International Journal for Quality in Health Care 2014; **Volume 26, Number 1**: pp. 16–25
Advance Access Publication: 13 December 2013
10.1093/intqhc/mzt090

How many diagnosis fields are needed to capture safety events in administrative data? Findings and recommendations from the WHO ICD-11 Topic Advisory Group on Quality and Safety

SASKIA E. DRÖSLER[1], PATRICK S. ROMANO[2], VIJAYA SUNDARARAJAN[3,4], BERNARD BURNAND[5], CYRILLE COLIN[6], HAROLD PINCUS[7,8,9] AND WILLIAM GHALI[10,11] FOR THE WORLD HEALTH ORGANIZATION QUALITY AND SAFETY TOPIC ADVISORY GROUP

[1]Faculty of Health Care, Niederrhein University of Applied Sciences, Reinarzstrasse 49, Krefeld 47805, Germany, [2]Division of General Medicine, University of California–Davis School of Medicine, Sacramento, CA, USA, [3]Department of Medicine, St. Vincent's Hospital, University of Melbourne, Australia, [4]Department of Medicine, Southern Clinical School, Monash University, Australia, [5]Institute of Social and Preventive Medicine (IUMSP), Lausanne University Hospital, Lausanne, Switzerland, [6]Department of Medical Information, Health Evaluation and Clinical Research, University Lyon I, Hospices Civils de Lyon, France, [7]Department of Psychiatry, Columbia University and the New York State Psychiatric Institute, New York, NY, USA, [8]Irving Institute for Clinical and Translational Research at Columbia University and New York-Presbyterian Hospital, New York, NY, USA, [9]RAND Corporation, Pittsburgh, PA, USA, [10]Department of Community Health Sciences and the Institute for Public Health, University of Calgary, Calgary, Alberta, and [11]Faculty of Medicine, University of Calgary, Calgary, Alberta, Canada

Address reprint requests to: Saskia Drősler, Competence Center for Routine Data in Healthcare, Niederrhein University of Applied Sciences, Reinarzstrasse 49, D-47805 Krefeld, Germany. Fax: +49-2151-822-6660; E-mail: saskia.droesler@hs-niederrhein.de

Accepted for publication 19 November 2013

Abstract

Objective. As part of the WHO ICD-11 development initiative, the Topic Advisory Group on Quality and Safety explores meta-features of morbidity data sets, such as the optimal number of secondary diagnosis fields.

Design. The Health Care Quality Indicators Project of the Organization for Economic Co-Operation and Development collected Patient Safety Indicator (PSI) information from administrative hospital data of 19–20 countries in 2009 and 2011. We investigated whether three countries that expanded their data systems to include more secondary diagnosis fields showed increased PSI rates compared with six countries that did not. Furthermore, administrative hospital data from six of these countries and two American states, California (2011) and Florida (2010), were analysed for distributions of coded patient safety events across diagnosis fields.

Results. Among the participating countries, increasing the number of diagnosis fields was not associated with any overall increase in PSI rates. However, high proportions of PSI-related diagnoses appeared beyond the sixth secondary diagnosis field. The distribution of three PSI-related ICD codes was similar in California and Florida: 89–90% of central venous catheter infections and 97–99% of retained foreign bodies and accidental punctures or lacerations were captured within 15 secondary diagnosis fields.

Conclusions. Six to nine secondary diagnosis fields are inadequate for comparing complication rates using hospital administrative data; at least 15 (and perhaps more with ICD-11) are recommended to fully characterize clinical outcomes. Increasing the number of fields should improve the international and intra-national comparability of data for epidemiologic and health services research, utilization analyses and quality of care assessment.

Keywords: world health organization, international classification of diseases, quality indicators, patient safety, risk adjustment, diagnosis-related groups

Introduction

International comparisons of patient safety in acute care hospitals are performed and published biannually through the Health Care Quality Indicators Project of the Organization for Economic Co-Operation and Development (OECD) [1, 2]. These comparisons are based on an indicator set developed and maintained by the US Agency for Health Care Research and Quality (AHRQ), known as the AHRQ Patient Safety Indicators (PSIs). PSI definitions rely on administrative hospital data, are in the public domain and were harmonized for international use [3]. The OECD adopted several of these PSIs based on their scientific soundness and international applicability through a systematic and evidence-based process [4]. Administrative hospital data are available in many OECD countries and do not require additional data collection efforts, thanks to the broad international use of the International Classification of Diseases (ICD).

In 2015, the World Health Organization (WHO) plans to release the 11th revision of the ICD [5]. Topic advisory groups (TAGs) [6] 'serve as the planning and coordinating advisory body for specific issues which are key topics in the update and revision process' [7]. The Quality and Safety TAG was charged with proposing concepts and defining terms to support the quality and safety use case for ICD-11 implementation. As part of this process, the TAG discussed desirable meta-features of morbidity data sets, such as the preferred definition of the principal or primary diagnosis, the reporting of diagnosis timing and the number of diagnosis fields. In this article, we explore how many secondary diagnosis fields should be provided in hospital data sets to consistently capture patient safety-related events. Previous studies have demonstrated selective under-ascertainment of comorbidities among patients who died, and bias in observed associations with mortality, if records are limited to 5 or 9 instead of 25 secondary diagnoses [8, 9].

