International Journal for Quality in Health Care, 2023, **35(4)**, 1–7
DOI: https://doi.org/10.1093/intqhc/mzad092
Advance Access Publication Date: 9 November 2023
Original Research Article

Validity of 16 AHRQ Patient Safety Indicators to identify in-hospital complications: a medical record review across nine Swiss hospitals

Michael M. Havranek[1,*], Florian Rüter[2], Selina Bilger[2], Yuliya Dahlem[3], Leonel Oliveira[2], Daniela Ehbrecht[4], Rudolf M. Moos[5], Christian Westerhoff[6], Thomas Beck[7,‡], Marie-Annick Le Pogam[8,‡]

[1]Competence Center for Health Data Science, Faculty of Health Sciences and Medicine, University of Lucerne, Frohburgstrasse 3, Lucerne 6002, Switzerland
[2]University Hospital Basel, Petersgraben 4, Basel 4031, Switzerland
[3]University Hospital Zurich, Rämistrasse 100, Zurich 8006, Switzerland
[4]Zug Cantonal Hospital, Landhausstrasse 11, Zug 6340, Switzerland
[5]Cantonal Hospital Winterthur, Brauerstrasse 15, Winterthur 8400, Switzerland
[6]Hirslanden Private Hospital Group, Boulevard Lilienthal 2, Zurich 8152, Switzerland
[7]University Hospital Berne (Inselspital), Freiburgstrasse, Berne 3010, Switzerland
[8]Department of Epidemiology and Health Systems, Unisanté (University Center for Primary Care and Public Health), University of Lausanne, Route de la Corniche 10, Lausanne 1010, Switzerland

[*]Corresponding author. Competence Center for Health Data Science, Faculty of Health Sciences and Medicine, University of Lucerne, Frohburgstrasse 3, Lucerne 6002, Switzerland. E-mail: michael.havranek@unilu.ch
[‡]These authors contributed equally.
Handling Editor: Dr Anthony Staines

Abstract
The validity of the Agency for Healthcare Research and Quality's Patient Safety Indicators (PSIs) has been established in the USA and Canada. However, these indicators are also used for hospital benchmarking and cross-country comparisons in other nations with different health-care settings and coding systems as well as missing present on admission (POA) flags in the administrative data. This study sought to comprehensively assess and compare the validity of 16 PSIs in Switzerland, where they have not been previously applied. We performed a medical record review using administrative and electronic medical record data from nine Swiss hospitals. Seven independent reviewers evaluated 1245 cases at various hospitals using retrospective data from the years 2014–18. True positives, false positives, positive predictive values (PPVs), and reasons for misclassification were compared across all investigated PSIs, and the documentation quality of the PSIs was examined. PSIs 6 (iatrogenic pneumothorax), 10 (postoperative acute kidney injury), 11 (postoperative respiratory failure), 13 (postoperative sepsis), 14 (wound dehiscence), 17 (birth trauma), and 18 and 19 (obstetric trauma with or without instrument) showed high PPVs (range: 90–99%) and were not strongly influenced by missing POA information. In contrast, PSIs 3 (pressure ulcer), 5 (retained surgical item), 7 (central venous catheter-related bloodstream infection), 8 (fall with hip fracture), and 15 (accidental puncture/laceration) showed low PPVs (range: 18–49%). In the case of PSIs 3, 8, and 12 (perioperative embolism/thrombosis), the low PPVs were largely due to the lack of POA information. Additionally, it was found that the documentation of PSI 3 in discharge letters could be improved. We found large differences in validity across the 16 PSIs in Switzerland. These results can guide policymakers in Switzerland and comparable health-care systems in selecting and prioritizing suitable PSIs for quality initiatives. Furthermore, the national introduction of a POA flag would allow for the inclusion of additional PSIs in quality monitoring.

Keywords: patient safety; present on admission; validity; chart review; documentation; quality indicators

Introduction

The Patient Safety Indicators (PSIs) are a set of quality indicators developed by the Agency for Healthcare Research and Quality (AHRQ) providing information on potential hospital complications and adverse events after surgeries, procedures, and childbirth [1, 2]. They have been used for the past two decades in the USA for monitoring potentially preventable patient safety events in the inpatient setting through the automated screening of readily available administrative data [3, 4]. However, the PSIs have also been implemented and adopted in several other countries [5, 6]. One prominent example is their use for cross-country comparisons by the Organisation for Economic Co-operation and Development (OECD) [7, 8].

