

©Health Research and Educational Trust
DOI: 10.1111/j.1475-6773.2011.01290.x
RESEARCH ARTICLE

International Comparability of Patient Safety Indicators in 15 OECD Member Countries: A Methodological Approach of Adjustment by Secondary Diagnoses

Saskia E. Drösler, Patrick S. Romano, Daniel J. Tancredi, and Niek S. Klazinga

Objective. To improve the international comparability of patient safety indicators based on administrative hospital data, adjustment of country-specific rates by a proxy measure of diagnostic coding intensity was tested.

Data Sources. Secondary data (numerator and denominator counts of patient safety indicators) based on adults discharged from acute care hospitals between 2006 and 2008 was used.

Study Design. A retrospective cross-sectional study using hospital administrative data was performed.

Data Collection. Belgium, Canada, Denmark, Germany, Italy, Ireland, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the United Kingdom, and the United States provided data according to detailed instructions.

Principal Findings. Age- and sex-standardized rates varied across countries. An ordinary least squares regression model was estimated for each Patient Safety Indicator (PSI) using the mean number of secondary diagnoses among denominator cases as the predictor ($R^2 = 23$ percent to 56 percent). Estimated country-specific residuals were linearly transformed into adjusted PSI rates. Variation among age-sex standardized PSI rates decreased substantially after this adjustment.

Conclusions. International comparisons of health system performance based on unadjusted patient safety indicators are problematic due to suspected coding or ascertainment bias. The model could be an interim approach to provide comparable information on hospital quality, with a long-term goal of improving international consistency in diagnostic reporting in administrative data.

Key Words. Patient safety, quality indicators, international classification of diseases

There is increasing interest in international comparisons of health system performance, as reflected in recent work led or supported by the World

Health Organization (Murray and Evans 2003; Groene et al. 2008; Nolte and McKee 2008), the Commonwealth Fund (Schoen et al. 2007; Davis, Schoen, and Stremikis 2010), the European Commission (Kramers 2003; <http://www.echim.org/>), and the Organization for Economic Co-Operation and Development (OECD 2009). The findings of some of these comparisons received attention during the recent health reform debate in the United States (Muennig and Glied 2010), although patient safety was not within the scope of this previous work.

Collecting information on patient safety events from OECD member countries is now part of the conceptual framework of the OECD's work on comparing health systems (Arah et al. 2006) and is regarded as an important module in the OECD Health Care Quality Indicators Project (<http://www.oecd.org/health/hcqi>). To advance this work, the OECD convened an international expert panel in 2004, which rigorously evaluated 59 candidate indicators of patient safety and endorsed 21 for international use (Millar and Mattke 2004; McLoughlin et al. 2006) based on the following criteria: impact on health (including clear gaps between actual and potential levels of health), policy importance, susceptibility to being influenced by the health care system (independent of confounders like patient risk), face validity (including the basic clinical rationale for the indicator and past usage in national or other quality reporting activities), content validity, data availability on the international level, and reporting burden. Twelve of these 21 indicators came from a larger set developed and maintained by the U.S. Agency for Health Care Research and Quality (U.S. Agency for Health Care Research and Quality [AHRQ]), known as the AHRQ Patient Safety Indicators (PSIs). PSI definitions are in the public domain and were harmonized for international use by a separate collaborative group that included experts from six OECD countries (Quan et al. 2008).

Up to 16 countries participated in two previous rounds of PSI data collection in 2007 and 2008, demonstrating the feasibility of the data collection methods and the usability of the technical manual prepared by the authors.

Address correspondence to Saskia E. Drösler, M.D., Competence Center for Routine Data in Healthcare, Niederrhein University of Applied Sciences, Reinarzstraße 49, D-47805 Krefeld, Germany; e-mail: saskia.droesler@hs-niederrhein.de. Patrick S. Romano, M.D., M.P.H., and Daniel J. Tancredi, Ph.D., are with the UC Davis Division of General Medicine and Center for Healthcare Policy and Research, Sacramento, CA. Niek S. Klazinga, M.D., Ph.D., is with the Organization for Economic Co-operation and Development, Paris, France, and is also with the Department of Social Medicine, Academic Medical Centre/University of Amsterdam, Amsterdam, the Netherlands.

These previous data analyses demonstrated that PSI definitions could be applied to data from countries using either ICD-10 or ICD-9 coded data (Drösler et al. 2009b). However, we found substantial international variation in rates, more than could reasonably be attributed to differences in health system performance alone, stimulating additional efforts to understand and to improve comparability. Hospital administrative data from various countries can differ in terms of who is responsible for code assignment, strength, and scope of incentives for coding, implementation of coding guidelines, and data storage limitations, all of which are discussed in detail later. This article focuses on variation in the completeness of hospital coding, represented by the mean number of secondary diagnoses among patients at risk, as this variation was identified as the single strongest correlate of PSI rates among 37 U.S. states (Raetzman et al. 2008). As complete harmonization of international data collection methods is not feasible in the short term, we introduce a mathematical model to adjust for quantitative discrepancies in diagnostic coding across countries to attain more comparable indicator rates.

