UC Davis

UC Davis Previously Published Works

Title

Incorporating harms into the weighting of the revised Agency for Healthcare Research and Quality Patient Safety for Selected Indicators Composite (Patient Safety Indicator 90).

Permalink

https://escholarship.org/uc/item/2d14g633

Journal

Health services research, 57(3)

ISSN

0017-9124

Authors

Zrelak, Patricia A
Utter, Garth H
McDonald, Kathryn M
et al.

Publication Date

2022-06-01

DOI

10.1111/1475-6773.13918

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

DOI: 10.1111/1475-6773.13918

METHODS ARTICLE

HSR Health Services Research

Incorporating harms into the weighting of the revised Agency for Healthcare Research and Quality Patient Safety for Selected Indicators Composite (Patient Safety Indicator 90)

Patricia A. Zrelak PhD, RN, NEA-BC, SCRN, CNRN[1,2] | Garth H. Utter MD, MSc[1,3] | Kathryn M. McDonald PhD, MM[4,5] | Robert L. Houchens PhD[6] | Sheryl M. Davies MS[4] | Halcyon G. Skinner PhD, MPH[6,7] | Pamela L. Owens PhD[8] | Patrick S. Romano PhD, MPH[1,9,10]

[1]Center for Healthcare Policy and Research, University of California Davis, Sacramento, California, USA

[2]Kaiser Permanente, Sacramento, CA, USA.

[3]Department of Surgery, Outcomes Research Group, University of California, Davis, Medical Center, Sacramento, California, USA.

[4]Stanford University Center for Health Policy, Stanford University, Stanford, California, USA

[5]Johns Hopkins University Schools of Nursing and Medicine, Baltimore, MD, USA.

[6]Truven Analytics, Ann Arbor, Michigan, USA.

[7]Lehigh University College of Health, Bethlehem, PA, USA.

[8]Agency for Healthcare Research and Quality, Rockville, Maryland, USA

[9]Department of Internal Medicine, University of California, Davis, Medical Center, Sacramento, California, USA

[10]Department of Pediatrics, University of California, Davis, Medical Center, Sacramento, California, USA

Correspondence
Garth H. Utter, Department of Surgery, Outcomes Research Group, University of California, Davis, Medical Center, 2335 Stockton Boulevard, North Addition Office Building, Room 5027, Sacramento, CA 95817, USA.
Email: ghutter@ucdavis.edu

Funding information
Agency for Healthcare Research and Quality (AHRQ), Grant/Award Number: HHSAA290201200003I

Abstract

Objective: To reweight the Agency for Healthcare Research and Quality Patient Safety for Selected Indicators Composite (Patient Safety Indicator [PSI] 90) from weights based solely on the frequency of component PSIs to those that incorporate excess harm reflecting patients' preferences for outcome-related health states.

Data Sources: National administrative and claims data involving hospitalizations in nonfederal, nonrehabilitation, acute care hospitals.

Study Design: We estimated the average excess aggregate harm associated with the occurrence of each component PSI using a cohort sample for each indicator based on denominator-eligible records. We used propensity scores to account for potential confounding in the risk models for each PSI and weighted observations to estimate the "average treatment effect in the treated" for those with the PSI event. We fit separate regression models for each harm outcome. Final PSI weights reflected both the disutilities and the frequencies of the harms.

Data Collection/Extraction Methods: We estimated PSI frequencies from the 2012 Healthcare Cost and Utilization Project State Inpatient Databases with present on admission data and excess harms using 2012–2013 Centers for Medicare & Medicaid Services Medicare Fee-for-Service data.

Principal Findings: Including harms in the weighting scheme changed individual component weights from the original frequency-based weighting. In the reweighted composite, PSIs 11 ("Postoperative Respiratory Failure"), 13 ("Postoperative Sepsis"), and 12 ("Perioperative Pulmonary Embolism or Deep Vein Thrombosis") contributed the greatest harm, with weights of 29.7%, 21.1%, and 20.4%, respectively. Regarding reliability, the overall average hospital signal-to-noise ratio for the reweighted PSI 90 was 0.7015. Regarding discrimination, among hospitals with greater than median volume, 34% had significantly better PSI 90 performance, and 41% had significantly worse performance than benchmark rates (based on percentiles).

 © 2021 Health Research and Educational Trust

Conclusions: Reformulation of PSI 90 with harm-based weights is feasible and results in satisfactory reliability and discrimination, with a more clinically meaningful distribution of component weights.

KEYWORDS
AHRQ, composite, harm, patient safety, PSI 90, quality indicator

What is known on this topic

- The Agency for Healthcare Research and Quality (AHRQ) Quality Indicator "Patient Safety and Adverse Events Composite" (Patient Safety Indicator [PSI] 90) factors into the value-based purchasing programs of the Centers for Medicare & Medicaid Services.
- PSI 90 was originally formulated as a volume-weighted composite indicator, with each occurrence of a component PSI event ascribed equal value.

What this study adds

- Reweighting PSI 90 to reflect not only the volume of component events but also the likelihood of patient harm associated with each type of component event is feasible and better aligns the measure with a focus on preventing patient harm.
- The harm-based weighting approach for PSI 90 results in satisfactory reliability and discrimination and more equally distributes weights across the component PSIs.

1 | INTRODUCTION

The Agency for Healthcare Research and Quality (AHRQ) Quality Indicator (QI) entitled "Patient Safety and Adverse Events Composite" (Patient Safety Indicator [PSI] 90, formerly known as the "Patient Safety Indicator Composite for Selected Indicators") was developed to provide a simple and transparent single metric that can be used to better understand, communicate, and track patient safety in US hospitals. This indicator, which has been endorsed by the National Quality Forum (NQF),[1] is currently comprised of 11 component PSIs which are calculated using readily available and routinely collected administrative data.[2,3]

Conceptually, PSI 90 is a formative composite, with each component reflecting a different aspect of patient safety. Historically, PSI 90 scores were based on a combination of the reliability-adjusted (smoothed), risk-standardized observed-to-expected (O/E) ratios for each PSI, with component weights determined by the relative national frequency of the PSI events.[4] One of the shortcomings of this approach is that it attributes an equal amount of harm to each occurrence of a PSI event, although some types of events are more harmful than others. For software version 5.0 (v5), PSI 15 ("Accidental Puncture or Laceration") and PSI 12 ("Perioperative Pulmonary Embolism or Deep Vein Thrombosis") received the majority of the weight (43.9% and 33.8%, respectively) because they are frequent events, while PSI 08 ("Postoperative Hip Fracture"), as an infrequent event, received a very small percentage (0.18%), despite greater associated short- and long-term patient suffering, complications, and mortality per event.

Therefore, we sought to reweight PSI 90 to better reflect actual harm to patients by considering both the estimated harm (as determined by utility estimates) of each PSI event and the event frequency. Our approach involved determining the harms that were relevant to each type of component PSI event, quantifying the likelihood of these harms, then aggregating this information into a summary harm for each component PSI. We then planned to determine a new harm-based PSI 90 weight for each component PSI based on the summary harm and the national volume of the PSI.