Their validity has been investigated in various medical record reviews that were mainly conducted in the USA and Canada [9–12]. However, other health-care systems have different clinical practices, administrative data, and coding systems than the USA and Canada. In particular, many countries do not yet have a present on admission (POA) flag in their

Received 16 April 2023; Editorial Decision 10 October 2023; Revised 25 August 2023; Accepted 22 October 2023
© The Author(s) 2023. Published by Oxford University Press on behalf of International Society for Quality in Health Care.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/intqhc/article/35/4/0/7394665 by Universitaetsbibliothek Bern user on 23 November 2023

administrative data. The POA flag indicates whether each secondary diagnosis was pre-existing or not and thus allows to distinguish the in-hospital complications represented by the PSIs from pre-existing conditions [13]. In addition, the PSIs' coding accuracy may vary between countries and hospitals [6], especially if their documentation quality is not monitored. Despite these concerns, very few scientific studies on the validity of the PSIs have been conducted outside North America, and those that have only focused on a limited number of PSIs [14, 15]. Consequently, there exists a significant knowledge gap regarding the validity of the PSIs in many countries, even though they have been widely used for years.

As part of a large initiative to adopt additional quality indicators in Switzerland (as mentioned later), we have assessed the validity of 16 PSIs in a large record review study across nine hospitals. Our primary objective was to estimate the positive predictive values (PPVs) of the PSIs to determine which ones can accurately identify relevant in-hospital complications based on Swiss administrative data. Our secondary objective was to examine the impact of the missing POA flag on the frequency of false positives (FPs) and investigate potential variations in documentation quality across different hospital types.

Methods

Study design and data

This retrospective study was part of a large collaborative research project funded by the Swiss Innovation Agency (Innosuisse) aiming to translate, examine, and adopt international quality indicators into the Swiss health-care and medical coding system (research grant number 40160.1 IP-SBM). The results presented herein are from a manual record review using administrative and electronic medical record data for the fiscal years 2014–18 from nine hospitals that participated in the study: three university (i.e. academic teaching) hospitals, three private hospitals, and three regional cantonal hospitals.

The administrative dataset [16] contained all inpatient stays treated by the hospitals during the study period, with up to 50 diagnosis codes for each stay (from the International Statistical Classification of Diseases and Related Health Problems, 10th revision, German Modification, ICD-10 GM [17]), up to 100 procedure codes (from the Swiss classification of surgical interventions, CHOP [17]), the diagnosis-related group (from the SwissDRG system [18]), other clinically relevant variables such as admission and discharge conditions, and patients' demographic information. Electronic medical records were accessed directly by the reviewers at the hospitals and contained all available information from the patient documentation (e.g. discharge letters; surgery reports; charts; medications; and imaging, laboratory, and other results). In addition, the full national administrative dataset including all Swiss hospitals was used from another study that investigated the excess costs of PSIs in Switzerland in the year 2019 [19] to determine the frequency of PSIs across all hospitals.

PSI definitions

The PSI specifications were translated into the Swiss medical coding systems (ICD-10 GM and CHOP) based on the AHRQ's original definitions (version 2020 [20]). This was done by the authors in close collaboration with medical coding experts from the hospitals, subsequently checked by two independent medical coders, and validated as part of this study. In addition, for the PSIs that rely on POA coding, several additional inclusion/exclusion criteria from the AHRQ's 2015 definitions were used to identify and exclude POA cases [20].

Sampling and record review

A random sample of inpatient stays from those that were flagged as "cases" (i.e. stays with a potentially preventable adverse event) by at least one of the 16 investigated PSIs (Table 1) was drawn for the medical record review. The random sampling was performed across the PSI cases of all participating hospitals aiming to select an equal number of cases from the different hospitals for each PSI, but it was limited by the availability of relevant cases in the participating hospitals. For this reason, the number of reviewed cases is lower for some PSIs compared to others (see, e.g., PSI 15 in Table 2). In addition, for PSI 4 (death after complication), we separately reviewed the subcategories for shock (PSI 4.1), sepsis (PSI 4.2), pneumonia (PSI 4.3), thrombosis/embolism (PSI 4.4), and gastrointestinal hemorrhage (PSI 4.5).