Up to 20 OECD member states participated in several calculation rounds from 2007 to 2011. Previous studies showed significant positive correlations between the mean number of secondary diagnoses actually coded among eligible cases and the country's age–sex standardized PSI rate [10, 11]. These comparative reporting efforts, and similar efforts within large countries such as the United States and Canada, may be biased by variation in the maximum number of diagnoses [12, 13].

To assess the importance of this problem and to offer recommendations regarding the minimum number of secondary diagnosis fields in hospital data sets, we reviewed three sources of data. First, we obtained data from two OECD PSI calculation rounds in 2009 and 2011. In the 2009 calculation round, countries were asked to provide the distribution of PSI numerator diagnoses across the secondary diagnosis data fields in their data sets. In both calculation rounds, countries provided information on the available number of fields for secondary diagnoses. Between 2009 and 2011, three countries increased the number of available data fields, offering the opportunity for a natural experiment on the impact of this policy change on country-specific PSI rates. Secondly, we obtained 2011 data from California's health data agency on the distribution of PSI numerator diagnoses across the 25 secondary diagnosis fields used in California. We then replicated this analysis using the 2010 State Inpatient Database (SID) of Florida, which has 30 secondary diagnosis fields. Together, these investigations allowed us to explore whether states or countries can improve their tracking of quality and safety indicators by increasing the number of secondary diagnosis fields in their administrative data sets.

Methods

Patient safety indicators

The AHRQ PSIs were developed by a team at the University of California San Francisco, the University of California Davis and Stanford University, based on previous research by Iezzoni and colleagues [14]. Complete documentation on indicator definitions and on AHRQ's selection process, development methods and expert panel review findings is available online [15]. The AHRQ PSIs exclusively rely on routinely collected hospital data such as diagnoses, procedures and selected patient characteristics related to each hospitalization. AHRQ PSI definitions refer to the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) for use in the USA; harmonized definitions were provided to countries using ICD-10 [3, 16]. The following patient safety indicators used in this investigation were selected by the Health Care Quality Indicators Expert Group, which guides the OECD's Health Care Quality Indicators Project

- catheter-related bloodstream infection (previously known as 'selected infections due to medical care'),
- postoperative pulmonary embolism (PE) or deep vein thrombosis (DVT)
- postoperative sepsis,
- accidental puncture or laceration,
- foreign body left in during procedure.

The definitions of all of the indicators listed above rely on secondary diagnoses, displayed in Table 1. Cases with a PSI-related sensitive principal diagnosis are excluded because most countries define the principal diagnosis as present on admission, whereas the OECD project focuses on events that occurred during hospitalization. Obstetric PSIs are omitted from this analysis because obstetric lacerations may be coded in the principal diagnosis field, and because the mean number of secondary diagnoses is low for obstetric records.

Participating OECD countries and PSI calculation

Overall, 19 countries participated in the 2009 PSI calculation project and 20 countries in 2011 [2, 17]. As not all the countries submitted technical information on their data systems or provided distributions of their numerator diagnoses, data from nine countries are included in this analysis (Table 2): Belgium, Canada, Germany, Italy, New Zealand, Singapore, Spain, Sweden and the USA. Four of these countries use ICD-10 or country-specific versions thereof (e.g. ICD-10 CA in Canada, ICD-10 AM in New Zealand, ICD-10 GM in Germany). Five countries use ICD-9 CM, but only the USA uses the current annual version. Hospital discharge data were collected in 2006 or 2007 and submitted in 2009, or were collected in 2008 or 2009 and submitted in 2011 (except that Belgian cases were from 2006 to 2007). Each country provided summary data based on its own analysis of either a probability sample of all hospitalized patients (20% in the USA, based on the Nationwide Inpatient Sample and 10% in Germany) or a complete sample of eligible discharges, although one country (i.e. Spain) excluded non-public hospitals. In no case did these excluded hospitals account for >15% of a country's inpatient hospitalizations. Numerator and denominator counts for

Table 1 PSI underlying ICD codes (numerator definitions) [16]