2 | METHODS

2.1 | Data sources

We used data from the Centers for Medicare & Medicaid Services (CMS) and the AHRQ Healthcare Cost and Utilization Project (HCUP). Specifically, we used the 2012 and 2013 CMS Limited Dataset (LDS) Inpatient Standard Analytic File (SAF), the LDS Outpatient SAF (Base Claims File and Revenue Center File), the LDS Skilled Nursing Facility SAF, and the LDS Denominator SAF (2012 only) to estimate the excess risk of the harm outcomes (risk difference) that occurred in association with each component indicator patient safety event.[5] These files contain diagnosis codes (International Classification of Diseases, Ninth Revision, Clinical Modification [ICD-9-CM]), procedure codes (ICD-9-CM or Healthcare Common Procedure Coding System [HCPCS] codes), service dates, cost and revenue codes, provider identifiers, and beneficiary information. We used the CMS datasets because they allow for tracking of patients over time and because Medicare beneficiaries comprise a substantial proportion (43%–80%) of hospitalizations flagged by PSI 90's component indicators.[6] We limited the CMS datasets to a subset of hospital claims to mirror those

TABLE 1 Description of patient harms captured in the revised AHRQ Patient Safety and Adverse Events Composite (PSI 90) (v6.0)

Outcome	Description of events captured	Applicable Patient Safety Indicator (PSI)											Disutility[a]
		03	06	07	08	09	10	11	12	13	14	15	
Pressure ulcer treatment	Debridement of a pressure ulcer and/or surgical soft tissue flap procedure during the hospitalization when the pressure ulcer developed due to tissue damage	X											0.009
Readmission for a pressure ulcer-related complication	Readmission to a hospital for management of a pressure ulcer or a related complication (e.g., infection, failure of a soft tissue flap) within >30–180 days after discharge	X											0.119
Intubation and mechanical ventilation	Intubation or mechanical ventilation during the hospitalization		X										0.007
Pneumothorax treatment	Chest tube placement or needle aspiration of the pleural space during the hospitalization		X										0.008
Readmission for a nonsurgical hip fracture complication	Readmission to a hospital for a mechanical or infectious hip fracture complication not requiring surgery within >30–90 days after discharge				X								0.160
Hip reoperation	Readmission to the hospital for reoperation on the hip within >30–90 days after discharge				X								0.265
Readmission for avascular necrosis	Readmission to the hospital for aseptic or avascular necrosis within >30–365 days after discharge				X								0.222
Anoxic brain damage or shock	Development of brain (cerebral) anoxia and/or shock					X							0.007[8]
Acute renal failure requiring dialysis	Development of acute kidney injury/failure (stage V) requiring dialysis while hospitalized					X				X			0.005[9,10]
One-year all-cause hospital readmission	All-cause hospital readmission within >30–365 days after discharge						X						0.084
Postdischarge dialysis	Ongoing need for dialysis postdischarge due to persistent or worsening renal failure within 180 days of discharge					X	X			X			0.326[9,10]
Extubation delay	Delay in extubation or need for reintubation during hospitalization						X						0.008
Tracheostomy	Tracheostomy during the hospitalization							X					0.114

TABLE 1 (Continued)

Outcome	Description of events captured	Applicable Patient Safety Indicator (PSI)											Disutility[a]
		03	06	07	08	09	10	11	12	13	14	15	
Readmission for a bleeding complication	Readmission to a hospital for a bleeding complication related to anticoagulation within >30–180 days after discharge								X				0.008[11]
Emergency department visit for a thrombotic complication	Emergency department visit related to a thrombotic event such as pulmonary embolus, deep vein thrombosis, or postphlebitic syndrome within >30–180 days after discharge								X				0.009[11]
Readmission for an enterocutaneous fistula	Readmission to a hospital for enterocutaneous fistula within >30–180 days after discharge										X		0.160[12]
Readmission for an incisional hernia	Readmission to a hospital (including observational stays) for an incisional hernia or reclosure of postoperative disruption of the abdominal wall within >30–180 days after discharge										X		0.075
Readmission for an intraabdominal abscess or enterocutaneous fistula	Readmission to a hospital for an intra-abdominal abscess or enterocutaneous fistula up to >30–180 days after discharge											X	0.281[12]
30-Day all-cause mortality	Death due to any cause within 30 days after discharge		X	X	X	X						X	1.000
180-Day all-cause mortality	Death due to any cause within 180 days after discharge	X					X	X	X	X	X		1.000
Excess hospital days	Excess hospital length of stay (in days)	X	X	X	X	X	X	X	X	X	X	X	0.0003/day[9]
Long-term skilled nursing facility stay	Skilled nursing facility stay ≥26 days	X	X	X	X	X	X	X	X	X	X	X	0.250
Short-term skilled nursing home days	Cumulative skilled nursing facility days within one year after discharge	X	X	X	X	X	X	X	X	X	X	X	0.0002/day
30-Day all-cause hospital readmission	All-cause hospital readmission within 30 days after discharge	X	X	X	X	X	X	X	X	X	X	X	0.010

Note: Harms are listed above even if there was no excess harm empirically observed for the outcome of interest (i.e., negative point estimate in Table S2).
[a]Citations indicate sources of related disutility information, when available. We derived disutilities using available literature-based disutilities, clinician input, and polynomial regression, as described in Section 2.

included in the HCUP dataset. This eliminated initial claims associated with noncommunity hospitals, such as rehabilitation, psychiatric, or specialty hospitals, identified using the American Hospital Association (AHA) Survey Database. The data included hospitals and providers, both paid under the Prospective Payment System and other payment mechanisms (e.g., those that apply to critical access hospitals, certain cancer-focused hospitals, and hospitals in Maryland).

The 2012 HCUP State Inpatient Databases (SID)[7] from 36 states with present on admission data (representing 82% of US community hospitals and approximately 30 million adult hospital discharges) were used to calculate the volume (count) of component patient safety events. These billing data contain diagnosis and procedure codes (ICD-9-CM), Medicare Severity Diagnosis Related Groups (MS-DRGs), revenue center codes, service days, admission source, discharge disposition, and patient demographic information.