In total, 1245 cases were reviewed at the participating hospitals, by seven independent reviewers comprising five medical doctors and two quality managers with a background in nursing and health sciences. In this sample, 182 (15%) cases were duplicates that were reviewed by two independent reviewers to assess inter-rater reliability (IRR). All reviewers underwent standardized training to familiarize themselves with the PSI definitions and learn the structured review process (as mentioned later). Their assessments were collected using a standardized online questionnaire specifically designed for this research in a previous pilot study.

The structure of the review process was as follows: for all flagged cases, the reviewers first had to assess whether the PSI-related adverse event was present according to the AHRQ's definitions (version 2020 [20]) [i.e. true positives (TPs)] or not (i.e. FPs). For FPs, the reviewers were then asked to indicate the reason for misclassification, including POA-related FPs. For TPs, they were requested to indicate whether the adverse event was also correctly documented in the patient's discharge letter (or whether they had to review other sources in the patient's electronic medical record to find it). This was assessed as a measure for the quality of documentation in patients' discharge letters. In addition, the reviewers were asked to provide their level of subjective certainty with regard to each question they answered (e.g. "How certain are you about this decision?") based on a Likert scale ranging from 1 ("very uncertain") to 10 ("very certain").

Analysis

As part of our statistical analyses, we compared the frequency of TPs, FPs, and the PPVs [where PPV = TPs/(TPs + FPs)] across all PSIs. In addition, we assessed the percentage of FPs that could have been avoided with a POA flag, were due to coding mistakes, or arose through other reasons specific to certain PSIs (as explained later). Furthermore, we identified the percentage of TPs without correct documentation in patients' discharge letters. Finally, we measured IRR as the percentage of agreement on the distinction between TPs

Downloaded from https://academic.oup.com/intqhc/article/35/4/0/7394665 by Universitaetsbibliothek Bern user on 23 November 2023

Table 1. The investigated PSIs.

	AHRQ description[a]	Short form[b]
PSI 2	Death in low-mortality diagnosis-related groups	n/a[c]
PSI 3	Pressure ulcer	Pressure ulcer
PSI 4	Death among surgical inpatients with serious treatable complications	Death after complication
PSI 5	Retained surgical item or unretrieved device fragment count	Retained surgical item
PSI 6	Iatrogenic pneumothorax	Iatrogenic pneumothorax
PSI 7	CVC-related bloodstream infection	CVC bloodstream infection
PSI 8	In-hospital fall with hip fracture	Fall with hip fracture
PSI 9	Postoperative hemorrhage or hematoma	Postoperative hemorrhage/hematoma
PSI 10	Postoperative acute kidney injury requiring dialysis	Postoperative acute kidney injury
PSI 11	Postoperative respiratory failure	Postoperative respiratory failure
PSI 12	Perioperative pulmonary embolism or deep vein thrombosis	Perioperative embolism/thrombosis
PSI 13	Postoperative sepsis	Postoperative sepsis
PSI 14	Postoperative wound dehiscence	Wound dehiscence
PSI 15	Unrecognized abdominopelvic accidental puncture or laceration	Accidental puncture/laceration
PSI 17	Birth trauma—injury to neonate	Birth trauma
PSI 18	Obstetric trauma—vaginal delivery with instrument	Obstetric trauma with instrument
PSI 19	Obstetric trauma—vaginal delivery without instrument	Obstetric trauma without instrument

[a]The AHRQ has discontinued PSIs 1 and 16.
[b]Subsequently, we will use these abbreviations to refer to the PSIs in the main text.
[c]We did not conduct a medical record review for PSI 2 as it does not focus on specific in-hospital complications.

Table 2. The frequency of TPs and FPs, PPVs with CIs, POA status, and undocumented cases across all investigated PSIs