PSI title	Numerator definition: ICD-9 CM	Numerator definition: ICD-10 WHO
Catheter-related bloodstream infection (previously known as 'selected infections due to medical care')	999.3 Complications of medical care, not elsewhere classified: other infection 996.62 Infection and inflammatory reaction due to internal prosthetic device, implant and graft: due to vascular device, implant and graft	T80.2 Infections following infusion, transfusion and therapeutic injection T82.7 Infection and inflammatory reaction due to other cardiac and vascular devices, implants and grafts T88.0 Infection following immunization
Postoperative pulmonary embolism (PE) or deep vein thrombosis (DVT)	415.11 Iatrogenic pulmonary embolism and infarction 415.19 Pulmonary embolism and infarction, other 451.11 Phlebitis and thrombosis of femoral vein (deep) (superficial) 451.19 Phlebitis and thrombophlebitis of deep vessel of lower extremities—other 451.2 Phlebitis and thrombophlebitis of lower extremities 451.81 Phlebitis and thrombophlebitis of iliac vein 451.9 Phlebitis and thrombophlebitis of other sites—of unspecified site 453.40 DVT-emblsm lower ext nos 453.41 DVT-emb prox lower ext 453.42 DVT-emb distal lower ext 453.8 Other venous embolism and thrombosis of other spec. veins 453.9 Other venous embolism and thrombosis of unspec. site	I26.0 Pulmonary embolism with mention of acute cor pulmonale I26.9 Pulmonary embolism without mention of acute cor pulmonale I80.1 Phlebitis and thrombophlebitis of femoral vein I80.2 Phlebitis and thrombophlebitis of other deep vessels of lower extremities I80.3 Phlebitis and thrombophlebitis of lower extremities, unspecified I80.8 Phlebitis and thrombophlebitis of other sites I80.9 Phlebitis and thrombophlebitis of unspecified site I82.8 Embolism and thrombosis of other specified veins I82.9 Embolism and thrombosis of unspecified vein
Postoperative sepsis	038.0 Streptococcal septicaemia 038.10 Staphylococcal ependence, unspecified 038.11 Staphylococcus aureus septicaemia 038.19 Other staphylococcal septicaemia 038.2 Pneumococcal ependence (streptococcus pneumoniale ependence) 038.3 Septicaemia due to anaerobes 785.52 Septic shock 785.59 Other shock w/o mention of trauma 998.0 Postoperative shock 038.40 Gram-negative organism, unspecified 038.41 *Hemophilus influenzae* 038.42 *Escherichia coli* 038.43 *Pseudomonas* 038.44 Serratia 038.49 Septicaemia due to other gram-negativeorganisms 038.8 Other specified septicaemias 038.9 Unspecified septicaemia 995.91 Systemic inflammatory response syndrome due to infectious process w/o organ Dysfunction 995.92 Systematic inflammatory response syndrome due to infectious process w/organ dysfunction	A40.0 Septicaemia due to streptococcus, Group a A40.1 Septicaemia due to streptococcus, Group b A40.2 Septicaemia due to streptococcus, Group d A40.3 Septicaemia due to streptococcus pneumoniae A40.8 Other streptococcal septicaemia A40.9 Streptococcal septicaemia, unspecified A41.0 Septicaemia due to *Staphylococcus aureus* A41.1 Septicaemia due to other specified staphylococcus A41.2 Septicaemia due to unspecified staphylococcus A41.3 Septicaemia due to haemophilus influenzae A41.4 Septicaemia due to anaerobes A41.5 Septicaemia due to other gram-negative organisms A41.8 Other specified septicaemia A41.9 Septicaemia, unspecified R57.8 Other shock T81.1 Shock during or resulting from a procedure, not elsewhere classified

(*continued*)

Table 1 Continued

PSI title	Numerator definition: ICD-9 CM	Numerator definition: ICD-10 WHO
Accidental puncture or laceration[a]	998.2 Accidental puncture or laceration during a procedure	T81.2 Accidental puncture and laceration during a procedure, not elsewhere classified
Foreign body left in during procedure[a]	998.4 Foreign body accidentally left during a procedure 998.7 Acute reactions to foreign substance accidentally left during a procedure	T81.5 Foreign body accidentally left in body cavity or operation wound following a procedure T81.6 Acute reaction to foreign substance accidentally left during a procedure

Ext., extremities; spec., specified; unspec., unspecified.
[a]Additional ICD codes for external causes are not displayed.

each indicator were reported by 5-year age and sex strata, starting with 15–19 years and ending with 85 or more years. For each indicator, we aggregated the age–sex group specific denominator counts provided by each country to produce an (internal) standard population, which was then used along with each country's age–sex group indicator rates to form direct age–sex standardized rates [18].

Collection and analysis of additional PSI Data from OECD countries

In both calculation rounds, countries were asked to report the available number of data fields for secondary diagnoses, which revision of ICD was in use, and other features of their administrative data sets. We estimated the difference in rates for each PSI in each country between the 2009 and 2011 rounds. We then compared weighted mean PSI rate differences between the three countries that increased secondary diagnosis fields and the other countries that did not do so, using the non-parametric Mann–Whitney-*U*-test for independent samples.

The distribution of PSI-defining secondary diagnoses (e.g. foreign body retained in the patient) across the data fields was evaluated within the 2009 PSI calculation process. Countries were asked to report counts of numerator-defining diagnoses, starting with '1st Secondary Diagnosis field in your database' up to '6th+ Secondary Diagnosis field in your database'. Six countries, Germany, Ireland, Italy, Spain, New Zealand and the USA, provided these data.

Estimates using data from the California Office of Statewide Health Planning and Development and the Florida SID

We further estimated the effects of truncation of secondary diagnosis fields using 2011 data from the California Office of Statewide Health Planning and Development [19] and 2010 data from the Florida SID [20]. The PSI relevant ICD-9 CM codes 999.31 (Infection due to central venous catheter), 998.2 (accidental puncture or laceration during a procedure), and 998.4 (Foreign body accidentally left during a procedure) were flagged; we analysed the proportional distributions of those codes across 24 secondary diagnosis fields in California and across 30 secondary diagnosis fields in Florida.