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 for the AHRQ. Precise documentation on indicator definitions and on their selection process, development, and continuous review is available online (McDonald et al. 2002; <http://www.qualityindicators.ahrq.gov/>). 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 United States; additional harmonized definitions were provided to countries using ICD-10 (Drösler 2008). Based on the experience accumulated through previous data collections, the following Patient Safety Indicators were selected by the Health Care Quality Indicators Expert Group, which guides the OECD's Health Care Quality Indicators Project, for the 2009 calculation round:

- 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
- Obstetric trauma—vaginal delivery with instrument (i.e., forceps or vacuum)
- Obstetric trauma—vaginal delivery without instrument

Those indicators were selected from the original set of 12 endorsed by the OECD's international expert panel (Millar and Mattke 2004; McLoughlin et al. 2006). The AHRQ PSI for "Postoperative Hip Fracture" was dropped because it requires procedure codes and dates to indicate the timing of operating room procedures. As a number of different procedure classifications are in use internationally, it was impossible to provide all participating countries with a comparable list of procedure codes for hip fracture repair. The AHRQ PSIs for "Complications of Anesthesia" and "Decubitus Ulcer" were dropped because of evidence of poor validity, at least in the absence of a diagnosis timing (present on admission) indicator (Houchens, Elixhauser, and Romano 2008). The AHRQ PSI for "Transfusion Reaction" was dropped because of its extremely low frequency. Finally, the AHRQ PSI for "Birth Trauma" was dropped because of difficulty establishing an internationally comparable denominator definition of inborn neonates.

With the exception of the obstetric indicators, the definitions are intended to be applied to records on adult patients older than 17 years. For some PSIs, AHRQ offers two versions with different denominator definitions: provider-level and area-level indicators. Provider-level indicators were used in the project; the denominator reflects the population of hospitalized cases at risk.

The numerator definitions of the nonobstetric indicators mainly rely upon secondary diagnoses. A case is counted as positive if the defined critical event appears as a coded secondary diagnosis in the data. If the critical event is coded as the primary (principal) diagnosis, then the case is excluded from the denominator under the assumption that the event was the cause of the hospitalization. Only events that occurred during inpatient treatment are supposed to be captured and flagged as possibly related to patient safety; however, none of the participating countries (except Canada) was able to provide nationally representative data with "present on admission" or diagnosis timing flags.

Participating Countries

The following 15 countries participated in the investigation: Belgium (BEL), Canada (CAN), Denmark (DNK), Germany (DEU), Ireland (IRL), Italy (ITA), New Zealand (NZL), Norway (NOR), Portugal (PRT), Singapore (SGP), Spain (ESP), Sweden (SWE), Switzerland (CHE), the United Kingdom (GBR, data from England), and the United States of America (U.S.A.). Nine 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). Six countries use ICD-9-CM, but only the U.S.A. uses the current annual version. Hospital discharge data were collected in 2006 or 2007 (except that Danish cases were from 2008). Each country provided summary data based on its own analysis of data representing either a probability sample of all hospitalized patients (20 percent in the United States, based on the Nationwide Inpatient Sample, and 10 percent in Germany) or a complete sample of eligible discharges, although two countries (i.e., Ireland and Spain) excluded nonpublic hospitals. In no case did these excluded hospitals account for more than 15 percent of a country's inpatient hospitalizations. Numerator and denominator counts for each indicator were reported by 5 year age and sex strata, starting with 15–19 years and ending with 85 or more years (Drösler, Romano, and Wei 2009a).

Total denominator populations of the PSI “Foreign body left in during procedure” lie in the range between 241,178 (SGP) and 33,298,777 (U.S.A.). This indicator is well-suited for comparing patient populations across countries as its denominator by definition covers all hospitalized patients aged over 17 years. Table 1 shows the minimum and maximum denominator populations for all indicators. Due to the unique definition of each indicator (e.g., operative cases only are used in computing the indicator “Postoperative PE or DVT”), these denominator populations vary substantially across PSIs.