	Cases	TPs	FPs	PPV (%)	Low CI (%)	High CI (%)	POA	POA %	Not Doc	Not Doc %
PSI 3	66	18	48	27	17	38	34	71	6	33
PSI 4	190	156	34	82	77	88	23	68	7	4
PSI 4.1	50	47	3	94	87	100	2	67	2	4
PSI 4.2	39	33	6	85	73	96	5	83	0	0
PSI 4.3	37	32	5	86	75	98	2	40	2	6
PSI 4.4	35	21	14	60	44	76	10	71	1	5
PSI 4.5	29	23	6	79	65	94	4	67	2	9
PSI 5	51	25	26	49	35	63	15	58	3	12
PSI 6	69	68	1	99	96	100	0	0	3	4
PSI 7	58	27	31	47	34	59	8	6	3	11
PSI 8	45	8	37	18	7	29	32	86	1	13
PSI 9	60	41	19	68	57	80	0	0	3	7
PSI 10	50	48	2	96	91	100	2	100	1	2
PSI 11	53	50	3	94	88	100	1	33	3	6
PSI 12	64	40	24	63	51	74	21	88	0	0
PSI 13	73	66	7	90	84	97	2	29	3	5
PSI 14	59	55	4	93	87	100	4	100	2	4
PSI 15	35	15	20	43	26	59	0	0	2	13
PSI 17	59	56	3	95	89	100	0	0	4	7
PSI 18	66	65	1	98	96	100	0	0	1	2
PSI 19	65	64	1	98	95	100	0	0	0	0
Full Sample	1063	802	261	75	73	78	142	54	42	5

Cases = reviewed cases, low CI = lower bound of the 95% CI for the PPV, high CI = upper bound of the 95% CI for the PPV, POA % = percentage of FPs that were POA, Not Doc = not correctly documented in discharge letters, Not Doc % = percentage of TPs not correctly documented, PSI 3 = pressure ulcer, PSI 4 = death after serious complication, PSI 5 = retained surgical item, PSI 6 = iatrogenic pneumothorax, PSI 7 = CVC bloodstream infection, PSI 8 = fall with hip fracture, PSI 9 = postoperative hemorrhage/hematoma, PSI 10 = postoperative acute kidney injury, PSI 11 = postoperative respiratory failure, PSI 12 = perioperative embolism/thrombosis, PSI 13 = postoperative sepsis, PSI 14 = wound dehiscence, PSI 15 = accidental puncture/laceration, PSI 17 = birth trauma, PSI 18 = obstetric trauma with instrument, PSI 19 = obstetric trauma without instrument.

and FPs between two independent reviewers across all PSIs (IRR = $N_{agreement}/N_{total}$).

To compare the frequency of TPs and FPs (underlying the calculation of the PPV), the frequency of FPs due to a missing POA flag, and the frequency of cases without correct documentation in the discharge letters across the different PSIs, we calculated odds ratios (ORs) and their 95% confidence intervals (CI) and used Fisher's exact tests comparing each PSI with the rest of the PSIs combined (i.e. in a separate 2 × 2 contingency table for each comparison). In addition, to explore potential differences between hospital types (i.e. university, private, and cantonal hospitals) in the frequency of TPs and FPs, the frequency of FPs due to a missing POA flag, and the frequency of cases without correct documentation, we used chi-squared (χ^2) tests to compare the actual frequencies with the expected frequencies. Notably, the expected frequencies were calculated based on the number of reviewed PSIs in the respective hospitals. This was necessary because, as shown later, the frequencies differed across the PSIs and the number of reviewed PSIs slightly differed across the participating hospitals (e.g. one hospital reviewed slightly more PSI 3 cases, while another hospital reviewed slightly more PSI

Downloaded from https://academic.oup.com/intqhc/article/35/4/0/7394665 by Universitaetsbibliothek Bern user on 23 November 2023

4 cases). All statistical analyses were performed in Python (version 3.8.8), and results were considered statistically significant if $P < .05$ (with the Bonferroni correction for multiple comparisons across the PSIs, $P < .003$ [21]).

Results

In total, 1063 unique cases (1245 minus the 182 duplicates used to assess IRR) were reviewed, spanning all of the investigated PSIs, after excluding eight cases that could not be evaluated based on the available patient documentation (i.e. two cases each from PSIs 3 and 5, and one case each from PSIs 4.2, 4.4, 13, and 17). Of the 1063 unique cases, the reviewers judged 802 (75%) as TPs and 261 (25%) as FPs. Among the FPs, 142 (54%) were assessed as POA-related, 27 (10%) were indicated as coding errors, 52 (20%) were due to reasons specific to the particular PSIs (as mentioned later), and 40 (15%) were because of other reasons (not falling in any of the pre-determined categories). Across all PSIs, the mean certainty expressed by reviewers in their decisions to distinguish TPs from FPs was 9.36 (SD = 1.29; on a scale from 1 to 10) and the IRR was 93%. Appendix Table S1 presents the frequency of all PSIs in the full national sample including all Swiss hospitals.