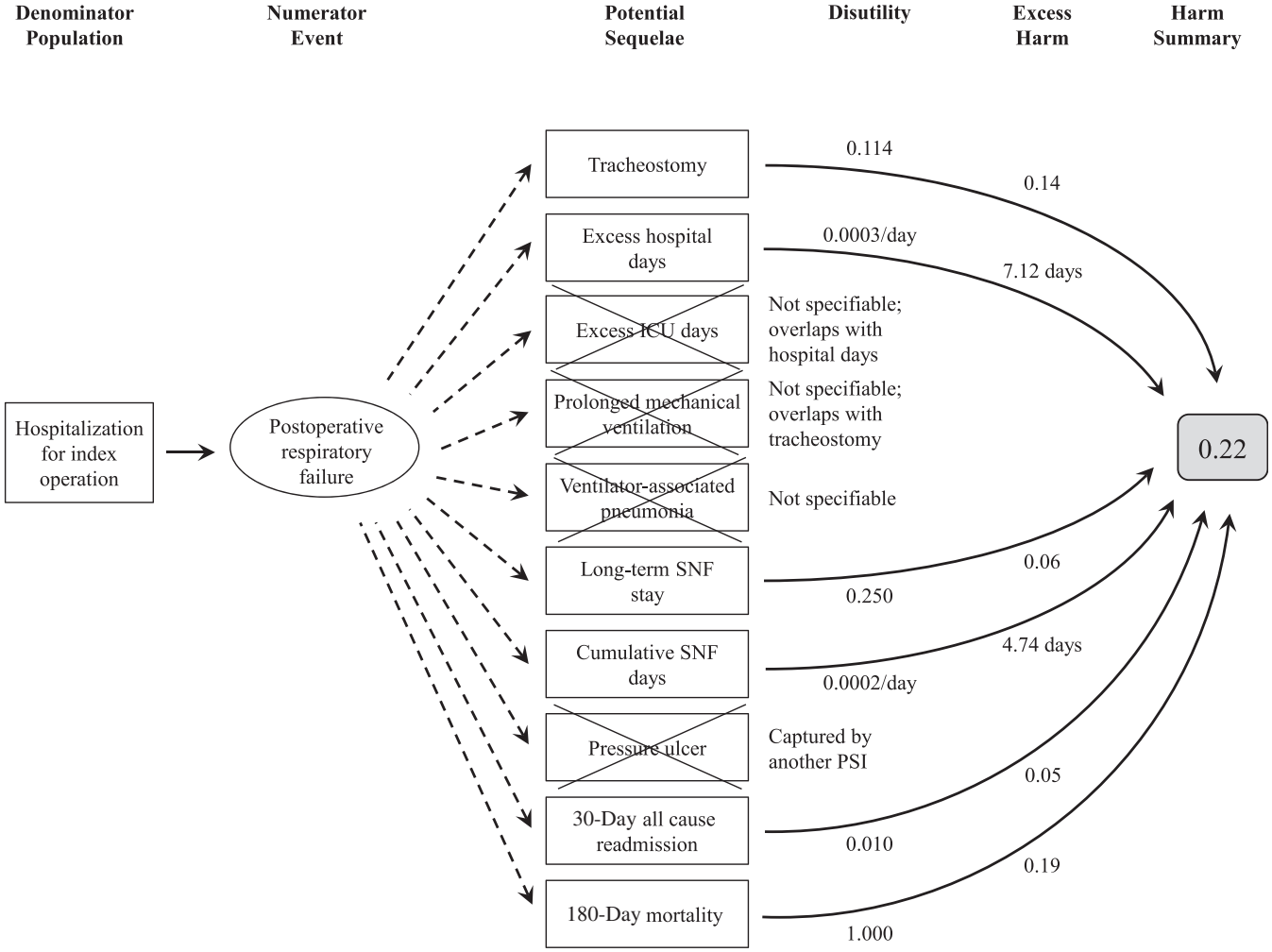

FIGURE 1 Schematic depicting the derivation of harms for a particular adverse event, combined into an overall harm summary, exemplified with Patient Safety Indicator (PSI) 11 "Postoperative Respiratory Failure." Excess harms are multiplied by their associated disutilities, then the products are summed to achieve the "harm summary." Some potential harms did not contribute to the harm summary because they cannot be specified from available data sources (i.e., Medicare fee-for-service datasets) or overlap with other harms or PSIs. *Source*: 2012 and 2013 Centers for Medicare & Medicaid Services Limited Dataset (LDS) Inpatient Standard Analytic File (SAF), the LDS Outpatient SAF (Base Claims File and Revenue Center File), the LDS Skilled Nursing Facility SAF, and the LDS Denominator SAF (2012 only)

2.2 | Defining potential harms

We convened a group of expert clinicians (who used literature review and their clinical expertise) to identify an initial set of potential harms associated with each PSI that could reasonably be specified using CMS and AHRQ HCUP administrative data sets (Table 1). We refined the initial list to eliminate overlap of certain harm states, such as transient harms that were captured by other specified harms (exemplified in Figure 1). For example, for some PSIs, short-term events such as intubation and mechanical ventilation, chest tube placement, delay in extubation, and return to the operating room are not usually associated with long-term sequelae and typically resolve by the time of hospital discharge or shortly thereafter; as such, the harm attributable to these events would be captured in the hospital length of stay or other harms. For all PSIs, we included hospital length of stay, 30-day hospital all-cause readmission, skilled nursing facility length of stay, 30- or 180-day all-cause mortality, and long-term care placement (skilled nursing facility stay ≥26 days). For PSIs limited to short-term sequelae, we used 30-day all-cause mortality; for all others, we determined all-cause mortality at 180 days. We did not include intensive care unit (ICU) use as a harm, as it was a dichotomous variable in the dataset, and we were unable to determine the timing of ICU care relative to the PSI event or calculate the excess number of ICU days.

2.3 | Excess harms

We calculated the excess harms for each PSI using 2012 and 2013 CMS datasets. We used 2012 CMS inpatient data to identify index hospitalizations and PSI events and 2012–2013 inpatient, outpatient, and skilled nursing data to identify harms subsequent to index hospitalizations. To estimate the excess harms associated with a PSI event,

we first constructed propensity models for each component PSI using the following predictors: age (years), sex, race/ethnicity (six categories), Medicare eligibility status (five categories, including aged with/without end-stage renal disease [ESRD], disabled with/without ESRD, and ESRD alone), and the log odds predictor from the PSI's standard risk-adjustment model. The last covariate incorporated age (5-year categories), sex, age-sex interactions, AHRQ/Elixhauser comorbidities, and modified Diagnosis Related Groups (DRGs), which were constructed by aggregating MS-DRGs with/without comorbidities and complications. (Our goal was to enhance the ability of the propensity scoring to account for social risk factors that are deliberately omitted from the standard PSI risk-adjustment models; we incorporated race/ethnicity and Medicare eligibility as partial adjustments for such factors.) We then weighted observations to estimate the "average treatment effect in the treated" (weight of 1 for those with the PSI event of interest and weight of the propensity odds for those without).[13,14] For each PSI, we then fit separate regression models for each harm outcome, accounting for other PSI events as confounders (to avoid exaggerating the effects of PSIs that are associated with each other, illustrated in Table S1) and clustering observations within hospitals. We used linear probability models for binary outcomes and linear models for the length of stay.

2.4 | Health utility weights

We sought to account for the relative disutility of each of the specified harm states based on previously reported utilities or disutilities. The advantage of such a utility approach is that it adopts a commonly used scale from 0 to 1 that can be converted to a harm scale (1 – utility) to weight the relative quality-of-life effects on patients of a diverse set of PSI-related harms. Events of no significance to a patient are not given any weight because there would be no disutility. We also considered harms with negative coefficients (based on the aforementioned empirical derivation) to have no disutility.

