Lessons from the AHRQ PSI Validation Pilot Project

AHRQ QI Users Meeting
September 10, 2008

Presenter: Patrick S. Romano, MD MPH; Professor, UC Davis Center for Healthcare Policy and Research

Team: Mamatha Pancholi and Marybeth Farquhar (AHRQ); Jeffrey Geppert and Teresa Schaaf (Battelle Memorial Institute); Sheryl Davies and Kathryn McDonald (Stanford); Garth Utter, Richard White, Daniel Tancredi, Patricia Zrelak, Ruth Baron, and Banafsheh Sadeghi (UC Davis)

Outline

- What do we mean by "validation"?
- Goals
- Methods
- Results
- Future plans
- Implications

Validation

A valid measure accurately represents the true state of the phenomenon being measured (i.e., "free of systematic error").

- Content (aka consensual) validity is the degree to which a measure "on its face" adequately samples all relevant domains of the concept of interest.

- Criterion (aka concurrent) validity is the degree to which a measure generates data that agree with data from a better ("gold standard") approach to measuring the same characteristic.

- Predictive validity is the degree to which a measure successfully predicts an outcome of interest.

- Construct (aka convergent/discriminant) validity is the degree to which a measure correlates with other measures, based on a construct that is grounded in prior literature or a sound conceptual framework

Content validity established

- Modified RAND/UCLA Appropriateness Method
- Physicians of various specialties/subspecialties, nurses, other professionals (e.g., midwife, pharmacist) were nominated
- Each potential indicator was assigned to one or two panels
- All panelists rated all assigned indicators (1-9) on overall usefulness and other dimensions
- Pre-conference ratings and comments were collected
- Individual ratings were returned to panelists with distribution of ratings and other panelists' comments
- Telephone conference call(s) focused on high-variability items and panelists' suggestions, which were adopted only by consensus
- Post-conference ratings and comments were collected
- Excluded indicators rated "Unclear," "Unclear-," or "Unacceptable":
  - Median score <7, OR
  - At least 2 panelists rated the indicator in each of the extreme 3-point ranges

Content validity

- 48 indicators reviewed in total
  - 37 reviewed by multispecialty panel
  - 15 of those reviewed by surgical panel
- 20 "accepted" based on face validity
  - 2 dropped due to operational concerns
- 17 "experimental" or promising indicators
- 11 rejected

OECD Expert panel review of PSI usefulness and "preventability"[c]

Acceptable	Acceptable (-)	Unclear	Unclear (-)
Decubitus ulcer	Complications of anesthesia	Death in low mortality DRG	Failure to rescue
Foreign body left in	Selected infections due to medical care	Postop hemorhage/ hematoma	Postop physiologic/ metabolic derangement
Iatrogenic pneumothorax[a]	Postop PE or DVT[b]	Postop respiratory failure	
Postop hip fracture[a]	Transfusion reaction	Postop abdominopelvic wound dehiscence	
Technical difficulty with procedure	Birth trauma	Postop sepsis	
Obstetric trauma (all delivery types)			

[a] Panel ratings were based on definitions different than final definitions. For "Iatrogenic pneumothorax," the rated denominator was restricted to patients receiving thoracentesis or central lines; the final definition expands the denominator to all patients (with same exclusions). For "In-hospital fracture" panelists rated the broader Experimental indicator, which was replaced in the Accepted set by "Postoperative hip fracture" due to operational concerns.
[b] Vascular complications were rated as Unclear (-) by surgical panel; multispecialty panel rating is shown here.
[c] "Acceptable" and "acceptable-" both have median score 7-9; "acceptable" means no more than 1 or 2 panelists rated indicator below 7; "acceptable-" means no more than 1 or 2 panelists rated indicator in 1-3 range.