Table 2 shows the frequency of TPs and FPs, the PPV, the number of POA-related FPs, and the number of TPs without correct documentation in the discharge letter for all investigated PSIs. Comparing the frequencies of TPs and FPs across all PSIs revealed that the following PSIs had a significantly lower frequency of TPs compared to FPs, relative to the other PSIs combined: PSIs 3 [pressure ulcer; OR = 0.10, 95% CI = 0.06–0.18; *P*-value from Fisher's exact test (*P*) <.001], 5 (retained surgical item; OR = 0.29, CI = 0.16–0.51; $P < .001$), 7 [central venous catheter (CVC) bloodstream infection; OR = 0.26, CI = 0.15–0.44; $P < .001$], 8 (fall with hip fracture; OR = 0.06, CI = 0.03–0.13; $P < .001$), and 15 (accidental puncture/laceration; OR = 0.23, CI = 0.12–0.46; $P < .001$). Conversely, the following PSIs had a significantly higher frequency of TPs to FPs compared to the other PSIs combined: PSIs 6 (iatrogenic pneumothorax; OR = 24.09, CI = 3.33–174.36; $P < .001$), 10 (postoperative acute kidney injury; OR = 8.24, CI = 1.99–34.16; $P < .001$), 11 (postoperative respiratory failure; OR = 5.72, CI = 1.77–18.49; $P < .001$), 13 (postoperative sepsis; OR = 3.25, CI = 1.47–7.18; $P = .002$), 14 (wound dehiscence; OR = 4.73, CI = 1.70–13.18; $P < .001$), 17 (birth trauma; OR = 6.46, CI = 2.00–20.80; $P < .001$), 18 (obstetric trauma with instrument; OR = 22.93, CI = 3.17–166.10; $P < .001$), and 19 (obstetric trauma without instrument; OR = 22.55, CI = 3.11–163.35; $P < .001$).

Comparing the frequency of POA-related FPs revealed that PSIs 3 (pressure ulcer; OR = 2.36, CI = 1.20–4.65; $P = .016$, however, this result did not remain significant after the Bonferroni correction), 8 (fall with hip fracture; OR = 6.63, CI = 2.49–17.64; $P < .001$), and 12 (perioperative embolism/thrombosis; OR = 6.71, CI = 1.95–23.10; $P < .001$) had significantly more POA-related FPs. In contrast, significantly fewer POA-related FPs were found for PSIs 7 (CVC bloodstream infection; OR = 0.25, CI = 0.11–0.58; $P = .001$), 9 (postoperative hemorrhage/hematoma; OR = 0.00, CI = 0.00–0.00; $P < .001$), and 15 (accidental puncture/laceration; OR = 0.00, CI = 0.00–0.00; $P < .001$). Besides missing POA information and coding inaccuracies (as mentioned earlier), PSI-specific reasons for FPs included 24 instances (77%) of infection due to other devices or implants in the vascular system unrelated to a CVC (PSI 7), 15 cases (79%) with postoperative hemorrhage or hematoma that could be treated conservatively without surgical intervention (PSI 9) whereby the flagged procedure code of the repair surgery was unrelated to the adverse event in question, and 13 cases (65%) with accidental abdominal puncture/laceration that was repaired during the first surgery and did not require additional surgical revision for that purpose (PSI 15). In addition, among the TPs, we found a significantly higher frequency of missing documentation in the discharge letters for PSI 3 (OR = 10.26, CI = 3.64–28.91; $P < .001$) compared to the other PSIs.

Table 3 shows the frequency of TPs and FPs, the PPVs, the frequency of POA-related FPs, and the frequency of TPs without correct documentation in the discharge letters grouped by hospital types, alongside the expected frequencies based on the number of specific PSIs reviewed in the hospitals. There

Table 3. A comparison of the frequency of TPs vs. FPs, POA vs. not POA, and documented vs. not documented PSIs by hospital types

	TPs obs[a]	FPs obs	PPV obs (%)	TPs exp[a]	FPs exp	PPV exp (%)
Cantonal	269	83	76	273	79	78
Private	192	52	79	185	59	76
University	341	126	73	344	123	74
	POA obs	Not POA obs	POA % obs	POA exp	Not POA exp	POA % exp
Cantonal	55	28	66	48	35	58
Private	21	31	40	24	28	46
University	66	60	52	68	58	54
	Not Doc obs	With Doc obs	Not Doc % obs	Not Doc exp	With Doc exp	Not Doc % exp
Cantonal	5	263	2	12	256	4
Private	17	169	9	11	175	6
University	20	319	6	19	320	6