However, because validated utility values were not available in the literature for each health state, we utilized a two-step approach in which we elicited rankings of harm states from clinicians, then one investigator (KM) fit these rankings to known literature-based disutilities[8–12] using polynomial regression to estimate patient-sourced disutilities. In the first step, we convened a convenience group of clinicians (three nurses and eight physicians), each of whom ranked, in a blinded fashion, a predetermined list of all harms we considered across the PSIs. In the second step, we normalized rankings elicited from the clinicians (rank divided by the total number of harm states) and averaged these values across the 11 clinicians. For health states that had utility values reported in the literature, we derived a best fit polynomial regression equation, similar to a previously described approach,[15] to express the empirically observed relationship between the elicited average normalized ranking (X; with values within the range 0–1) to utility values (Y; also with rounded values within the range 0–1):

$$Y = (1.74 \times X^3) - (3.72 \times X^2) + (2.98 \times X) + 0.004.$$

We then used this equation to estimate utility values (Y) for harm states without literature-based utilities available. We conducted a final review of harm states and the calculated utilities with input from the clinical ranking group to identify any concerns with face validity. This final step led to modifying several disutilities considered too high for the relative duration of the harm state. In our approach, each disutility represents a summary estimate of the cumulative burden on the patient's quality of life applicable to their remaining life expectancy. Thus, an outcome that had negligible impact on the patient's quality of life, either because the harm was (1) persistent but minuscule or (2) clinically significant but very transient, would have a disutility of 0 or near 0. Conversely, inpatient death would correspond to a disutility of 1. Importantly, we decided that many short-term harms would be encompassed by the harm associated with short-term hospitalization (i.e., length of stay) and thus did not warrant retention as separate harms. Harms that fall between the extremes of transient events during the index hospitalization and death, such as the prolonged need for dialysis or a prolonged stay at a skilled nursing facility, were of particular interest. When possible (i.e., for lengths of stay), we calculated these prolonged harms as having a constant per-day burden on the patient's quality of life. As with decision modeling applications of utilities, we incorporated the concept of tolls for transient harm states (e.g., each day in intensive care is equivalent to a disutility equal to death for that day). Because we estimated harms from Medicare data, we used the average age of Medicare beneficiaries (71 years), coupled with the average life expectancy for a person of that age (79 years), to determine the expected years of additional life that could potentially be lost (8 years) as the basis for tolls.

2.5 | Component indicator weights

For each component indicator in the PSI 90 composite, we estimated two sets of values. The first was the excess risk of the harm outcomes (risk difference) that occurred in association with the indicator patient safety event (described in Section 2.3). We multiplied these by the disutility scores to obtain the harm weights. The second was the set of PSI numerator weights, which we calculated from the volume (count) of component events in an HCUP US reference population. We calculated the final weight for each component PSI indicator, q, that is part of PSI 90 as:

$$\text{Weight}_q = \frac{\text{volume}_q \sum_{h=1}^{H} \text{harm}_{qh}\, \text{disutility}_{qh}}{\sum_{q=1}^{Q} \left(\text{volume}_q \sum_{h=1}^{H} \text{harm}_{qh}\, \text{disutility}_{qh} \right)},$$

where:

Q is the total number of component quality indicators, q, in PSI 90.

H is the total number of outcome types (harms), *h*, related to each component indicator.

Volume is the numerator count or the number of total PSI events within the component indicator in the reference population.

Harm is the excess risk (risk difference) of each type of outcome (i.e., harm) within each component indicator estimated from a model comparing people with PSI events to those without PSI events in an "at risk" cohort.

Disutility is the complement of a utility weight (1 – utility_wt) assigned to each excess occurrence of each type of outcome within each component indicator.

2.6 | Hospital scoring

We calculated hospital-specific PSI 90 scores using the 2012 HCUP databases. Risk models for individual PSI events included age, sex, the reason for hospitalization (modified DRG), comorbidities (modified AHRQ's Elixhauser Comorbidity Software[16]), transfer-in status, point of origin, and days to the procedure to create a parsimonious set of covariates, retaining only those that were significant predictors ($p < 0.05$). For each component PSI, we determined an O/E ratio based on that measure's risk model. To account for the uncertainty (noise) in a hospital's performance due to reliability concerns stemming from low volume, we calculated smoothed rates using the approach described in the AHRQ document "Quality Indicator Empirical Methods."[17]

We considered fewer than three qualifying records per hospital for an individual PSI component denominator as missing information. Whenever a component was missing, we substituted a value of 1.0 for the smoothed O/E ratio. This is consistent with the individual PSI models, which also shrink O/E ratios to a target of 1.0. Thus, a hospital with missing components is neither rewarded nor penalized for missing information regarding the PSIs. Additionally, the variance calculations incorporate the uncertainty associated with this imputation method.

We repeated the analyses using a prerelease edition of version 6.0 (v6.0) of the AHRQ software.[18] There were several specification changes from v5.0 to 6.0 with impacts on PSI 90 weights independent of the implementation of the harms-based weighting approach (see Supplemental Methods). Based on requests from users, we modeled a version of PSI 90 with and without PSI 07 ("Central Line Related Bloodstream Infection"), in part due to overlap with a similar indicator administered by the Centers for Disease Control and Prevention National Healthcare Safety Network, used in several federally mandated programs.

For most empirical testing, including reliability and discrimination tests, we used HCUP data from 2012 to 2013. To describe the distribution of hospital-level observed rates of PSI 90, we used 2011–2013 HCUP data. We assessed the reliability of the component PSIs using signal-to-noise ratios, comparing the degree to which rates were different from hospital to hospital (the signal) to how stable the rates were within hospitals (the noise).[17] We evaluated the discrimination of each component PSI using the concordance statistic ("*c*-statistic") to measure discrimination, defined as how well the risk adjustment model distinguishes events from nonevents. We computed the *c*-statistic by assigning each observation a predicted probability of the outcome from the risk-adjustment model based on the values of the observation's covariates. The *c*-statistic is a measure of the proportion of discordant pairs (one event and one nonevent) of observations for which the observation with the event had a higher predicted probability from the risk-adjustment model than the nonevent.

To assess whether each of the component measures contributed to PSI 90 as an overall construct of hospital-level quality of care, we

TABLE 2 Patient, hospital, and discharge counts for the AHRQ Patient Safety and Adverse Events Composite (PSI 90 v6.0), by component indicators

Patient Safety Indicator	Patient Safety Indicator title	Individual records (*n*)	Hospitals (*n*)	Patient stays (*n*)
03	Pressure Ulcer	2,367,595	4615	3,153,685
06	Iatrogenic Pneumothorax	6,252,940	4702	9,947,207
07	Central Venous Catheter-Related Blood Stream Infection	4,276,737	4697	6,303,846
08	In-Hospital Fall With Hip Fracture Rate	1,343,305	4117	1,531,045
09	Perioperative Hemorrhage or Hematoma	2,226,409	4154	2,532,614
10	Postoperative Acute Kidney Injury Rate	1,330,871	3886	1,449,502
11	Postoperative Respiratory Failure	1,066,909	3863	1,151,960
12	Perioperative Pulmonary Embolism or Deep Vein Thrombosis	2,393,597	4167	2,746,255
13	Postoperative Sepsis	284,142	3580	296,483
14	Postoperative Wound Dehiscence	296,139	3798	315,973
15	Unrecognized Abdominopelvic Accidental Puncture/Laceration Rate	383,362	3835	410,698
Total		6,529,709	4705	10,552,935

Source: 2012 Centers for Medicare & Medicaid Services (CMS) Limited Dataset Inpatient and Denominator Standard Analytic Files.