As former analyses of both U.S. data (Raetzman et al. 2008) and international data (Drösler et al. 2009a, b) demonstrated marked correlations between nonobstetric PSI rates and the mean number of coded secondary diagnoses, countries were also asked to compute and report the mean number of secondary diagnoses and the mean length of stay among the denominator cases of each indicator. We also surveyed country representatives on other factors that might affect PSI reporting, including reimbursement-related incentives, definitions of all relevant data elements, and diagnosis and procedure classification systems. Countries were asked to exclude same-day cases or cases with a length of stay less than 24 hours (based on the definitions used in

Table 1: Range of Age-Sex Standardized Patient Safety Indicator Rates and Range of Mean Number of Secondary Diagnoses Reported across 15 Participating OECD Countries

	Countries	Missing Due to Data Problem	Denominator Population Minimum	Denominator Population Maximum	PSI Rate Minimum (%)	PSI Rate Maximum (%)	R ² %* (p-Value)	Secondary Diagnoses Minimum	Secondary Diagnoses Maximum
Catheter-related bloodstream infection	15		165,676	21,891,1250	0.005 (ITA)	0.442 (NZL)	41 (0.010)	1.45	6.61
Postoperative pulmonary embolism or deep vein thrombosis	14	SGP	70,544	6,674,866	0.116 (PRT)	1.459 (USA)	56 (0.002)	1.3	5.63
Postoperative sepsis	14	NOR	7,859	1,014,591	0.148 (ITA)	1.704 (BEL)	43 (0.010)	0.87	6.66
Accidental puncture or laceration	12	NOR PRT SGP	358,617	28,505,655	0.013 (ITA)	0.403 (CAN)	31 (0.049)	2.26	7.02
Foreign body left in during procedure	13	NOR PRT	241,178	33,298,777	0.002 (DNK)	0.011 (CHE)	23 (0.096)	1.5	6.72
Obstetric trauma—vaginal delivery with instrument	14	PRT	834	251,934	2.307 (ITA)	16.870 (USA)	1 (>0.5)	1.79	6.88
Obstetric trauma—vaginal delivery without instrument	14	PRT	10,772	2,656,529	0.494 (ESP)	3.768 (USA)	1 (>0.5)	1.4	6.62

*The R² statistic here represents the percentage of variation in PSI rates across countries that was explained by the mean number of secondary diagnoses reported among patients in the denominator at risk for each indicator, in a single-variable ordinary least squares regression model.

each country's data system) from the calculation. This request was based on previous analyses showing that international variation in the use of different clinical environments for short-term treatment may lead to variation in the proportion of day cases across countries, and hence variation in the rates of PSIs for which ascertainment is related to length of stay. Not all countries were able to provide data on all indicators. Columns two and three of Table 1 show how many countries provided data and which countries were unable to report certain indicators. As indicator rates were calculated separately from each other, and every country either used a representative sample or the complete hospital population, our analyses are not affected by different numbers of participating countries per PSI.

Adjustment Model

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 (Rothman and Greenland 2008). Countries submitted numerator and denominator counts by sex for 15 five-year age groups starting from 15 to 85 years and older. This approach adjusts not only for differences in general population structure across countries but also for differences in the age and gender distribution of hospitalized patients, which may be attributable to international differences in hospitalization practices. In other words, we were concerned that there might be greater variation in the age–gender structure of hospitalized patients across developed OECD countries, due to variation in end-of-life care and other clinical practices, than in the age–gender structure of the general population.

Based on the consistent country-level association between the mean number of secondary diagnoses (among denominator cases) and rates of all nonobstetric 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, based on the assumption that the prevalence of any PSI-related diagnosis is linearly related to the mean number of diagnoses reported for eligible patients. Data from Norway were excluded from the analyses of “Accidental Puncture or Laceration,” “Foreign Body,” and “Postoperative Sepsis,” because Norway reported zero or implausibly high (i.e., nearly five times higher than the next highest) rates on these three indicators. Earlier discussions with country representatives confirmed that zero numerator counts for an entire year probably reflect systematic

nonreporting. Data from Portugal were removed from the analysis of “Accidental Puncture or Laceration,” because Portugal reported an implausibly low mean number of secondary diagnoses (0.10). The outcome variable in these regression models was a country’s age–sex standardized 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 standardized rates.