Not POA = not present on admission, POA % = percentage of cases that were POA, Not Doc = not correctly documented in discharge letters, With Doc = with correct documentation, Not Doc % = percentage of cases not correctly documented, Cantonal = regional cantonal hospitals, Private = private hospitals, University = university (i.e. academic teaching) hospitals.
[a]The observed frequencies (obs) and expected frequencies (exp) are given, with the latter calculated based on the PSIs reviewed in the respective hospitals as described in the Methods section (i.e. these are not the expected frequencies assumed by the null hypothesis of no differences in frequency but the expected frequencies based on the number of reviewed PSIs in the respective hospitals).

Downloaded from https://academic.oup.com/intqhc/article/35/4/0/7394665 by Universitaetsbibliothek Bern user on 23 November 2023

were no significant differences between university, private, or cantonal hospitals concerning the frequency of TPs and FPs or the proportion of POA cases among FPs. However, the participating cantonal hospitals exhibited a significantly lower percentage of undocumented adverse events in the discharge letters than expected compared to the other hospital types ($\chi^2(2,1063) = 4.27$, $P = .039$).

Discussion

To the best of our knowledge, this is the first scientific study outside the USA and Canada to evaluate the validity of all PSIs and compare them comprehensively. Across the PSIs investigated, we found acceptable PPVs even though Switzerland uses a different medical coding system and the Swiss administrative data does not yet include a POA flag. Nevertheless, there were some significant differences between PSIs concerning both PPV and the reliance on the POA flag.

Statement of principal findings

Two previous studies from Germany and Italy investigated differences in PSI rates for Europe and the USA and cautioned against identifying adverse events based on coded administrative data due to coding inaccuracies and differences between countries [14, 15]. However, our study went beyond this previous research by comprehensively examining which adverse events represented by PSIs can be accurately identified in a health-care system where the PSIs are not systematically monitored. We found that PSIs 6, 10, 11, 13, 14, 17, 18, and 19 (i.e. iatrogenic pneumothorax, postoperative acute kidney injury, postoperative respiratory failure, postoperative sepsis, wound dehiscence, birth trauma, and obstetric trauma with or without instrument) showed high validity in terms of PPV and were not significantly influenced by the missing POA flag. On the contrary, PSIs 3, 5, 7, 8, and 15 (i.e. pressure ulcer, retained surgical item, CVC bloodstream infection, fall with hip fracture, and accidental puncture/laceration) exhibited comparatively low validity (i.e. PPVs), which, in the case of PSIs 3 and 8, was mainly due to the missing POA flag in Swiss administrative data. In addition, in the case of PSI 3, we also found that the documentation in the discharge letters of the participating hospitals could be improved (i.e. pressure ulcers should more reliably be listed as diagnoses in the discharge letters).

Interpretation within the context of the wider literature

Our results largely confirm previous findings from record reviews in the USA and Canada, which, for example, also found good validity for PSIs 6, 10, 11, 14, 18, and 19; low validity for PSIs 3, 5, 7, and 8; and reliance on POA information for PSIs 3 and 8 [9–12]. However, the present findings differ from previous results for PSIs 13 (where we found a better PPV), 15 (where we found a worse PPV), and 12 (where we identified a greater reliance on POA information). These differences may stem from updates to the PSI definitions that have occurred since 2013, when the previous validation studies were performed [4], or from variations in coding practices or health-care delivery between countries. However, no differences were evident in coding accuracy (reflected by the PPV) or the relevance of the POA flag (for the distinction between TPs and FPs) across the different hospital types in Switzerland. We only identified a difference in the documentation quality, with cantonal hospitals appearing to document PSI-related complications more frequently in the discharge letters than the other hospital types.

Implications for policy, practice, and research

From a practical perspective, our results provide insights into which PSIs should be used for quality monitoring initiatives, provider comparisons, or even country comparisons. In particular, PSIs showing high PPVs despite missing POA information—such as PSIs 6, 10, 11, 13, 14, 17, 18, and 19—should primarily be chosen for quality comparisons. Conversely, PSIs with lower validity—including PSIs 3, 5, 7, 8, and 15—should be confined to case-finding and not be employed for provider or country comparisons. These practical insights are relevant for Swiss health policymakers but may also be helpful for the OECD, which uses a selection of PSIs for country comparisons [7, 8]. Among the OECD PSI indicators, PSI 5 (retained surgical item) requires critical reappraisal given its low PPV and reliance on the availability of POA information found in this study and, to a lesser extent, so does PSI 12 (perioperative embolism/thrombosis) given its reliance on POA information.