TABLE 3 Excess harms for AHRQ Patient Safety and Adverse Events Composite (PSI 90) (v6.0), by component indicators

Harm	Excess harm	95% lower confidence limit	95% upper confidence limit
PSI 03 (Pressure Ulcer)			
Pressure ulcer treatment	0.0493	0.0395	0.0590
30- to 180-day readmission for a pressure ulcer complication	0.0737	0.0590	0.0885
180-Day mortality	0.2696	0.2478	0.2913
Excess hospital days	9.2831	7.9464	10.6197
Long-term skilled nursing facility stay greater than or equal to 26 days	0.0930	0.0726	0.1135
Cumulative skilled nursing facility days	8.8465	7.1941	10.4990
30-Day all-cause readmission	0.0504	0.0350	0.0658
PSI 06 (Iatrogenic Pneumothorax)			
Intubation and ventilation	0.2090	0.1936	0.2245
Pneumothorax treatment	0.6351	0.6188	0.6515
30-Day mortality	0.1300	0.1156	0.1444
Excess hospital days	4.6300	4.2879	4.9720
Long-term skilled nursing facility stay greater than or equal to 26 days	−0.0072	−0.0195	0.0051
Cumulative skilled nursing facility days	−0.7479	−1.6253	0.1294
30-Day all-cause readmission	−0.0124	−0.0242	−0.0006
PSI 07 (Catheter-Related Blood Stream Infection)			
30-Day mortality	0.1329	0.1114	0.1544
Excess hospital days	18.8503	17.9183	19.7822
Long-term skilled nursing facility stay greater than or equal to 26 days	0.1143	0.0876	0.1410
Cumulative skilled nursing facility days	8.2946	6.1763	10.4129
30-Day all-cause readmission	0.0829	0.0599	0.1058
PSI 08 (Postoperative Hip Fracture)			
Nonsurgical fracture complication within 30 to 90 days	−0.0004	−0.0005	−0.0003
Hip reoperation due to complication within 30 to 90 days	0.0012	−0.0037	0.0060
30- to 365-day readmission for avascular necrosis	−0.0001	−0.0018	0.0017
30-Day mortality	0.0731	0.0491	0.0971
Excess hospital days	4.5179	3.8248	5.2110
Long-term skilled nursing facility stay greater than or equal to 26 days	0.2534	0.2200	0.2868
Cumulative skilled nursing facility days	18.6191	15.9928	21.2455
30-Day all-cause readmission	0.0905	0.0629	0.1181
PSI 09 (Perioperative Hemorrhage or Hematoma)			
New onset renal failure requiring dialysis	0.0950	0.0863	0.1038
Complication of anoxia and/or shock	0.1069	0.0994	0.1144
180-Day persistent or worsening renal failure or related complications	−0.0068	−0.0111	−0.0025
30-Day mortality	0.0452	0.0398	0.0505
Excess hospital days	5.0801	4.8218	5.3384
Long-term skilled nursing facility stay greater than or equal to 26 days	0.0307	0.0239	0.0376
Cumulative skilled nursing facility days	2.5312	2.0154	3.0470
30-Day all-cause readmission	0.0471	0.0397	0.0546

(Continues)

TABLE 3 (Continued)

Harm	Excess harm	95% lower confidence limit	95% upper confidence limit
PSI 10 (Postoperative Metabolic and Physiologic Derangement)			
Extubation delay	0.6170	0.5928	0.6413
180-Day persistent or worsening renal failure or related complications	0.0428	0.0251	0.0605
365-Day all-cause readmission	−0.1107	−0.1311	−0.0902
180-day mortality	0.3269	0.2939	0.3599
Excess hospital days	11.3501	10.4157	12.2845
Long-term skilled nursing facility stay greater than or equal to 26 days	0.0294	0.0057	0.0531
Cumulative skilled nursing facility days	1.8030	0.1227	3.4833
30-Day all-cause readmission	0.0632	0.0360	0.0904
PSI 11 (Postoperative Respiratory Failure)			
Tracheostomy	0.1399	0.1310	0.1488
180-Day mortality	0.1861	0.1769	0.1954
Excess hospital days	7.1257	6.8263	7.4252
Long-term skilled nursing facility stay greater than or equal to 26 days	0.0626	0.0559	0.0692
Cumulative skilled nursing facility days	4.7455	4.2563	5.2347
30-Day all-cause readmission	0.0517	0.0447	0.0586
PSI 12 (Perioperative Pulmonary Embolism or Deep Vein Thrombosis)			
30- to 180-day readmission for bleeding complications	0.0108	0.0067	0.0150
180-Day emergency department visit for thrombotic complications	0.4757	0.4595	0.4919
180-Day mortality	0.1343	0.1257	0.1429
Excess hospital days	8.0335	7.7180	8.3489
Long-term skilled nursing facility stay greater than or equal to 26 days	0.0503	0.0423	0.0582
Cumulative skilled nursing facility days	3.8904	3.3041	4.4768
30-Day all-cause readmission	0.0575	0.0495	0.0654
PSI 13 (Postoperative Sepsis)			
New onset renal failure requiring dialysis	0.4435	0.4315	0.4556
180-Day persistent or worsening renal failure or related complications	−0.0226	−0.0278	−0.0175
180-Day mortality	0.2857	0.2717	0.2998
Excess hospital days	12.0275	11.4729	12.5821
Long-term skilled nursing facility stay greater than or equal to 26 days	0.0650	0.0544	0.0755
Cumulative skilled nursing facility days	4.8970	4.1201	5.6740
30-Day all-cause readmission	0.0479	0.0365	0.0593
PSI 14 (Postoperative Wound Dehiscence)			
30- to 180-day readmission for incisional hernia	0.0138	0.0044	0.0233
30- to 180-day readmission for enterocutaneous fistula	0.0174	0.0057	0.0290
180-Day mortality	0.1076	0.0758	0.1393
Excess hospital days	12.1829	11.1740	13.1918
Long-term skilled nursing facility stay greater than or equal to 26 days	0.1020	0.0719	0.1321
Cumulative skilled nursing facility days	8.2244	5.9132	10.5356
30-Day all-cause readmission	0.0839	0.0529	0.1148

TABLE 3 (Continued)

