, Loveland SA, Romano PS, Itani KMF, Silber JH, Even-Shoshan OO, Halenar MJ, Teng Y, Zhu J, Volpp KG. (2009), "Effects of resident duty hour reform on surgical and procedural Patient Safety Indicators among hospitalized Veterans Health Administration and Medicare patients", *Medical Care*, Vol. 47, No. 7, pp. 723-731.
- Scanlon MC, Miller M, Harris JM, Schulz K, Sedman A. (2006), "Targeted chart review of pediatric patient safety events identified by the Agency for Healthcare Research and Quality's Patient Safety Indicators methodology", *Journal of Patient Safety*, Vol. 2, No. 4, pp. 191-197.
- Scanlon MC, Harris JM, Levy F, Sedman A. (2008), "Evaluation of the Agency for Healthcare Research and Quality Pediatric Quality Indicators", *Pediatrics*, Vol. 121, No. 6, pp. e1723-1731.
- Sedman A, Harris JM 2nd, Schulz K, Schwalenstocker E, Remus D, Scanlon M, Bahl V. (2005), "Relevance of the Agency for Healthcare Research and Quality Patient Safety Indicators for children's hospitals", *Pediatrics*, Vol.115, No.1, pp.135-145.
- Stevanovic V. (2009), "Technical Analysis of the Validity and Comparability of the Patient Safety Indicators: Impact of the AHRQ Exclusions", presented at the Patient Safety Experts Subgroup of the HCQI Project, OECD, Paris, 23 October.
- Stone PW, Horan TC, Shih HC, Mooney-Kane C, Larson E. (2007), "Comparisons of health care-associated infections identification using two mechanisms for public reporting", *American Journal of Infection Control*, Vol. 35, No. 3, pp. 145-149.
- Taylor B. (1998), "Common bile duct injury during laparoscopic cholecystectomy in Ontario: does ICD-9 coding indicate true incidence?", *CMAJ*, Vol.158, No.4, pp.481-485.
- Thornlow DK, Merwin E. (2009), "Managing to improve quality: The relationship between accreditation standards, safety practices, and patient outcomes", *Health Care Management Review*, Vol. 34, No. 3, pp. 267-272.

- Utter GH, Zrelak PA, Baron R, Tancredi DJ, Sadeghi B, Geppert JJ, Romano PS. (2009), "Positive predictive value of the AHRQ Accidental Puncture or Laceration Patient Safety Indicator", *Medical Care*, epub ahead of print.
- Valinsky LJ, Hockey RL, Hobbs MS, Fletcher DR, Pikora TJ, Parsons RW, Tan P. (1999), "Finding bile duct injuries using record linkage: a validated study of complications following cholecystectomy", *Journal of Clinical Epidemiology*, Vol.52, No.9, pp.893-901.
- Vartak S, Ward MM, Vaughn TE. (2008), "Do postoperative complications vary by hospital teaching status", *Medical Care*, Vol. 46, No. 1, pp. 25-32.
- Weingart SN, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, Phillips RS, Davies DT Jr, Banks NJ. (2000), "Use of administrative data to find substandard care: validation of the complications screening program", *Medical Care*, Vol.38, No.8, pp.796-806.
- Weller WE, Gallagher BK, Cen L, Hannan EL. (2004), "Readmissions for venous thromboembolism: Expanding the definition of Patient Safety Indicators", *Joint Commission Journal on Quality and Safety*, Vol. 30, No. 9, pp. 497-504.
- Wheeler TL 2nd, Richter HE. Delivery method, anal sphincter tears and fecal incontinence: new information on a persistent problem. *Curr Opin Obstet Gynecol*. 2007 Oct;19(5):474-9.
- White RH, Sadeghi B, Tancredi DJ, Zrelak P, Cuny J, Sama P, Utter GH, Geppert JJ, Romano PS. (2009), "How valid is the ICD-9-CM based AHRQ Patient Safety Indicator for Postoperative Venous Thromboembolism?" *Medical Care*, Vol. 47, epub ahead of print.
- WHO (2009), "WHO Guidelines on Hand Hygiene in Health Care: First Global Patient Safety Challenge - Clean Care is Safer Care", WHO, Geneva.
- Yao H, Greenberg MD, Haviland AM, Farley DO. (2009), "'Canary measures' among the AHRQ Patient Safety Indicators", *American Journal of Medical Quality*, Vol. 24, epub ahead of print.
- Zhan C, Miller MR. (2003), "Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization", *JAMA*, Vol.290, No.14, pp.1868-1874.
- Zhan C, Battles J, Chiang Y, Hunt D. (2007), "The validity of ICD-9-CM codes in identifying postoperative deep vein thrombosis and pulmonary embolism", *Joint Commission Journal on Quality and Patient Safety*, Vol. 33, No. 6, pp. 326-331.
- Zrelak PA, Sadeghi B, Utter GH, Baron R, Tancredi DJ, Geppert JJ, Romano PS. (2009), "Positive predictive value of the AHRQ Patient Safety Indicator for Central Line Associated Bloodstream Infection (Selected Infections Due to Medical Care)", under review.