Results

Available numbers of secondary diagnosis fields in the data sets

The numbers of available data fields for secondary diagnoses ranged from 5 (Italy) to an unlimited number in Belgium, Singapore and Sweden (Table 2). New Zealand, Spain and Sweden increased the number of secondary diagnosis fields between the 2009 and 2011 calculation rounds. Among these three countries (shown in Fig. 1), Sweden expanded its data fields from 6 to an unlimited number, and showed increased rates for all PSIs. Rates in New Zealand (which expanded from 24 to 98) and Spain (which expanded from 12 to 20) increased for postoperative sepsis, but not consistently for other PSIs. Figure 2 shows variable trends in PSI rates across the six countries that did not alter available diagnosis fields; for example, the rate of 'Postoperative Sepsis' substantially decreased in Belgium but slightly increased in other countries.

Among the three countries that added secondary diagnosis fields between the first and second rounds, the mean age–sex standardized rates of 'Foreign Body', 'Catheter-related Blood stream Infection', 'Postoperative PE or DVT', 'Postoperative Sepsis' and 'Accidental Puncture or Laceration' increased by 0.021 per 10 000 patients, and by 0.016, 0.064, 0.331 and 0.018 per 100 patients, respectively. In comparison, among the six countries that did not add secondary diagnosis fields, the mean age–sex standardized rates of the same five PSIs increased by 0.0078 per 10 000 patients, and by 0.0007, 0.122, −0.014 and 0.086 per 100 patients, respectively. None of the mean 'differences in differences' between the three countries in Fig. 1 and the six countries in Fig. 2 was statistically significant at $P < 0.10$.

Distribution of PSI-related ICD codes within the data fields in international data

Data from Italy were removed from this analysis because Italy provided only five data fields for secondary diagnoses. Mean

Table 2 Data from participating OECD countries

Country	Year of OECD data collection		ICD revision in use	Number of available secondary diagnosis fields		Mean number of secondary diagnoses (among population at risk for PSI foreign body)		Population at risk for PSI foreign body		
	2009 PSI calculation round	2011 PSI calculation round		2009 questionnaire	2011 questionnaire	2009 questionnaire	2011 questionnaire	2009 PSI calculation round	2011 PSI calculation round	Estimated sample size (% of all adult hospitalizations)
Belgium	2006	2007	ICD-9 CM	No limit	No limit	6.72	6.40	1 516 867	1 184 800	100%
Canada	2007	2009	ICD-10 CA	24	24	3.11	3.22	1 705 508	1 737 321	100%
Germany	2007	2009	ICD-10 GM	89	89	5.31	5.23	1 273 816	1 300 365	10%
Italy	2006	2009	ICD-9 CM	5	5	1.50	1.48	7 345 711	6 565 127	100%
New Zealand	2007	2009	ICD-10 AM	**24**	**98**	3.93	3.39	428 316	502 740	100%
Singapore	2007	2009	ICD-9 CM	No limit	No limit	4.00	4.39	241 178	244 632	100%
Spain	2007	2009	ICD-9 CM	**12**	**20**	3.71	4.20	3 288 840	3 365 483	90%
Sweden	2007	2009	ICD-10	**6**	**No limit**	2.50	3.70	1 263 707	1 201 276	100%
USA	2006	2008	ICD-9 CM	15	15	6.02	7.07	33 298 777	33 219 050	Estimation from a 20% sample

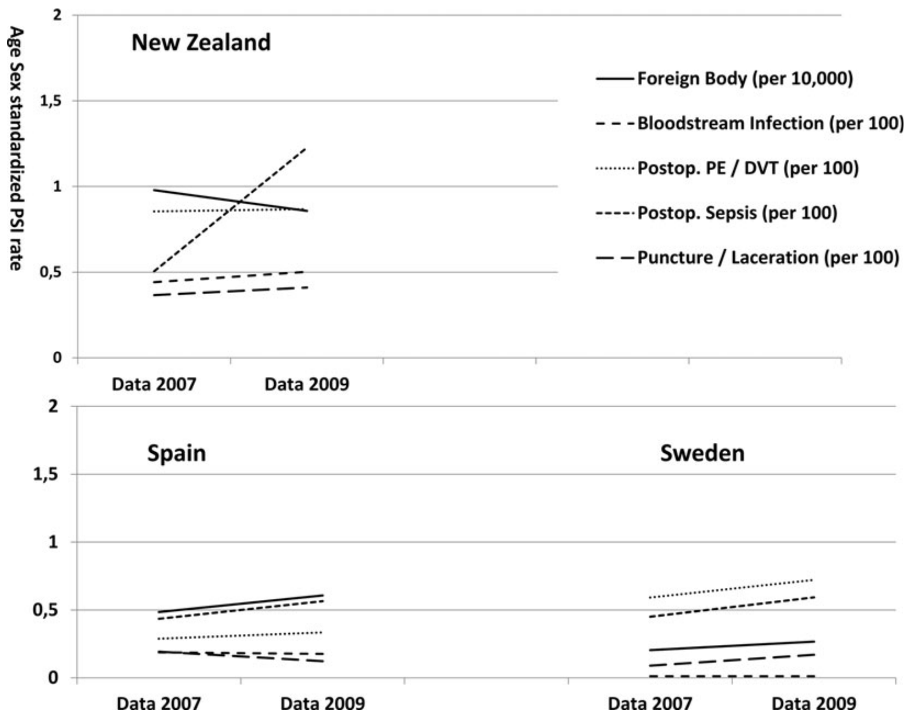

Figure 1 Age–sex standardized PSI rates from 2007 to 2009 among three countries that increased secondary diagnosis fields.