Harm	Excess harm	95% lower confidence limit	95% upper confidence limit
PSI 15 (Accidental Puncture or Laceration)			
30- to 180-day readmission for intra-abdominal abscess or enterocutaneous fistula	0.0939	0.0622	0.1255
30-Day mortality	0.1006	0.0572	0.1441
Excess hospital days	14.1948	12.0065	16.3831
Long-term skilled nursing facility stay greater than or equal to 26 days	0.0543	0.0118	0.0968
Cumulative skilled nursing facility days	4.7647	1.5702	7.9591
30-Day all-cause readmission	0.0732	0.0274	0.1191

Source: 2012 and 2013 Centers for Medicare & Medicaid Services Limited Dataset (LDS) Inpatient Standard Analytic File (SAF), the LDS Outpatient SAF (Base Claims File and Revenue Center File), the LDS Skilled Nursing Facility SAF, and the LDS Denominator SAF (2012 only).

calculated Pearson and Spearman (rank) correlations between each hospital's PSI 90 scores and its smoothed, risk-adjusted component PSI rates. We also evaluated pairwise Spearman rank-correlations among the risk-adjusted component indicator rates to determine whether performance in one indicator generally tracked with performance in the other indicators.

To evaluate the effects of uncertainty in disutility estimates on hospital scoring, we conducted a sensitivity analysis in which we varied the disutility estimates randomly up to ±15% from the point estimate for each disutility, repeating these perturbations in 1000 simulations.

2.7 | PSI 90 reliability

Using the 2012 HCUP databases, we evaluated the reliability of PSI 90 using test–retest analysis and signal-to-noise ratios. We assessed test–retest reliability by randomly dividing records at individual hospitals into two equal sized groups, applying the measure to each group separately, and calculating an intraclass correlation coefficient (ICC).[19] To adjust for the reduction in sample size (resulting from the set of observations being split into two), we used the Spearman–Brown Prophecy formula[20] to estimate the reliability after changing the number of items (sample size).

Among the entire population of hospitals in the database, we calculated the signal-to-noise ratio for each hospital. We then aggregated these into an overall signal-to-noise estimate weighted by hospital size, in turn, calculated as a weighted average of the denominators (at-risk group) of the component PSIs (equal to the final PSI 90 weight).[17] Weighting by hospital size reduces the impact of hospitals that have very small denominators (numbers of patients at risk).

2.8 | PSI 90 discrimination

We used the *c*-statistic to measure the discrimination of PSI 90. We did not employ common "goodness of fit" tests because these tests tend not to be informative with large samples.

To determine if the indicator can discriminate the best performing hospitals from the poorest performing ones, we assessed the probability that a hospital was higher or lower than a benchmark or threshold, given hospital size. For this analysis, "benchmark" refers to the 20th percentile of the indicator rate estimated for the reference population (i.e., 20% of hospitals have a lower complication rate or better performance). "Threshold" refers to the 80th percentile of the indicator rate estimated for the reference population (i.e., 80% have lower complications or better performance). A gamma distribution was estimated to establish the 20th and 80th percentile marks for the reference population.

We then estimated a gamma distribution separately for each hospital to compute the 95% confidence limits for each hospital's smoothed rate. We ranked hospitals on size and grouped them into 10 equal categories of size (deciles). We compared the benchmark and threshold rates to the gamma distribution of the smoothed rates for each hospital to determine whether the hospital rate was better or worse than the benchmark and threshold rates with 95% probability.

3 | RESULTS

The aggregate CMS dataset for determining harms related to PSI 90 included 6,529,709 patients, 4705 hospitals, and 10,552,935 hospital stays (Table 2). Denominator numbers by PSI indicator for 2012 ranged from 284,142 to 6,252,940.

Individual harms differed by component PSI (Table 1), and excess harms varied across the PSIs (Table 3). In 1000 simulated perturbations of the disutility values, over 98% of hospitals remained in the same quartile of performance for PSI 90. Final weights, including the harm weight (excess harm and disutility) and volume weights for each PSI, varied across the PSIs from 0.0068 for PSI 15 to 0.2972 for PSI 11 (Table 4). Harm-based weights roughly corresponded to previous volume-based weights (v5.0) except for PSIs 12 and 15 (which underwent substantial modification), as well as PSI 13. Compared to hypothetical volume-based weights incorporating changes to the component indicator specifications in v6.0 (i.e., aside from the

TABLE 4 Comparison of volume-based weights to harm-based weights and related inputs for PSI 90 (implemented and theoretical versions), by component indicators

	PSI 03	PSI 06	PSI 07	PSI 08	PSI 09	PSI 10	PSI 11	PSI 12	PSI 13	PSI 14	PSI 15
Version 5.0 specifications											
Volume-based weight	0.0330	0.0751	0.0377	0.0018	–	–	–	0.3379	0.0573	0.0182	0.4390
Version 6.0 specifications											
Harm summary	0.308	0.138	0.171	0.144	0.057	0.358	0.222	0.156	0.310	0.144	0.147
Volume (numerator)	2957	5898	2305	1067	22,792	1742	20,918	20,429	10,600	1402	724
Final harm-based weight (v6.0) with PSI 07	0.0583	0.0521	0.0252	0.0098	0.0832	0.0400	0.2972	0.2037	0.2106	0.0129	0.0068
Final harm-based weight without PSI 07 (v6.0) (NQF endorsed)	0.0598	0.0535	–	0.0101	0.0853	0.0410	0.3049	0.2090	0.2160	0.0133	0.0070
Hypothetical v6.0-concordant volume-based weight without PSI 07[a]	0.0334	0.0666	–	0.0120	0.2574	0.0197	0.2363	0.2308	0.1197	0.0158	0.0082

[a]For illustrative purposes only. Volume-based weights are not used for PSI 90 v6.0.

Source for volume: Healthcare Cost and Utilization Project (HCUP) State Inpatient Databases (SID), 2012–2013. Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (www.hcup-us.ahrq.gov/sidoverview.jsp) (subset of 36 states with "present on admission" information). *Source for harm summary:* Centers for Medicare & Medicaid Services (CMS) Inpatient and Outpatient Medicare Fee-For-Service data in the 100% standard analytical files (SAF). A separate cohort sample was drawn for each component PSI that was defined by that PSI's denominator criteria among inpatient stays in 2012. We used data from 2013 for follow-up only, to ensure that a full 365 days of follow-up were available for each observation.

harm-based weighting), PSI 09 and to a lesser degree, PSI 12 were down-weighted while PSIs 10, 11, and 13 were up-weighted.

The mean (standard deviation) distribution of observed hospital-level PSI 90 rates per 1000 for pooled 2011–2012 and 2012–2013 data, respectively, were 1.016 (0.193) and 0.996 (0.195) (Table S2). A majority of hospitals (70.3%) had ≥3 denominator-eligible hospitalizations for each of the component PSIs (Table S3). The most frequent missing PSI components were PSIs 10 (22.16%), 11 (22.46%), and 13 (27.69%) (Table S4), presumably reflecting hospitals that do not routinely offer surgical procedures or systematically misclassify elective status.