OECD HEALTH WORKING PAPERS

A full list of the papers in this series can be found on the OECD website: www.oecd.org/els/health/workingpapers

- No. 47 *HEALTH CARE QUALITY INDICATORS PROJECT: PATIENT SAFETY INDICATORS REPORT 2009* (2009) Saskia Drösler, Patrick Romano, Lihan Wei
- No. 46 *EDUCATION AND OBESITY IN FOUR OECD COUNTRIES* (2009) Franco Sassi, Marion Devaux, Jody Church, Michele Cecchini and Francesca Borgonovi
- No. 45 *THE OBESITY EPIDEMIC: ANALYSIS OF PAST AND PROJECTED FUTURE TRENDS IN SELECTED OECD COUNTRIES* (2009) Franco Sassi, Marion Devaux, Michele Cecchini and Elena Rusticelli
- No. 44 *THE LONG-TERM CARE WORKFORCE: OVERVIEW AND STRATEGIES TO ADAPT SUPPLY TO A GROWING DEMAND* (2009) Rie Fujisawa and Francesca Colombo
- No. 43 *MEASURING DISPARITIES IN HEALTH STATUS AND IN ACCESS AND USE OF HEALTH CARE IN OECD COUNTRIES* (2009) Michael de Looer and Gaetan Lafortune
- No. 42 *POLICIES FOR HEALTHY AGEING: AN OVERVIEW* (2009) Howard Oxley
- No. 41 *THE REMUNERATION OF GENERAL PRACTITIONERS AND SPECIALISTS IN 14 OECD COUNTRIES: WHAT ARE THE FACTORS EXPLAINING VARIATIONS ACROSS COUNTRIES?* (2008) Rie Fujisawa and Gaetan Lafortune
- No. 40 *INTERNATIONAL MOBILITY OF HEALTH PROFESSIONALS AND HEALTH WORKFORCE MANAGEMENT IN CANADA: MYTHS AND REALITIES* (2008) Jean-Christophe Dumont, Pascal Zurn, Jody Church and Christine Le Thi
- No. 39 *PHARMACEUTICAL PRICING & REIMBURSEMENT POLICIES IN GERMANY* (2008) Valérie Paris and Elizabeth Docteur
- No. 38 *MIGRATION OF HEALTH WORKERS: THE UK PERSPECTIVE TO 2006* (2008) James Buchan, Susanna Baldwin and Miranda Munro
- No. 37 *THE US PHYSICIAN WORKFORCE: WHERE DO WE STAND?* (2008) Richard A. Cooper
- No. 36 *MIGRATION POLICIES OF HEALTH PROFESSIONALS IN FRANCE* (2008) Roland Cash and Philippe Ulmann
- No. 35 *NURSE WORKFORCE CHALLENGES IN THE UNITED STATES: IMPLICATIONS FOR POLICY* (2008) Linda H. Aiken and Robyn Cheung
- No. 34 *MISMATCHES IN THE FORMAL SECTOR, EXPANSION OF THE INFORMAL SECTOR: IMMIGRATION OF HEALTH PROFESSIONALS TO ITALY* (2008) Jonathan Chaloff
- No. 33 *HEALTH WORKFORCE AND INTERNATIONAL MIGRATION: CAN NEW ZEALAND COMPETE?* (2008) Pascal Zurn and Jean-Christophe Dumont
- No. 32 *THE PREVENTION OF LIFESTYLE-RELATED CHRONIC DISEASES: AN ECONOMIC FRAMEWORK* (2008) Franco Sassi and Jeremy Hurst

RECENT RELATED OECD PUBLICATIONS

HEALTH AT A GLANCE 2009: OECD INDICATORS (2009).

See <http://www.oecd.org/health/healthataglance> for more information

ACHIEVING BETTER VALUE FOR MONEY IN HEALTH CARE (2009), OECD HEALTH POLICY STUDIES

OECD HEALTH DATA 2009 (2009), available on CD-ROM (in single-user or network installations). The database can be queried in English, French, German, Spanish and Italian. Japanese and Russian are also available but exclusively in the online version. www.oecd.org/health/healthdata

OECD REVIEWS OF HEALTH SYSTEMS - TURKEY (2009)

THE LOOMING CRISIS IN THE HEALTH WORKFORCE: CAN OECD COUNTRIES RESPOND? (2008)

PHARMACEUTICAL PRICING POLICIES IN A GLOBAL MARKET (2008)

OECD REVIEWS OF HEALTH SYSTEMS - SWITZERLAND (2006)

LONG-TERM CARE FOR OLDER PEOPLE (2005), *OECD HEALTH PROJECT SERIES*

HEALTH TECHNOLOGIES AND DECISION MAKING (2005), *OECD HEALTH PROJECT SERIES*

OECD REVIEWS OF HEALTH CARE SYSTEMS - FINLAND (2005)

OECD REVIEWS OF HEALTH CARE SYSTEMS - MEXICO (2005)

PRIVATE HEALTH INSURANCE IN OECD COUNTRIES (2004), *OECD HEALTH PROJECT SERIES*

TOWARDS HIGH-PERFORMING HEALTH SYSTEMS - POLICY STUDIES (2004), *OECD HEALTH PROJECT SERIES*

TOWARDS HIGH-PERFORMING HEALTH SYSTEMS (2004), *OECD HEALTH PROJECT SERIES*

OECD REVIEWS OF HEALTH CARE SYSTEMS - KOREA (2003)

A DISEASE-BASED COMPARISON OF HEALTH SYSTEMS: WHAT IS BEST AND AT WHAT COST? (2003)

MEASURING UP: IMPROVING HEALTH SYSTEMS PERFORMANCE IN OECD COUNTRIES (2002)

A SYSTEM OF HEALTH ACCOUNTS (2000)

SHA Revision information and proposals at <http://www.oecd.org/health/sha/revision>

For a full list, consult the OECD On-Line Bookstore at www.oecd.org,
or write for a free written catalogue to the following address:

OECD Publications Service
2, rue André-Pascal, 75775 PARIS CEDEX 16