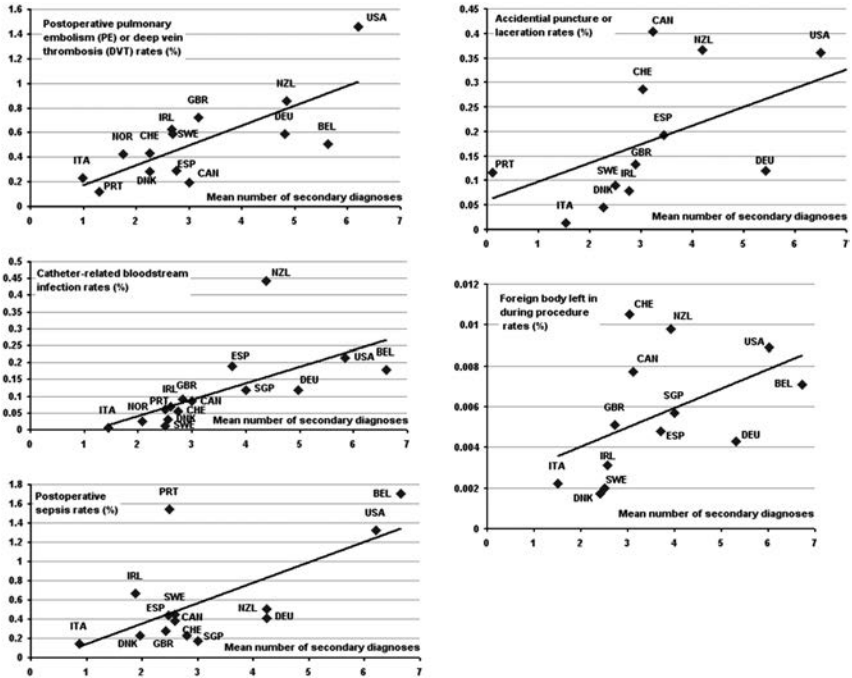
Coefficients of variation (CV, ratio of standard deviation to mean) and maximum-to-minimum ratios for each Patient Safety Indicator are used to analyse the adjustment effect. Regression adjustment is expected to reduce the CV of the outcome by $^2\sqrt{(1 - R^2)}$ (Cohen et al. 2003). The regression adjustment can be straightforwardly extended to include additional covariates. To assess whether ICD version (9 versus 10) or mandatory external cause of injury (E) coding would improve the fit of the simple model that included the mean number of secondary diagnoses, we explored whether these covariates were associated with statistically significant reductions in unexplained variation.

RESULTS

Table 1 shows the range of age–sex standardized but nonadjusted rates, which vary substantially across countries. The mean numbers of secondary diagnoses also demonstrate substantial variation across countries, suggesting discrepancies in diagnostic coding of hospital cases. Based on the R^2 statistic, 23 to 56 percent of the observed variation at the country level in nonobstetric PSI rates is attributable to variation in diagnostic coding. Figure 1 shows a positive correlation between age–sex standardized indicator rates and the mean number of secondary diagnoses calculated from the eligible denominator cases for all of the nonobstetric indicators. In each case, the points above the regression line represent countries that have higher age–sex standardized PSI rates than would be expected based on the average thoroughness of diagnostic coding among patients at risk, while the points below the regression line represent countries with lower age–sex standardized PSI rates than would be expected from the same model.

Table 2 depicts the maximum-to-minimum ratios and the CV of unadjusted and adjusted rates. The adjusted PSI rates demonstrate far less

Figure 1: Scatter Plots of Mean Number of Secondary Diagnoses among Denominator Cases at Risk (*x*-Axis) and Age–Sex Standardized Patient Safety Indicator Rates (*y*-Axis) among 15 Participating OECD Countries



variation than the unadjusted but age–sex standardized rates, for all non-obstetric PSIs.

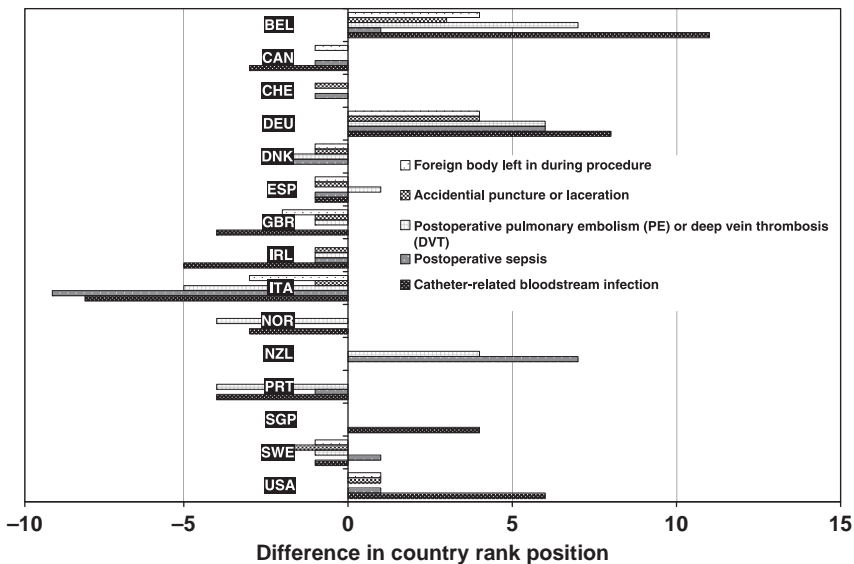
To demonstrate the impact of adjustment for the thoroughness of diagnostic coding, all countries are ranked by PSI rate in ascending order. The country with the lowest PSI rate before adjustment has position one. The differences in rank positions before and after adjustment are shown in Figure 2. It can be demonstrated, with singular exceptions, that rank differences for each country have the same direction for all PSIs. In general, countries with marked positive differences in rank order (shown in Figure 2) also reported high values of the mean number of secondary diagnoses for the same indicator (shown in Figure 1).