The signal-to-noise ratios of the component PSIs ranged from 0.3743 to 0.7564, and the *c*-statistics ranged from 0.7195 to 0.9260 (Table S5). The component PSIs were positively correlated with the composite score (Table S6) and with one another (Table S1). The three PSIs that were not previously components of PSI 90 (PSI 09, PSI 10, and PSI 11) had weighted Pearson correlation coefficients of 0.527, 0.247, and 0.479 and Spearman rank correlation coefficients of 0.484, 0.243, and 0.460, respectively.

3.1 | PSI 90 reliability

Regarding test–retest reliability, both the calculated and approximated (using the Spearman–Brown adjustment) ICC statistics for the overall composite were 0.76. Overall, the average signal-to-noise ratio ranged from 0.3436 for the smallest hospital decile to 0.8598 for the largest hospital decile, with an overall US average of 0.7015 (Table S7).

3.2 | PSI 90 discrimination

Among hospitals in the upper half of the volume distribution (i.e., 1875 of the 3749 tested hospitals), 34% were significantly better than the threshold value. Similarly, 41% were significantly worse than the benchmark value (Table S7). The number of statistically significant outliers was limited by the relatively tight distribution of performance. Hospitals that were random outliers on one or two of the component indicators were not outliers on the overall composite.

4 | DISCUSSION

Overall the new proposed approach to weighting appears to represent an important advance over the previous weighting approach and better aligns PSI 90 with the concept of "freedom from harm."[21] The new approach weights component indicators more when PSI-attributable harms are either more frequent or more burdensome to the patients who experience them. The weights in the new scheme are more evenly balanced across PSIs so that no single indicator carries more than one-third of the total weight. PSI events associated with worse health consequences such as PSIs 10 ("Postoperative Acute Kidney Injury"), 11 ("Postoperative Respiratory Failure"), and 13 ("Postoperative Sepsis") are now weighted accordingly, and events

with relatively modest consequences, on average, such as PSI 09 ("Perioperative Hemorrhage or Hematoma"), are down-weighted compared to a purely volume-based weighting. Overall, the new approach weights more harmful events more heavily than in the original composite, which reflected the frequency of events, regardless of harm. Partially by virtue of the focus on harms, PSI events that are easy to "prevent" by changes in coding practice, such as PSI 15 (v5.0), are no longer weighted more heavily than events that require careful attention to processes of care to achieve the same relative reduction in frequency, but a real reduction in harm. To the extent that some PSIs have false positives, flagged nonevents theoretically should reduce the corresponding harm-based weights.

Discrimination results demonstrate that PSI 90 can detect statistically meaningful differences across hospitals. Hospitals that are random outliers on one or two of the component indicators are no longer outliers on the overall composite. These findings are expected, but they do suggest that, for low-volume hospitals, PSI 90 should be used cautiously—or multiple years of data should be combined.

PSI 90 (v6.0) has strong reliability as measured by the high ICC of split samples (0.76 by both methods) and high signal-to-noise ratios (US average being 0.70). The high test–retest reliability indicates that the measure is capturing a hospital trait rather than potentially transient performance. Although a formal threshold for "adequate" reliability does not exist, typically, an ICC reliability score of 0.4 is considered moderate, much lower than the 0.76 obtained. Reliability measured by signal-to-noise ratios was sufficient across all size deciles indicating strong measurement precision, although reliability increased with hospital size. The AHRQ QI program generally considers signal-to-noise ratios between 0.4 and 0.8 as acceptable, although when a complication is very important (e.g., leads to great harm to the patient), lower reliability may be acceptable. It is rare to achieve reliability above 0.8. All but the lowest three deciles of hospital size exceeded the lower limits of acceptability.

PSI 90 scores are reliability-adjusted, that is, the risk-adjusted rates of hospitals with low signal-to-noise ratios are weighted more toward the overall population rate, whereas those of hospitals with high signal-to-noise ratios are weighted more toward the O/E ratio of that hospital. Nonetheless, other strategies to improve the reliability and discrimination of PSI 90 may be necessary. For example, PSI 90 could incorporate greater component volume thresholds to exclude low-volume hospitals with less reliably measured composite scores, and greater consideration might be given to pooling data over multiple years for low-volume hospitals.

4.1 | Limitations

Despite the conceptual and empirical attractiveness of PSI 90, as re-specified, we recognize several limitations. The updated method is more complex than the prior approach, and is therefore, not as easily understood by users. Resources constrained us in estimating utilities based on only a relatively small convenience sample of clinicians and information from a literature review. Utility assessment is more reliable when it involves a large and diverse panel of raters, including patients and clinicians from a wide variety of specialties, types of practices, and geographic areas. Although values were rescaled based on patient-reported utilities from the peer-reviewed literature, further work is necessary to refine these utility estimates, including ensuring that they are internally consistent, account for overlap between harms and the time-dependent nature of some harms, and reflect actual patient experiences as accurately as possible. Some relevant harms may have been omitted because there was no published information linking these harms to a prior PSI or because of limitations in the available data or analytic methods. Additionally, we did not attempt to adjust disutilities to reflect specific characteristics, such as age or comorbidities, of the patients who experienced the corresponding PSI events. Nonetheless, from a face validity standpoint, members of an expert workgroup[22] and NQF panelists[23] evaluated and agreed with the harms we identified. We assumed independence of multiple harms resulting from the same PSI event. It is possible that some PSIs are over-weighted because the stated harms were more encompassing compared to others for which the harms may be under-represented. The propensity score-based models we used to estimate the excess harms associated with each component PSI accounted for potentially confounding patient characteristics only through the propensity scores and did not also include these characteristics as covariates in the models; as a result, there may have been residual confounding. Additionally, testing was limited to 1-year data due to the limited availability of "present on admission" data. Finally, components weights and volumes may differ with the implementation of the International Classification of Diseases, Tenth Revision, Clinical Modification and Procedure Coding System (ICD-10-CM/PCS). Updated analyses involving such recent data, including evaluation of the construct validity of the harms-based weighting of PSI 90, are in progress.

4.2 | Conclusions

The new PSI 90 weighting approach—which accounts for both the frequency of harms associated with each potentially preventable patient safety event, as well as the severity (disutility) of those harms, achieving more optimal weighting among the component measures—is feasible and results in satisfactory reliability and discrimination, with a more clinically meaningful distribution of component weights. This new approach better aligns PSI 90 with the focus on preventing patient harm and not just the occurrence of safety events. PSI 90 (v6.0) may send a clearer signal to consumers, purchasers, and providers of health care, supporting better decision-making and resource allocation.

4.3 | Implications for policy or practice

PSI 90 is a high-profile quality metric used by providers, payers, and other stakeholders in programs such as the CMS Hospital Value-Based Purchasing Program and the Hospital-Acquired Conditions Reduction

Program. The revised weighting approach offers a better measure of harms experienced by patients with potentially preventable complications, supporting performance comparisons based on hospitals' success at keeping patients safe from these harms.