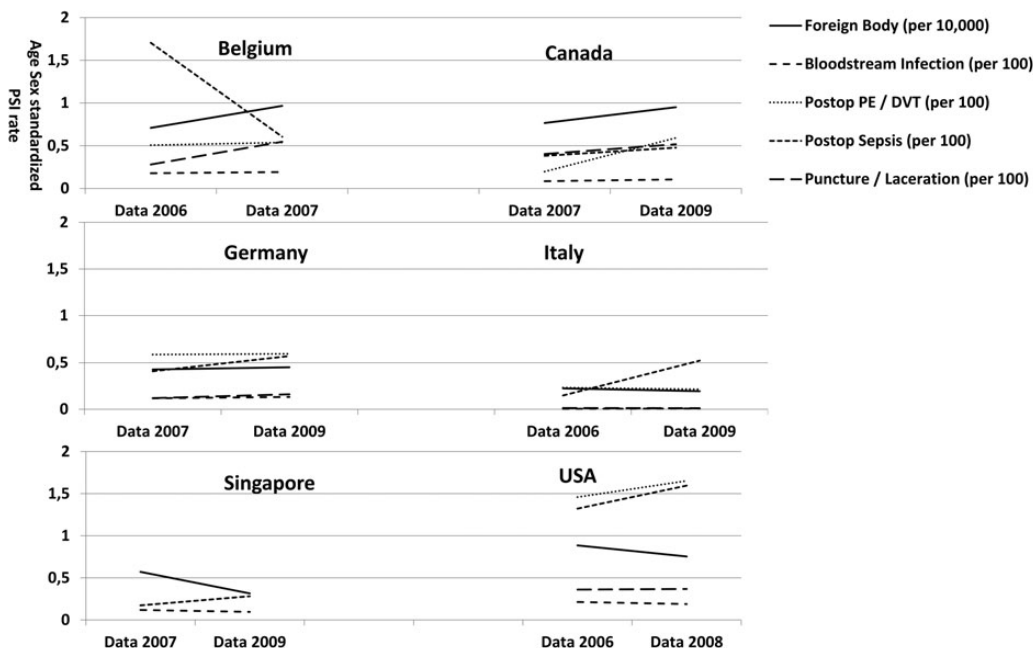

Figure 2 Age–sex standardized PSI rates from 2007 to 2009 (except as noted) among six countries that did not increase secondary diagnosis fields.

distributions by PSI across the remaining five countries are shown in Figure 3. Two PSIs, catheter-related bloodstream infection and postoperative sepsis show the greatest susceptibility to truncation bias, with high proportions of relevant ICD codes in the sixth secondary diagnosis field or beyond. At the country level, Spain and Germany found ~60% of catheter-related bloodstream infections and ~40% of postoperative sepsis in the sixth or higher secondary diagnosis field (data not displayed). Histograms by country show that PSI relevant ICD codes are less frequently assigned to the first five secondary diagnosis fields in Germany and Spain than in New Zealand and the USA (Fig. 3).