National Quality Forum evidence review

- National Voluntary Consensus Standards for Hospital Care: Additional Priorities, 2007

- Purpose was to seek additional voluntary consensus standards for measuring the performance of the nation's general acute care hospitals, including: 1) morbidity and mortality measures, 2) anesthesia and surgery measures, 3) measures for utilization rates for risky or often unnecessary procedures, 4) surgical volume and mortality measures, 5) readmission rates and length of stay (LOS) rates, 6) pain assessment, and 7) pediatric asthma.

- Convened Technical Assessment Panels for Patient Safety, Surgery and Anesthesia, Pediatric, Public Reporting, Length of Stay and Readmission

NQF review results

QI Name	Endorsed date
Death in Low-Mortality DRGs (PSI 2)	**5/15/08**
Death Among Surgical Patients w/ Treatable Complications (PSI 4)	**5/15/08**
Foreign Body Left During Procedure-Provider level (PSI 5)	**5/15/08**
Iatrogenic Pneumothorax-Provider level (PSI 6)	**5/15/08**
Postoperative DVT or PE (PSI 12)	**7/31/08**
Postoperative Wound Dehiscence-Provider level (PSI 14)	**5/15/08**
Accidental Puncture or Laceration-Provider level (PSI 15)	**5/15/08**
Transfusion Reaction-Provider level (PSI 16)	**5/15/08**

Yellow = proposed HACs; Cyan = proposed RHQDAPU

Predictive validity established:
Impact of preventing a PSI on mortality, LOS, charges
NIS 2000 analysis by Zhan & Miller, *JAMA* 2003;290:1868-74

Indicator	Δ Mort (%)	Δ LOS (d)	Δ Charge ($)
Postoperative septicemia	21.9	10.9	$57,700
Selected infections due to medical care	4.3	9.6	38,700
Postop abd/pelvic wound dehiscence	9.6	9.4	40,300
Postoperative respiratory failure	21.8	9.1	53,500
Postoperative physiologic or metabolic derangement	19.8	8.9	54,800
Postoperative thromboembolism	6.6	5.4	21,700
Postoperative hip fracture	4.5	5.2	13,400
Iatrogenic pneumothorax	7.0	4.4	17,300
Decubitus ulcer	7.2	4.0	10,800
Postoperative hemorrhage/hematoma	3.0	3.9	21,400
Accidental puncture or laceration	2.2	1.3	8,300

Excess mortality, LOS, and charges computed from mean values for PSI cases and matched controls.

Predictive validity established:
Impact of preventing a PSI on mortality, LOS, charges
VA PTF 2001 analysis by Rivard et al., *Med Care Res Rev*; 65(1):67-87

Indicator	Δ Mort (%)	Δ LOS (d)	Δ Cost ($)
Postoperative septicemia	30.2	18.8	$31,264
Selected infections due to medical care	2.7	9.5	13,816
Postop abd/pelvic wound dehiscence	11.7	11.7	18,905
Postoperative respiratory failure	24.2	8.6	39,745
Postoperative physiologic or metabolic derangement			
Postoperative thromboembolism	6.1	5.5	7,205
Postoperative hip fracture			
Iatrogenic pneumothorax	2.7	3.9	5,633
Decubitus ulcer	6.8	5.2	6,713
Postoperative hemorrhage/hematoma	5.1	3.9	7,863
Accidental puncture or laceration	3.2	1.4	3,359

Excess mortality, LOS, and charges computed from mean values for PSI cases and matched controls.

AHRQ PSI Validation Pilot Goals

- Gather evidence on the criterion validity of the PSIs based on medical record review

- Improve guidance about how to interpret & use the data

- Evaluate potential refinements to the specifications

- Develop medical record abstraction tools

- Develop mechanisms for conducting validation studies on a routine basis

Positive Predictive Value

- The **positive predictive value** or **post-test probability** is the proportion of flagged cases who actually had the event.