Strengths and limitations

One important limitation of our study is the fact that we only assessed TPs, FPs, and PPVs in flagged PSI cases since resource restrictions did not permit review of the hundreds of thousands of non-flagged cases necessary to reliably assess true negatives, false negatives, and negative predictive values (given the rarity of the complications represented by the PSIs). Therefore, we recommend that policymakers ensure and monitor the coding accuracy of PSIs across different hospitals if these indicators are included in national quality initiatives or value-based reimbursement systems. However, given the new opportunities presented by natural language processing techniques, future research may develop approaches to assess or mitigate coding inaccuracies for complications used in quality initiatives by detecting complications from patient discharge letters.

Another notable limitation is that we only assessed the validity of the PSIs in the Swiss health-care setting, which may limit the generalizability of our results to different health-care systems of other countries. This limitation also encompasses our specific translation of the PSI definitions into the Swiss coding system. Similarly, it must be noted that the OECD employs slightly different PSI specifications than those issued by the AHRQ (i.e. the ones we adopted in this study). However, we doubt that the finding of low PPV for PSI 5, for example, would differ drastically if the OECD specifications were used.

Conclusions

In summary, we have found significant differences in validity across various PSIs in the Swiss health-care setting. If PSIs were selected for quality initiatives in health-care systems comparable to Switzerland, then PSIs 6, 10, 11, 13, 14, 17,

Downloaded from https://academic.oup.com/intqhc/article/35/4/0/7394665 by Universitaetsbibliothek Bern user on 23 November 2023

18, and 19 (i.e. iatrogenic pneumothorax, postoperative acute kidney injury, postoperative respiratory failure, postoperative sepsis, wound dehiscence, birth trauma, and obstetric trauma with or without instrument) may be favored over others due to their greater validity. Furthermore, introducing a POA flag in hospital administrative data would likely allow additional PSIs, such as PSIs 3, 8, and 12 (i.e. pressure ulcer, fall with hip fracture, and perioperative embolism/thrombosis), to be included in quality initiatives. We hope that our findings will guide policymakers in Switzerland and other countries that are considering the adoption of PSIs for quality initiatives.

Acknowledgements

We would like to thank Annette Egger and Regula Heller for their continued support that made this project possible.

Author contributions

All authors contributed to the conceptualization of the study. M.M.H., F.R., S.B., Y.D., D.E., L.O., and T.B. conducted the investigation (i.e., performed the record reviews). M.M.H. conducted the formal analysis and the writing of the original draft. All authors contributed to the review & editing of the manuscript.

Supplementary data

Supplementary data is available at *IJQHC* online.

Conflict of Interest Statement

MH provides consulting and analysis services regarding quality indicators for the Swiss National Association for Quality Development in Hospitals and Clinics (ANQ), and their software partner INMED GmbH. However, these organizations were not involved in either the design, execution, analysis, and interpretation of the study, or the writing and publication of this manuscript. The other authors have no conflicts of interest to declare that are relevant to the content of this article.

Funding

This work was supported by Research Grant 40160.1 IP-SBM of the Swiss Innovation Agency (Innosuisse), which promotes science-based innovations commissioned by the Swiss Confederation.

Data Availability

The administrative data that support the findings of this study are available from the Swiss Federal Office of Statistics (contactable via gesundheit@bfs.admin.ch). However, the electronic medical records belong to the participating hospitals.

Ethics and other permissions

The study was approved in a jurisdictional inquiry by the Ethics Committee Northwest & Central Switzerland (27 January 2021; ID: Req-2019-00624). Informed consent from patients was not required because the study was conducted as a quality control project within hospitals.