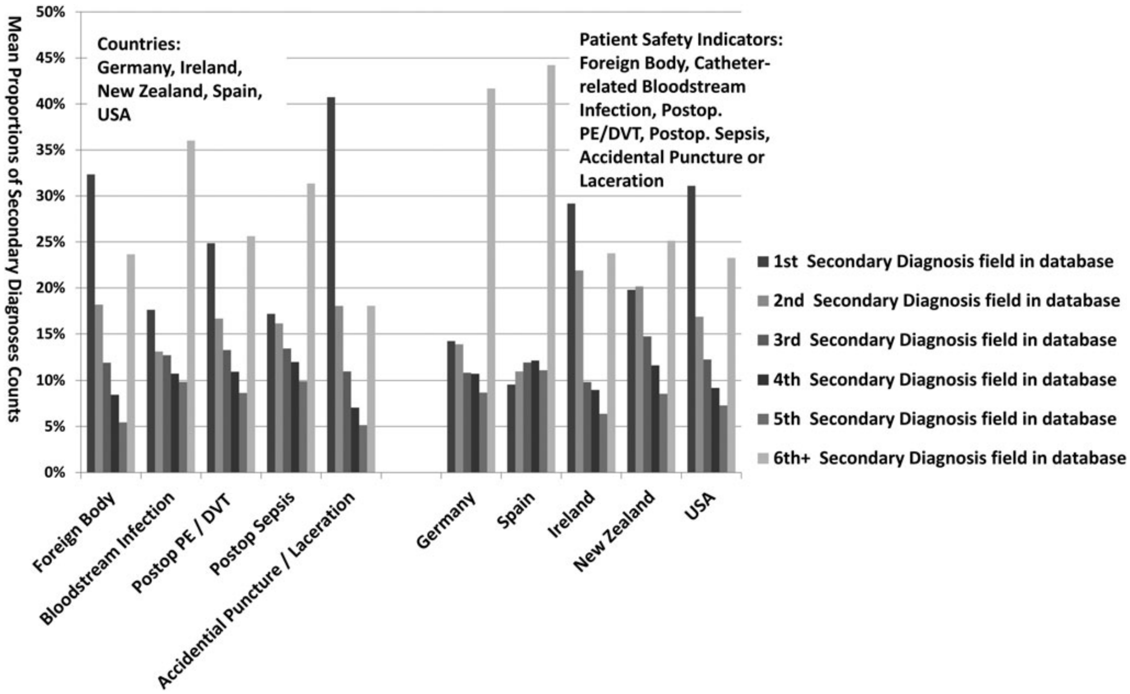

Figure 3 Mean distribution of PSI related ICD codes in data fields by indicator (Germany, Spain, Ireland, New Zealand and USA) and by country (2009 data collection).

Distribution of PSI-related ICD codes within the data fields in California and Florida data

Among 3.93 million discharges in California from 2011, we found 3624 instances of PSI-relevant ICD-9-CM 999.31 (Infection due to central venous catheter) as a secondary diagnosis (Table 3). About 65% of these events were captured within nine secondary diagnosis fields and ~90% were captured within 15 secondary diagnosis fields. This means that a limit of nine data fields would be associated with a loss of 35% of these critical events. Regarding the code 998.2 (accidental puncture or laceration during a procedure), 92.5% of 7352 reported events were captured within nine secondary diagnosis fields and 98.6% were captured within 15 secondary diagnosis fields. Finally, of the 242 retained foreign body events (ICD 998.4 Foreign body accidentally left during a procedure) coded as secondary diagnoses in the California data, 92.5% were listed in the first nine spaces and 97.5% were listed in the first 15. The 2010 data from Florida (2.64 million discharges) show similar results: 62.4% of 4480 central venous catheter infections were coded in fields 1–9, and ~89% in fields 1–15. The events 'Foreign Body' ($n = 138$) and 'Accidental Puncture' ($n = 4352$) show concordant results: ~93% of the relevant codes were found in secondary diagnosis fields 1–9 and 99% were in fields 1–15.

Discussion

Hard evidence on the optimal structure of health data is difficult to produce and obtain, as randomized controlled trials are not feasible. Despite this fact, the analyses presented here provide some insight into how the number of diagnosis fields available in hospital discharge data may affect reporting of adverse events arising in hospital. Through analysis of PSI rates in countries that changed the number of secondary diagnosis fields, we found no clear and consistent effect of increasing the number of fields on country-specific PSI rates over a 2-year period. However, this analysis was limited by the small number and self-selection of participating countries; each of the three countries involved in this natural experiment had unique circumstances. Specifically, New Zealand increased from 24 to 98 secondary diagnosis fields, which would not be expected to affect PSI rates significantly, because 24 was already an ample number (based on our other analyses). Spain increased from 12 to 20 secondary diagnosis fields, which might make a difference, except that Spain was underutilizing the available fields in both time periods (Table 2). So only the increase in Sweden, from 6 to unlimited, clearly represents conversion from an inadequate database to an adequate database. This change was associated with a consistent increase in PSI rates and the largest increase in the mean number of reported secondary diagnoses.

Our analyses of the distribution of safety-related diagnosis codes, using both international data and data from two large American states, demonstrate that the number of secondary diagnosis fields should be probably 15 or more. Otherwise, ICD codes for important events are often truncated and the clinical complexity of a hospitalization is not adequately reflected in its administrative data. PSI results from Italy (Fig. 2) support this conclusion, because Italy provides five

Table 3 Distributions of PSI numerator diagnoses in California (2011) and Florida (2010)

	California			Florida		
Hospitalizations	3 698 955			2 640 092		
Number of secondary diagnosis fields	24			30		
	Absolute number of coded events	% of events in data fields 1–9	% of events in data fields 1–15	Absolute number of coded events	% of events in data fields 1–9	% of events in data fields 1–15
ICD 999.31 (infection due to central venous catheter)	3624	64.71	90.26	4480	62.37	88.73
ICD 998.2 (accidental puncture or laceration during a procedure)	7352	92.50	98.67	4352	93.73	98.83
ICD 998.4 (foreign body accidentally left during a procedure)	242	92.56	97.52	138	92.75	98.55

data fields for secondary diagnoses and reported the lowest PSI rates of all participating countries. In most states and countries, however, the number of data fields seems adequate, but these fields are not used to the optimal extent, perhaps due to resource constraints within reporting hospitals or limited instruction and enforcement by health data agencies. Countries in which hospitals have an economic incentive, such as reimbursement by diagnosis-related groups at the case level, report higher mean numbers of secondary diagnoses and elevated PSI rates relative to countries without such payment schemes [10, 11]. Consistent with Iezzoni and colleagues' findings, increasing the number of data fields might not eliminate bias if hospitals do not utilize those fields [8].