- The Positive Predictive Value (PPV) can be further defined as:

$$PPV = \frac{\text{number of True Positives}}{\text{number of True Positives} + \text{number of False Positives}}$$

PSI Validation Pilot Methods

- Retrospective cross-sectional study design
- Volunteer sample of collaborative partners
  - Facilitating organizations (e.g., Arizona)
  - Hospital systems
  - Individual hospitals
- Sampling based on administrative data
- Sampling probabilities assigned using AHRQ QI software to generate desired sample size locally (30) and nationally (240 per indicator)

Data collection methods

- Each hospital identified chart abstractors
- Training occurred via webinars
- Medical record abstraction tools & guidelines
    - Pretested in the Sacramento area
    - Targeted the ascertainment of the event, risk factors, evaluation & treatment, and related outcomes
- Coordinating center entered data from paper forms

PSI Validation Pilot Timeline

- 10 indicators- divided into 2 phases of 5 each
- Phase I review-
  - Training early 2007
  - Chart review 4 month process
  - 2nd Qtr 2006 through 1st Qtr 2007 (but some hospitals had to reach back as far as 4th Qtr 2005)
- Phase II review –
  - Awaiting OMB approval
  - Pre-pilot (6 hospitals) now underway
- Phase III –sensitivity determination

PSI Validation Pilot Phases

Phase I	Phase II
Accidental puncture and laceration	Foreign body left in during procedure
Selected infection due to medical care	Postoperative hemorrhage or hematoma
Postoperative pulmonary embolism or deep vein thrombosis	Postoperative physiologic and metabolic derangement
Postoperative sepsis	Postoperative respiratory failure
Iatrogenic pneumothorax	Postoperative wound dehiscence

PSI Validation Pilot Samples

Phase I	Hospitals	Sample
Accidental puncture and laceration	43	249
Iatrogenic pneumothorax	38	205
Postoperative PE/DVT	37	155
Selected Infection due to Medical Care	37	189
Postoperative Sepsis	33	164
Overall	47	967

Accidental Puncture or Laceration

- N=249
- 91% (95% CI = 86-94%) PPV or true events
- 9% (n=23) were false positives
  - 7% (n=18) miscoded
    - 4 had disease-related lesions (perforated appendix or ischemic colon, ruptured AA, rectovesical fistula)
    - 7 had a different complication (4 bleeding due to operative conduct, 1 surgical site infection, 1 dislodged gastrostomy tube, 1 periprosthetic fracture)
    - 7 cases had no apparent event other than normal operative/procedural conduct (intentional, rule-out)
  - 2% POA (related to an earlier episode of care)

Accidental Puncture or Laceration

- Characteristics of confirmed cases (N=226)
  - 170 (75%) were potentially consequential
  - Most were related to an abdominal or pelvic procedure
    - 51 (30%) enterotomy or other perforation of the GI tract
    - 42 (25%) bladder injury
    - 33 (19%) dural tear
    - 27 (16%) vascular injury
  - 132 (78%) involved a reparative procedure at the time of occurrence
  - 19 (11%) required a return to the OR (one death)

Iatrogenic pneumothorax

- N=205
- 78% (95% CI = 73-82%) PPV or true events that occurred during the hospitalization
- 11% were false positives
  - 7% (n=14) present or suspected at admission, most (n=8) transferred in
  - 4% no documentation of event (miscoded), but some with suspicion (n=3)
- 11% had exclusionary diagnosis or procedure (e.g., trauma, metastatic cancer)

Iatrogenic pneumothorax

- Characteristics of confirmed cases (N=156)
  - 9 (6%) transthoracic needle aspiration or biopsy
  - 66 (47%) central venous catheter placement
    - Only 4 used sonographic and 7 fluoroscopic guidance
  - 59 (40%) other invasive procedures on or near the neck or chest wall
    - 37 catheterization, pacemaker insertion
    - 3 laparoscopic procedures
    - 8 nephrectomy/renal procedures
    - 2 operations involving the spinal canal
    - 9 Other procedures
  - 5 (5%) mechanical ventilation
  - 1 (1%) cardiopulmonary resuscitation