For the two obstetric indicators, there was no consistent association between the PSI rate and the mean number of secondary diagnoses at the country level ($R^2 < 2$ percent, $p > .5$). In addition, the regression line for this

Table 2: Variation Measures Applied to Nonadjusted Patient Safety Indicator Rates, and Rates Adjusted by the Mean Number of Secondary Diagnoses, across 15 Participating OECD Countries

	<i>Age-Sex Standardized Rates</i>		<i>Adjusted Age-Sex Standardized Rates</i>	
	<i>Ratio between highest and lowest PSI rate</i>	<i>Coefficient of variation</i>	<i>Ratio between highest and lowest PSI rate</i>	<i>Coefficient of variation</i>
Catheter-related bloodstream infection	84.94	99	16.10	76
Postoperative pulmonary embolism (PE) or deep vein thrombosis (DVT)	12.66	66	8.80	43
Postoperative sepsis	11.51	86	9.18	65
Accidental puncture or laceration	31.00	69	9.52	58
Foreign body left in during procedure	6.18	53	4.06	47

Figure 2: Position Differences of Country Ranking after Adjustment of Age-Sex Standardized Patient Safety Indicator Rates by Mean Number of Secondary Diagnoses among Denominator Cases at Risk *x*-Axis: Difference in Country Rank Position



weak association had an intercept greater than zero ($p = .15$ for laceration with instrumentation, $p = .02$ for laceration without instrumentation), suggesting that these codes often appear in the principal diagnosis field and are therefore relatively immune to truncation bias due to underreporting of diagnoses in some countries.

Mean length of stay data among denominator-eligible cases for each indicator were reported by all countries (available from the authors upon request). In no case was there a statistically significant ($p < .10$) association between mean length of stay and the age–sex standardized PSI rate across countries. Similarly, country-specific use of ICD-9-CM (instead of ICD-10) was not significantly ($p < .10$) associated with any PSI rate, and mandatory E code reporting was not significantly associated with rates of the two PSIs that include E codes in their numerator definitions (i.e., “Accidental Puncture or Laceration” and “Foreign Body”). In multiple regression models that adjusted for the mean number of secondary diagnoses, neither ICD version nor E code reporting was associated with statistically significant reduction in unexplained variation.

DISCUSSION

Nonobstetric PSI rates are positively correlated across countries with the intensity of diagnostic coding, expressed by the mean number of secondary diagnoses among patients at risk. A similar association was observed among 37 U.S. states for two of the same PSIs (“Postoperative Sepsis” and “Catheter-related Bloodstream Infection”); the other three nonobstetric PSIs considered in this study were not analyzed in that report (Raetzman et al. 2008). Therefore, comparative reporting of geographic PSI rates unadjusted for coding intensity could be misleading. These correlations have been found in all three calculation rounds initiated by the OECD, although they are reported in detail publicly for the first time here.

The most plausible explanation for these correlations is that complication-related diagnoses are more likely to be reported in areas where routine coding practice entails more thorough review of clinical documentation and more complete coding of documented diagnoses. However, it is possible that variation in the completeness of clinical documentation (i.e., physician notes) drives much of the observed variation in coding practice, especially among the majority of countries where diagnosis codes are assigned by trained health information professionals. To some extent, the number of secondary

diagnoses depends on the individual billing system of a country; for example, in the United States and Germany Diagnosis Related Groups are used for reimbursement at the individual patient level, creating a strong incentive for complete registration of complications and comorbidities. However, almost every country included in this study reported using some type of diagnosis-based grouping or case mix index for hospital payment. We were unable to ascertain the extent to which hospital budgets in each country are “at risk” from underreporting of comorbidities or complications.

Because it is unclear when and how hospitals in countries with a low mean number of secondary diagnoses will be motivated to code safety-related events, the proposed adjustment method may be a useful interim approach to reduce the confounding effects of coding bias on international comparisons. Variation in the thoroughness of diagnosis reporting is probably the strongest single source of bias, but several additional factors have to be taken into account in comparing PSIs internationally.

Other Factors Compromising International Comparisons of Administrative Data-Based Quality Indicators

Details such as coding guidelines and the locus of responsibility for the correct assignment of diagnosis and procedure codes vary across countries: Not all of the participating countries have standardized coding rules in use. This fact might have an impact on the comparability of the data because the PSI definitions were originally developed for the use in the United States. In the United States, definitions of primary and secondary diagnoses are regulated, whereas other countries apply different (e.g., Canada) or even no (several Scandinavian countries) definitions.

The definitions of the Patient Safety Indicators incorporate several denominator exclusions to eliminate patients who have greatly elevated risk; for example, cases with a principal or secondary diagnosis of immunocompromised state or cancer are excluded from the indicator “Catheter-related bloodstream infection.” Other exclusions (e.g., emergency admissions from the “Postoperative sepsis” denominator population) were designed by AHRQ to reduce the number of false-positive cases due to safety-related diagnoses that were actually present on admission. Those exclusions might impair comparability across countries. Additional investigations are underway by the OECD to quantify the impact of those exclusions.