In the USA, available spaces for secondary diagnoses vary among 28 states participating in the 2010 Healthcare Cost and Utilization Project from 8 (Nebraska and Utah) to 82 (South Dakota) [21]. The Centers for Medicare and Medicaid Services used to process only 9 diagnoses but updated its systems in 2011 to collect up to 25 diagnoses, with broad support from the hospital community [22]. Losing clinical information by truncating spaces for secondary diagnoses can be simulated using richer administrative data sets, such as those from California and Florida. Both states have similar coding habits, with earlier sequencing of accidental punctures or lacerations (998.2) and retained foreign bodies during procedures (998.4), but later sequencing of central venous catheter infections (999.31). It remains unclear to what extent clinical coders are concerned about limited data space and assign relevant information to the top spaces. Furthermore, the chronology of code assignment is unclear and might vary internationally. If secondary diagnoses are assigned as coders review each record, 'early' complications, such as those related to an operation, may be coded in earlier data fields than 'late' events, such as infections. In our data, ∼3% of all secondary diagnoses were assigned to data fields 20–24 (California) and 2.6% to data fields 20–30 (Florida), respectively. German data with 18.32 million hospitalizations in 2011 and 89 available fields show that 98.5% of all secondary diagnoses are assigned within 20 and 99.8% within 30 secondary diagnosis fields (Destatis, Federal Statistical Office, personal communication, 12/18/2012).

Besides the assessment of patient safety, there are other reasons for an adequate number of secondary diagnosis fields: health services research based on administrative hospital data strongly relies on case mix and risk adjustment [23, 24]. Common risk adjustment algorithms such as the Charlson or Elixhauser index require accurate data on comorbidity, which can only be collected and electronically stored if sufficient diagnosis fields are available. Hospitalizations with comorbidities related to different organ systems cannot be fully classified with a very limited number of secondary diagnoses. Furthermore, the concept of the ICD classification contains double coding, meaning that one clinical condition may need to be captured using more than one code (e.g. 995–999 ICD-9 CM complication codes often carry instructions to 'use additional code' to describe the specific complication). Some country-specific versions of ICD-10, such as ICD-10 CM, offer more precoordinated 'combination codes' to minimize this problem, but the need for multiple coding is likely to increase with ICD-11.

Our findings suggest that data systems with fewer than 15 spaces for secondary diagnoses (and perhaps 20 spaces as new code clustering mechanisms in ICD-11 increase the number of codes needed to capture some clinical concepts) are likely to lose relevant clinical information. The WHO has created a framework for evidence-informed decision-making around the development of ICD-11 and the data systems that will support it. In support of WHO's overriding goal of producing better, evidence-informed data systems to support health and health systems throughout the world, the Quality and Safety TAG recommends that health data agencies collect at least 15–20 secondary diagnoses fields for optimal capture of clinical information. Even more secondary diagnosis fields will be necessary in the future, as new code clustering mechanisms in ICD-11 may increase the number of diagnosis codes needed to capture complex clinical concepts.

Acknowledgments

This investigation was supported by the OECD as part of the Health Care Quality Indicators project. The first and second authors served as paid consultants to the OECD to design the data collection protocol, to prepare the technical manual, to analyse the data submitted by participating countries and to interpret the project findings. Appropriate health agencies in Belgium, Canada, Germany, Ireland, Italy, New Zealand, Singapore, Spain, Sweden and the USA supported this research by performing the calculations and providing the data in accordance with a technical manual distributed by the authors. Niek Klazinga, Head of the Health Care Quality Indicators (HCQI) Project, OECD Health Division is acknowledged for supporting this investigation. The WHO Topic Advisory Group for Quality and Safety membership includes: William Ghali (chair), Harold Pincus (co-chair), Marilyn Allen, Susan Brien, Bernard Burnand, Cyrille Colin, Saskia Drösler, Alan Forster, Yana Gurevich, James Harrison, Lori Moskal, Bill Munier, Donna Pickett, Hude Quan, Patrick Romano, Danielle Southern and Vijaya Sundararajan. The authors acknowledge the expertise and contributions of the following individuals at TAG meetings: David Van der Zwaag, Chris Chute, Eileen Hogan, Ginger Cox, Bedirhan Ustun and Nenad Kostanjek. They have all contributed to both the TAG work plan and some aspects of the paper.

Funding

This project was supported by Grant Number 1R13HS020543 from the Agency for Healthcare Research and Quality (AHRQ). Additional support was provided by the National Center for Advancing Translational Sciences, National Institutes of Health (NIH), through Grant Number UL1 TR000040. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality or the National Institutes of Health.