References

1. Miller MR, Elixhauser A, Zhan C *et al*. Patient safety indicators: using administrative data to identify potential patient safety concerns. *Health Serv Res* 2001;**36**:110–32.
2. Tedesco D, Moghavem N, Weng Y *et al*. Improvement in patient safety may precede policy changes: trends in patient safety indicators in the United States, 2000–2013. *J Patient Saf* 2021;**17**:e327–34. https://doi.org/10.1097/PTS.0000000000000615
3. Kaafarani HMA, Rosen AK. Using administrative data to identify surgical adverse events: an introduction to the Patient Safety Indicators. *Am J Surg* 2009;**198**:S63–8. https://doi.org/10.1016/j.amjsurg.2009.08.008
4. Narain W. Assessing estimates of patient safety derived from coded data. *J Healthc Qual* 2017;**39**:230–42. https://doi.org/10.1097/JHQ.0000000000000088
5. Drosler 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. https://doi.org/10.1093/intqhc/mzp018
6. Drosler 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**:275–92. https://doi.org/10.1111/j.1475-6773.2011.01290.x
7. Mattke S, Epstein AM, Leatherman S. The OECD health care quality indicators project: history and background. *Int J Qual Health Care* 2006;**18**:1–4. https://doi.org/10.1093/intqhc/mzl019
8. 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**:14–20. https://doi.org/10.1093/intqhc/mzl030
9. Quan H, Eastwood C, Cunningham CT *et al*. IMECCHI investigators. Validity of AHRQ patient safety indicators derived from ICD-10 hospital discharge abstract data (chart review study). *BMJ Open* 2013;**3**:e003716. https://doi.org/10.1136/bmjopen-2013-003716
10. Romano PS, Mull HJ, Rivard PE *et al*. Validity of selected AHRQ patient safety indicators based on VA National Surgical Quality Improvement Program data. *Health Serv Res* 2009;**44**:182–204. https://doi.org/10.1111/j.1475-6773.2008.00905.x
11. Romano PS, Yasmeen S, Schembri ME *et al*. Coding of perineal lacerations and other complications of obstetric care in hospital discharge data. *Obstet Gynecol* 2005;**106**:717–25. https://doi.org/10.1097/01.AOG.0000179552.36108.6d
12. Rosen AK, Itani KM. Validating the patient safety indicators in the Veterans Health Administration: are they ready for prime time? *J Am Coll Surg* 2011;**212**:921–3. https://doi.org/10.1016/j.jamcollsurg.2010.12.053
13. Winters BD, Bharmal A, Wilson RF *et al*. Validity of the Agency for Health Care Research and Quality Patient Safety Indicators and the Centers for Medicare and Medicaid Hospital-acquired Conditions: a systematic review and meta-analysis. *Med Care* 2016;**54**:1105–11. https://doi.org/10.1097/MLR.0000000000000550
14. Maass C, Kuske S, Lessing C *et al*. Are administrative data valid when measuring patient safety in hospitals? A comparison of data collection methods using a chart review and administrative data. *Int J Qual Health Care* 2015;**27**:305–13. https://doi.org/10.1093/intqhc/mzv045
15. Tedesco D, Hernandez-Boussard T, Carretta E *et al*. Evaluating patient safety indicators in orthopedic surgery between Italy and the USA. *Int J Qual Health Care* 2016;**28**:486–91. https://doi.org/10.1093/intqhc/mzw053

Downloaded from https://academic.oup.com/intqhc/article/35/4/0/7394665 by Universitaetsbibliothek Bern user on 23 November 2023

16. SFSO. *Medical Statistic of Hospitals*. 2020. https://www.bfs.admin.ch/bfs/de/home/statistiken/gesundheit/erhebungen/ms.html (16 April 2023, date last accessed).
17. SFSO. *Instruments of Medical Coding*. 2023. https://www.bfs.admin.ch/bfs/de/home/statistiken/gesundheit/nomenklaturen/medkk/instrumente-medizinische-kodierung.html (16 April 2023, date last accessed).
18. SwissDRG. *Swiss DRG Classification 2019*. 2020. https://www.swissdrg.org/de/akutsomatik/archiv-swissdrg-system/swissdrg-system-802019 (16 April 2023, date last accessed).
19. Giese A, Khanam R, Nghiem S *et al*. Assessing the excess costs of the in-hospital adverse events covered by the AHRQ's Patient Safety Indicators in Switzerland. *PLOS ONE* (in press). In submission.
20. AHRQ. *Patient Safety Indicator Measures*. 2022. https://qualityindicators.ahrq.gov/measures/psi_resources (16 April 2023, date last accessed).
21. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *BMJ* 1995;**310**:170. https://doi.org/10.1136/bmj.310.6973.170

Downloaded from https://academic.oup.com/intqhc/article/35/4/0/7394665 by Universitaetsbibliothek Bern user on 23 November 2023