In several countries, professional coders are responsible for coding inpatient diagnoses, whereas medical doctors who provided the treatment are responsible for the correct assignment of ICD codes in some other countries.

It is questionable whether physicians in the latter countries are trained in ICD use and motivated to assign codes accurately. Additionally, when there is no economic incentive affecting physicians, because they are not employees of the hospital, reporting of inpatient diagnoses might be reduced to a minimum.

U.S. studies of variation in PSI rates across states have shown positive correlations between the use of external cause of injury (E) codes and some PSI rates (Raetzman et al. 2008), even though these PSIs do not include E codes in their numerator definitions. Therefore, E code reporting appears to be a proxy for more thorough coding in general. Similarly, we found that mandatory E code reporting did not reduce the unexplained variation in PSI rates (for the PSIs that use E codes) after adjusting for the mean number of secondary diagnoses.

Data Systems

Data systems vary across countries in different respects. Participating countries reported some variation in the number of available data fields for secondary diagnoses, although the number of available fields greatly exceeded the mean number of reported diagnoses everywhere. One country was not able to participate in this investigation as only two data fields for secondary diagnoses were available in its database.

Recent evaluations of the criterion validity of the Patient Safety Indicators revealed that secondary diagnoses already present on hospital admission generate a significant number of false-positive cases (Houchens, Elixhauser, and Romano 2008; White et al. 2009). Patient Safety Indicators are designed to capture critical events during hospitalization. Some countries (e.g., Canada, U.S.A., New Zealand) have introduced a qualifying marker to distinguish whether or not a safety-related condition was present on admission. If used correctly, this marker will improve the validity of the data, but it may also compromise future comparability if it is not used consistently in all participating countries.

True Differences in Health Care Systems

Differences in health care systems such as the process of treatment might confound international comparisons. For example, average length of stay in OECD countries varies between 7.8 days (CHE, DEU) and 3.5 days (DNK) with an OECD average of 6.5 days (OECD 2009). Similarly, in this project, average length of stay varied between 8.76 days (DEU) and 4.49 days (DNK) among denominator cases of the indicator “Foreign body left in during

procedure.” The relatively high mean length of stay reflects the fact that countries were asked to eliminate same day cases. We hypothesized that countries with a longer average length of stay would have higher PSI rates because their patients have more time at risk for a patient safety event. Surprisingly, there was no association at the country level between mean length of stay and age–sex standardized PSI rates.

OECD analyses of Cesarean delivery rates reflect a high variability between 15.9 (NOR) and 39.7 (ITA) per 100 live births (OECD 2009). Although Italy reported rather low rates in both obstetric indicators and the highest Cesarean delivery rates, no statistically significant correlation was found between obstetric indicator rates and Cesarean delivery rates among 12 countries that provided these data. Severity of illness is another unmeasured factor that may vary systematically across countries, even after restricting analyses to patients “at risk” for specific complications and adjusting (as we did) for age and sex. At this point, it is unclear to what extent variations in severity of illness and treatment affect comparability of PSI rates.

CONCLUSION

In a globalized world, comparisons of health system performance among highly industrialized countries are of great interest for stakeholders as well as for patients. The OECD has played a leading role in identifying indicators of health system performance and collecting internationally comparable data. Regarding patient safety, we found that these comparisons are heavily compromised by the completeness of diagnostic coding. Mathematical adjustment of age and sex standardized rates by the mean number of secondary diagnosis is an approach to reduce obvious coding bias that may confound between-country comparisons in health-system performance, although it is based on the strong assumptions that the probability of a PSI-related diagnosis on an inpatient record is proportional to the total number of coded diagnoses, holding other determinants of health system performance (e.g., severity of illness) constant, and that the magnitude of this linear association is independent of these other determinants. Because the intensity of secondary coding is likely to be correlated with some of these determinants, our proposed method could lead to an overcorrection that could reduce the apparent magnitude of real between-country differences (Cohen et al. 2003; Ceyhan and Goad 2009). In other words, low PSI rates of countries with relatively few secondary diagnoses on each record may be overcorrected if these countries also have relatively