ACKNOWLEDGMENTS

This article was supported by Contract #HHSAA290201200003I from the Agency for Healthcare Research and Quality (AHRQ). Dr. Owens is an employee of AHRQ. Drs. Utter and Romano served on an AHRQ Quality Indicators Expert Workgroup, for which they received honoraria, and served as consultants to AHRQ for its Healthcare Cost and Utilization Project (HCUP). The views expressed in this article are those of the authors and do not necessarily reflect those of AHRQ or the US Department of Health and Human Services. No other disclosures. We acknowledge the HCUP Partner organizations that participated in the HCUP State Inpatient Databases: Arizona Department of Health Services, Arkansas Department of Health, California Office of Statewide Health Planning and Development, Colorado Hospital Association, Florida Agency for Health Care Administration, Georgia Hospital Association, Hawaii Laulima Data Alliance, Hawaii University of Hawai'i at Hilo, Illinois Department of Public Health, Indiana Hospital Association, Iowa Hospital Association, Kansas Hospital Association, Kentucky Cabinet for Health and Family Services, Maine Health Data Organization, Maryland Health Services Cost Review Commission, Massachusetts Center for Health Information and Analysis, Michigan Health & Hospital Association, Minnesota Hospital Association (provides data for Minnesota and North Dakota), Mississippi State Department of Health, Montana MHA—An Association of Montana Health Care Providers, Nebraska Hospital Association, Nevada Department of Health and Human Services, New Hampshire Department of Health & Human Services, New Jersey Department of Health, New Mexico Department of Health, New York State Department of Health, North Carolina Department of Health and Human Services, North Dakota (data provided by the Minnesota Hospital Association), Oklahoma State Department of Health, Oregon Association of Hospitals and Health Systems, Oregon Health Policy and Research, Pennsylvania Health Care Cost Containment Council, Rhode Island Department of Health, South Carolina Budget & Control Board, South Dakota Association of Healthcare Organizations, Tennessee Hospital Association, Texas Department of State Health Services, Vermont Association of Hospitals and Health Systems, Virginia Health Information, Washington State Department of Health, and Wisconsin Department of Health Services.

ORCID

Patricia A. Zrelak https://orcid.org/0000-0002-5977-6503
Garth H. Utter https://orcid.org/0000-0001-7747-3429
Kathryn M. McDonald https://orcid.org/0000-0002-4049-5744
Sheryl M. Davies https://orcid.org/0000-0001-6250-2565
Halcyon G. Skinner https://orcid.org/0000-0002-0220-9110
Pamela L. Owens https://orcid.org/0000-0002-9905-9664
Patrick S. Romano https://orcid.org/0000-0001-6749-3979

REFERENCES

1. National Quality Forum. Accessed September 22, 2021. https://www.qualityforum.org/QPS/0531
2. PSI 90 Fact Sheet. U.S. Agency for Healthcare Research and Quality website. Accessed September 20, 2021. https://qualityindicators.ahrq.gov/Archive/News2016.aspx
3. U.S. Agency for Healthcare Research and Quality website. Accessed September 20, 2021. http://qualityindicators.ahrq.gov/Modules/psi_resources.aspx
4. Geppert J, Rhoda D, Morara M. *Quality Indicator Empirical Methods* (Revised by Battelle, under Contract No. HHSA290201200001C). U.S. Agency for Healthcare Research and Quality; 2013.
5. Centers for Medicare & Medicaid Services website. Accessed September 20, 2021. http://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/LimitedDataSets/StandardAnalyticalFiles.html
6. U.S. Agency for Healthcare Research and Quality. Quality Indicators website. Accessed September 20, 2021. https://www.qualityindicators.ahrq.gov/Downloads/Modules/PSI/V2019/Version_2019_Benchmark_Tables_PSI.pdf
7. U.S. Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project website. Accessed September 20, 2021. https://hcup-us.ahrq.gov/db/state/siddbdocumentation.jsp
8. Pepper PV, Owens DK. Cost-effectiveness of the pneumococcal vaccine in healthy younger adults. *Med Decis Making*. 2002;22(5 Suppl):S45-S57.
9. Sackett DL, Torrance GW. The utility of different health states as perceived by the general public. *J Chronic Dis*. 1978;31(11):697-704.
10. Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. *Med Care*. 2000;38(6):583-637.
11. Duriseti RS, Shachter RD, Brandeau ML. Value of quantitative D-dimer assays in identifying pulmonary embolism: implications from a sequential decision model. *Acad Emerg Med*. 2006;13(7):755-766.
12. Lehmann LE, Herpichboehm B, Kost GJ, Kollef MH, Stuber F. Cost and mortality prediction using polymerase chain reaction pathogen detection in sepsis: evidence from three observational trials. *Crit Care*. 2010;14(5):R186.
13. SAS Institute website. Accessed September 20, 2021. https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.3/statug/statug_psmatch_details05.htm
14. Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychol Methods*. 2010;15(3):234-249.
15. Dolan P, Gudex C, Kind P, Williams A. *A Social Tariff for EuroQol: Results from a UK General Population Survey*. Centre for Health Economics Discussion Paper 138. Centre for Health Economics; 1995.
16. U.S. Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project website. Accessed September 20, 2021. https://hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp
17. Quality Indicator Empirical Methods. Prepared for Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. Revised July 2021. Accessed September 20, 2021. https://www.qualityindicators.ahrq.gov/Downloads/Resources/Publications/2021/Empirical_Methods_2021.pdf
18. U.S. Agency for Heathcare Research and Quality. Quality Indicators website. Accessed September 20, 2021. https://qualityindicators.ahrq.gov/Archive/PSI_TechSpec_ICD09_v60.aspx
19. Adams JL. *The Reliability of Provider Profiling: A Tutorial*. RAND Corporation; 2009 Accessed September 20, 2021. http://www.rand.org/pubs/technical_reports/TR653.html
20. DeSimone JA, LeBreton JM. Attenuation, correction for. In: Frey B, ed. *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*. Vol 1–4. SAGE Publications; 2018:138-139.
21. Berwick D, Shojania K, Atchinson B. *Free from Harm: Accelerating Patient Safety Improvement Fifteen Years after To Err Is Human*. National Patient Safety Foundation; 2015.
22. Federal Register website. Accessed September 20, 2021. https://www.federalregister.gov/articles/2013/01/28/2013-01348/ahrq-standing-workgroup-for-(quality-indicator-measure-specification

23. National Quality Forum Patient Safety 2015. Final Technical Report. Accessed April 21, 2021. http://www.qualityforum.org/Publications/2016/02/Patient_Safety_2015_Final_Report.aspx

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Zrelak PA, Utter GH, McDonald KM, et al. Incorporating harms into the weighting of the revised Agency for Healthcare Research and Quality Patient Safety for Selected Indicators Composite (Patient Safety Indicator 90). *Health Serv Res*. 2022;57(3):654-667. doi:10.1111/1475-6773.13918