Postoperative DVT or PE

- N = 155

- Coding perspective:
  - PPV = 83% (95% CI = 73-95%)
  - 17% were false positives
    - 10% (n=12) POA
    - 7% (n=8) no documentation of VTE (miscoded)

- Clinical perspective:
  - PPV = 48% (95% CI = 33-61%)
  - Additional false positives due to preoperative VTE (20%), upper extremity DVT (9%), superficial or unspecified vein (6%)

Comparing PPV estimates with UHC sample for postoperative DVT/PE

UHC Cohort (n=450)	Coding	Clinical
Sensitivity	80% (46- 00%)	100%
Specificity	99.5% (99.3-99.6%)	98.6% (98.6-99.2%)
Positive Predictive Value	72% (67-79%)	44% (36-52%)
Negative Predictive Value	99.6% (98.9-100%)	100%
AHRQ Cohort (n=121)		
Positive Predictive Value	83% (73-95%)	48% (33-61%)

Selected Infection due to Medical Care

- N=191
- 61% (95% CI = 51-71%) were true events that occurred during the index hospitalization
- 39% were false positives
  - 7% (n=14) had exclusionary diagnosis
  - 20% (n=39) were present on admission, with no new infection
  - 12% (n=23) had no documentation of infection

Selected Infection due to Medical Care

- N=115 with new infection
  - 106 with new infection
  - 9 with POA + new infection
- Majority related to central lines (n= 53)
  - 35 ** cases due to non-tunneled central lines (SC, IJ, Fem)
    - Mean 11 days (7 days SD), range 2-35 days (n=33)
  - 17 cases associated with PICC lines
    - Mean 13.4 days (7.7 days SD), range 2-35 days
- 26 cases related to other types of catheters (3 arterial lines, 4 long-term vascular access ports, 3 ET, 8 IV, 1 urinary, 2 other)

*   *Difficult review- some earlier termination by abstractors*
** *2 cases had extreme values of 101 and 374 days were excluded from calculation*

Postoperative Sepsis

- N=164
- 41% (95% CI = 28-54%) were true events that occurred during the hospitalization
- 59% were false positive
  - 17% had no documentation of bacteremia, septicemia, sepsis or SIRS
  - 17% had infection (=14%) or sepsis (=3%) POA
  - 25% did not have elective surgery

Summary of PPVs
Preliminary estimates

PSI	PPV%
Accidental puncture or laceration	91%
Iatrogenic pneumothorax	78%
Postoperative DVT/PE	48-83%
Selected infections due to medical care	61%
Postoperative sepsis	41%

PPVs high for NQF-endorsed indicators

Recognizing limitations

- Not all data elements of interest available via chart review

- Time constraints (minimize burden on collaborators)

- Inter-hospital variation in documentation and abstraction

- Volunteer sample; time periods varied slightly across hospitals

AHRQ QI Validation Pilot Future plans

- **Further analysis of potential preventability**
- **Evaluation of alternative ICD-9-CM specifications**
  - Can we improve PPV?
- **Establish ongoing infrastructure for validation**
- **Estimate sensitivity of 10 indicators (including Foreign Body, Pneumothorax, Wound Dehiscence, and Accidental Puncture and Laceration)**

Policy implications

- Coding changes needed to enhance specificity and PPV in some areas
  - AHRQ proposed new codes for DVT
  - CMS proposed new code for catheter-associated bloodstream infection
- With these changes, 3 of 5 PSIs tested in Phase 1 should have high PPV
- These indicators have been endorsed by NQF
- More data on sensitivity (false negatives) are needed to avoid rewarding hospitals that underreport

Acknowledgments

- **AHRQ project team**
  - Mamatha Pancholi &Marybeth Farquhar

- **Battelle training and support team**
  - Laura Puzniak & Lynne Jones

- **All of the validation pilot partners!**