healthy hospital populations. On balance, however, the wide variation in coding intensity suggests that the proposed method may be a useful interim approach to provide more comparable international information on patient safety with the goals of improving patient safety by mutual learning and informing the public. In the long term, more consistent documentation and coding of safety-related diagnoses and ongoing evaluation of indicator performance across developed countries, similar to recent efforts in the United States (White et al. 2009; Utter et al. 2011), the United Kingdom (Bottle and Aylin 2009), and Belgium (Gillet et al. 2008), would be very desirable.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: This investigation was initiated and 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 analyze the data submitted by participating countries, and to interpret the project findings. The last author is seconded to the OECD Secretariat but retains his primary appointment at the University of Amsterdam. No other OECD staff or consultants were involved in the preparation of this manuscript. Appropriate health agencies in Belgium, Canada, Denmark, Germany, Italy, Ireland, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the United Kingdom, and the United States of America supported this research by performing the calculations and providing the data in accordance with a technical manual distributed by the authors. We acknowledge the members of the Patient Safety Indicators interest group of the International Methodology Consortium for Coded Health Information (<http://www.imecchi.org>), who assisted the authors in developing internationally harmonized indicator definitions: Hude Quan, M.D., Ph.D.; Vijaya Sundararajan, M.D., M.P.H., F.A.C.P.; Eugene Wen, M.D., Ph.D.; Bernard Burnand, M.D., M.P.H.; Chantal Marie Couris, Ph.D.; Patricia Halfon, M.D.; Jean-Marie Januel, R.N., M.P.H.; Edward Kelley, Ph.D.; Ph.D.; Jean-Christophe Luthi, M.D., Ph.D.; Lori Moskal, C.H.I.M.; Eric Pradat, M.D.; Jennie Shephard, B.S.; Lawrence So, M.A.; Lalitha Sundaesan, M.B.B.S., M.P.H.; Linda Tournay-Lewis, M.H.S., C.H.I.M.; Béatrice Trombert-Paviot, M.D., Ph.D.; Greg Webster, M.Sc.; William A. Ghali, M.D., M.P.H.

Disclosure: None.

Disclaimer: None.

REFERENCES

- Arah, O. A., G. P. Westert, J. Hurst, and N. S. Klazinga. 2006. "A Conceptual Framework for the OECD Health Care Quality Indicators Project." *International Journal for Quality in Health Care* 18 (suppl 1): 5–13.
- Bottle, A., and P. Aylin. 2009. "Application of AHRQ Patient Safety Indicators to English Hospital Data." *Quality and Safety in Health Care*: 18: 303–8.
- Ceyhan, E., and C. L. Goad. 2009. "A Comparison of Analysis of Covariate-Adjusted Residuals and Analysis of Covariance." *Communications in Statistics-Simulation and Computation* 38 (10): 2019–38.
- Cohen, J., P. Cohen, S. G. West, and L. S. Aiken. 2003. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Davis, K., C. Schoen, and K. Stremikis. 2010. "Mirror, Mirror on the Wall: How the Performance of the U.S. Health Care System Compares Internationally, 2010 Update." The Commonwealth Fund [accessed April 1, 2011]. Available at <http://www.commonwealthfund.org/Content/Publications/Fund-Reports/2010/Jun/Mirror-Mirror-Update.aspx>
- Drösler, S. 2008. "Facilitating Cross-National Comparisons of Indicators for Patient Safety at the Health-System Level in the OECD Countries." OECD Health Technical Papers, No. 19 [accessed April 1, 2011]. Available at <http://www.oecd.org/dataoecd/24/48/40401929.pdf>
- Drösler, S., P. Romano, and L. Wei. 2009a. "Health Care Quality Indicators Project: Patient Safety Indicators Report 2009." OECD Health Working Papers, No. 47 [accessed April 1, 2011]. Available at <http://www.oecd.org/dataoecd/56/31/44192992.pdf>
- Drösler, S. E., N. S. Klazinga, P. S. Romano, D. J. Tancredi, M. A. Gogorcena Aoiz, M. C. Hewitt, S. Scobie, M. Soop, E. Wen, H. Quan, W. A. Ghali, S. Mattke, and E. Kelley. 2009b. "Application of Patient Safety Indicators Internationally: A Pilot Study among Seven Countries." *International Journal for Quality in Health Care* 21 (4): 272–8.
- Gillet, P., P. Kolh, W. Sermeus, A. Vleugels, J. Jacques, K. Van den Heede, S. Devriese, F. Vrijens, and S. Verelst. 2008. "Détection des événements indésirables dans les bases de données administratives. Health Services Research (HSR). Bruxelles: Centre fédéral d'expertise des soins de santé (KCE). 2008. KCE Reports 93B. (D/2008/10.273/74)" [accessed April 2, 2011]. Available at http://kce.fgov.be/index_nl.aspx?SGREF=5195&CREF=11883
- Groene, O., N. Klazinga, V. Kazandjian, P. Lombrail, and P. Bartels. 2008. "The World Health Organization Performance Assessment Tool for Quality Improvement in Hospitals (PATH): An Analysis of the Pilot Implementation in 37 Hospitals." *International Journal for Quality in Health Care* 20 (3): 155–61.
- Houchens, R. L., A. Elixhauser, and P. S. Romano. 2008. "How Often Are Potential Patient Safety Events Present on Admission?" *Joint Commission Journal on Quality and Patient Safety* 34 (3): 154–63.
- Kramers, P. G. 2003. "The ECHI Project: Health Indicators for the European Community." *European Journal of Public Health* 13: 101–6.