References

1. Organization for Economic Co-Operation and Development. http://www.oecd.org/health/hcqi (27 August 2013, date last accessed).
2. Organization for Economic Co-Operation and Development. Health at a glance 2011: OECD indicators. OECD Publishing. http://www.oecd.org/health/healthpoliciesanddata/49105858.pdf (27 August 2013, date last accessed).
3. Quan H, Drösler S, Sundararajan V *et al*. Adaptation of AHRQ patient safety indicators for use in ICD-10 administrative data by an international consortium. In: Henriksen K, Battles JB, Keyes MA, Grady ML (eds). *Advances in Patient Safety: New Directions and Alternative Approaches*. **Vol. 1** (assessment). Rockville, MD: Agency for Healthcare Research and Quality, 2008. http://www.ahrq.gov/downloads/pub/advances2/vol1/advances-quan_52.pdf (27 August 2013, date last accessed).
4. McLoughlin V, Millar J, Mattke S *et al*. Selecting indicators for patient safety at the health system level in OECD countries. *Int J Qual Health Care* 2006;**18** (Suppl. 1):14–20.
5. World Health Organization. http://www.who.int/classifications/icd/revision/en (27 August 2013, date last accessed).
6. World Health Organization. http://www.who.int/classifications/icd/ICDRevision.pdf (27 August 2013, date last accessed).
7. World Health Organization. http://www.who.int/classifications/icd/TAGs/en/index.html (27 August 2013, date last accessed).
8. Iezzoni LI, Foley SM, Daley J *et al*. Comorbidities, complications, and coding bias. Does the number of diagnosis codes matter in predicting in-hospital mortality? *JAMA* 1992;**267**:2197–203.
9. Romano PS, Mark DH. Bias in the coding of hospital discharge data and its implications for quality assessment. *Med Care* 1994;**32**:81–90.
10. Drösler SE, Klazinga NS, Romano PS *et al*. Application of patient safety indicators internationally: a pilot study among seven countries. *Int J Qual Health Care* 2009;**21**:272–8.
11. Drösler SE, Romano PS, Tancredi DJ *et al*. International comparability of patient safety indicators in 15 OECD member countries: a methodological approach of adjustment by secondary diagnoses. *Health Serv Res* 2012;**47**(1 Pt 1):275–92.
12. Agency for Healthcare Research and Quality. 2010 national healthcare quality report. AHRQ Publication No. 11-0004 2011. http://www.ahrq.gov/research/findings/nhqrdr/nhqrdr10/qrdr10.html (27 August 2013, date last accessed).
13. Canadian Institute for Health Information (CIHI). A snapshot of health care in Canada as demonstrated by Top 10 lists, 2011. https://secure.cihi.ca/free_products/Top10ReportEN-Web.pdf (27 August 2013, date last accessed).
14. Weingart SN, Iezzoni LI, Davis RB *et al*. Use of administrative data to find substandard care: validation of the complications screening program. *Med Care* 2000;**38**:796–806.
15. McDonald KM, Romano PS, Geppert J *et al*. Measures of patient safety based on hospital administrative data—the patient safety indicators. Report No.: 02-0038I. Rockville, MD: Agency for Healthcare Research and Quality (US), 2002. http://www.ncbi.

nlm.nih.gov/books/NBK43854/pdf/TOC.pdf (27 August 2013, date last accessed).

16. Drösler S. Facilitating cross-national comparisons of indicators for patient safety at the health-system level in the OECD countries. OECD Health Technical Paper No. 19. http://www.oecd.org/dataoecd/24/48/40401929.pdf (27 August 2013, date last accessed).
17. Drösler SE, Romano PS, Wei L. Health care quality indicators project: patient safety indicators. Report 2009. OECD Health Working Paper No. 47. http://www.oecd.org/dataoecd/56/31/44192992.pdf (27 August 2013, date last accessed).
18. Rothman KJ, Greenland S. *Modern Epidemiology.* Philadelphia: Wolters Kluwer Health. Lippincott Williams & Wilkins, 2008.
19. *The Office of Statewide Health Planning and Development*. Sacramento, CA. http://www.oshpd.ca.gov/HID/Products/PatDischargeData/PublicDataSet/index.html (27 August 2013, date last accessed).
20. HCUP Databases. Healthcare cost and utilization project (HCUP), 2012. Rockville, MD: Agency for Healthcare Research and Quality. www.hcup-us.ahrq.gov/sidoverview.jsp (27 August 2013, date last accessed).
21. HCUP Central Distributor SPSS Load Programs. Healthcare cost and utilization project (HCUP). Rockville, MD: Agency for Healthcare Research and Quality. http://www.hcup-us.ahrq.gov/spssload/spssload_list.jsp?state=FL&year=2011&db=SID (27 August 2013, date last accessed).
22. Federal Register/Vol. 75, No. 85/Tuesday, May 4, 2010/Proposed Rules. http://www.healthwatchusa.org/publications/2012-Documents/20100504-CMS-Not-Capturing-Events-23914.pdf (27 August 2013, date last accessed).
23. Bottle A, Sanders RD, Mozid A *et al*. Provider profiling models for acute coronary syndrome mortality using administrative data. *Int J Cardiol* 2012 **10**. pii: S0167-5273(12)01154-0.
24. Chikuda H, Yasunaga H, Horiguchi H *et al*. Mortality and morbidity in dialysis-dependent patients undergoing spinal surgery: analysis of a national administrative database in Japan. *J Bone Joint Surg Am* 2012;**94**:433–8.