- McDonald, K. M., P. S. Romano, J. Geppert, S. M. Davies, B. W. Duncan, K. G. Shojania, and A. Hansen. 2002. "Measures of Patient Safety Based on Hospital Administrative Data - The Patient Safety Indicators." Agency for Healthcare Research and Quality (US); AHRQ Technical Reviews and Summaries. Publication No. 02-0038 [accessed April 1, 2011]. Available at <http://www.ahrq.gov/downloads/pub/evidence/pdf/psi/psi.pdf>
- McLoughlin, V., J. Millar, S. Mattke, M. Franca, P. M. Jonsson, D. Somekh, and D. Bates. 2006. "Selecting Indicators for Patient Safety at the Health System Level in OECD Countries." *International Journal for Quality in Health Care* 18 (suppl 1): 14–20.
- Millar, J., and S. Mattke. 2004. "Selecting Indicators for Patient Safety at the Health Systems Level in OECD Countries". OECD Health Technical Papers, No. 18, OECD Publishing [accessed April 1, 2011]. Available at http://www.oecd-ilibrary.org/social-issues-migration-health/selecting-indicators-for-patient-safety-at-the-health-systems-level-in-oecd-countries_800266264370
- Muennig, P. A., and S. A. Glied. 2010. "What Changes in Survival Rates Tell Us about US Health Care." *Health Affairs* 29 (11): 2105–13.
- Murray, C. J. L., and D. B. Evans. 2003. "Health Systems Performance Assessment: Debates, Methods and Empiricism. World Health Organization" [accessed April 1, 2011]. Available at <http://whqlibdoc.who.int/publications/2003/9241562455.pdf>
- Nolte, E., and M. McKee. 2008. "Measuring the Health of Nations: Updating An Earlier Analysis." *Health Affairs* 27 (1): 58–71.
- Organization for Economic Co-Operation and Development. 2009. *Health at a Glance 2009. OECD Indicators*. Paris: OECD Publishing.
- Quan, H., S. Drösler, V. Sundararajan, E. Wen, B. Burnand, C. M. Couris, P. Halfon, J. M. Januel, E. Kelley, N. Klazinga, J. C. Luthi, L. Moskal, E. Pradat, P. S. Romano, J. Shepherd, L. So, L. Sundaresan, L. Tournay-Lewis, B. Trombert-Paviot, G. Webster, and W. A. Ghali. 2008. "Adaptation of AHRQ Patient Safety Indicators for Use in ICD-10 Administrative Data by an International Consortium." In *Advances in Patient Safety: New Directions and Approaches*, edited by K. Henriksen, J. B. Battles, M. A. Keyes, and D. I. Lewin. Assessment. Rockville, MD: Agency for Healthcare Research and Quality (AHRQ) [accessed April 1, 2011]. Available at: http://www.ahrq.gov/downloads/pub/advances2/vol1/advances-quan_52.pdf.
- Raetzman, S., E. Stranges, R. M. Coffey, M. L. Barrett, R. Andrews, E. Moy, and J. Brady. 2008. "Patient Safety in Hospitals in 2004: Toward Understanding Variation across States." HCUP Methods Series Report # 2008-02, U.S. Agency for Healthcare Research and Quality [accessed April 1, 2011]. Available at http://www.hcup-us.ahrq.gov/reports/2008_02.pdf
- Rothman, K. J., and S. Greenland. 2008. *Modern Epidemiology*. Philadelphia: Wolters Kluwer Health. Lippincott Williams & Wilkins.
- Schoen, C., R. Osborn, M. M. Doty, M. Bishop, J. Peugh, and N. Murukutla. 2007. "Toward Higher-Performance Health Systems: Adults' Health Care Experiences in Seven Countries, 2007." *Health Affairs Web Exclusive* 26 (6): w717–34.

- Utter, G. H., A. M. Borzecki, A. K. Rosen, P. A. Zrelak, R. Baron, J. Cuny, H. M. Kaafarani, B. Sadeghi, J. J. Geppert, and P. S. Romano. 2011. "Designing an Abstraction Instrument—Lessons from Efforts to Validate the AHRQ Patient Safety Indicators." *The Joint Commission Journal on Quality and Patient Safety* 37 (1): 20–8.
- White, R. H., B. Sadeghi, D. J. Tancredi, P. Zrelak, J. Cuny, P. Sama, G. H. Utter, J. J. Geppert, and P. S. Romano. 2009. "How Valid Is the ICD-9-CM Based AHRQ Patient Safety Indicator for Postoperative Venous Thromboembolism?" *Medical Care* 47 (12): 1237–43.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.