

Convergent Validity of Three Methods for Measuring Postoperative Complications

Bradley A. Fritz, M.D., Krisztina E. Escallier, M.D., Arbi Ben Abdallah, D.E.S., Ph.D., Jordan Oberhaus, Jennifer Becker, B.S., Kristin Geczi, Sherry McKinnon, B.S., Dan L. Helsten, M.D., Anshuman Sharma, M.D., Troy S. Wildes, M.D., Michael S. Avidan, M.B.B.Ch.

ABSTRACT

Background: Anesthesiologists need tools to accurately track postoperative outcomes. The accuracy of patient report in identifying a wide variety of postoperative complications after diverse surgical procedures has not previously been investigated.

Methods: In this cohort study, 1,578 adult surgical patients completed a survey at least 30 days after their procedure asking if they had experienced any of 18 complications while in the hospital after surgery. Patient responses were compared to the results of an automated electronic chart review and (for a random subset of 750 patients) to a manual chart review. Results from automated chart review were also compared to those from manual chart review. Forty-two randomly selected patients were contacted by telephone to explore reasons for discrepancies between patient report and manual chart review.

Results: Comparisons between patient report, automated chart review, and manual chart review demonstrated poor-to-moderate positive agreement (range, 0 to 58%) and excellent negative agreement (range, 82 to 100%). Discordance between patient report and manual chart review was frequently explicable by patients reporting events that happened outside the time period of interest.

Conclusions: Patient report can provide information about subjective experiences or events that happen after hospital discharge, but often yields different results from chart review for specific in-hospital complications. Effective in-hospital communication with patients and thoughtful survey design may increase the quality of patient-reported complication data.

(ANESTHESIOLOGY 2016; 124:1265-76)

IN order to assess and improve perioperative care, it is essential to track postoperative patient outcomes. Together with other healthcare professionals, anesthesiologists have a strong incentive to monitor these outcomes. This is becoming ever more pressing with the imminent introduction of pay-for-performance models and with the advent of the perioperative surgical home concept, where the role of the anesthesiologist at some centers is expanding to involve all aspects of perioperative patient management, including postoperative care. To justify this expanded role, anesthesiologists must provide evidence that the perioperative surgical home model improves patient outcomes, including reduced complication rates both before and after hospital discharge.^{1,2} Patient-reported outcomes (PROs) will likely play a role in achieving this aim. Ever since Guyatt *et al.*³ published an influential 1993 article on the measurement of health-related quality of life, a variety of PROs have been developed to measure overall health or to measure disease-specific domains. The proliferation of studies reporting PROs led to an extension of the Consolidated Standards of Reporting Trials guidelines to specifically address the use of PROs in randomized trials⁴ and instructions from the Food and Drug Administration regarding the use of PROs in trials

What We Already Know about This Topic

- Patient-reported outcomes will play an important role in healthcare research and reimbursement, but their concordance with medical record data in the perioperative period has not been examined

What This Article Tells Us That Is New

- In a study of more than 1,500 subjects more than 30 days after surgery, patient-reported outcomes, compared to automated or manual chart review, demonstrated poor-to-moderate positive agreement (0 to 58%) and excellent negative agreement (82 to 100%)
- Discrepancies frequently reflected patients reporting events that occurred outside the time period of interest, suggesting that more effective in-hospital communication and thoughtful survey design may improve the value of patient-reported outcomes

for labeling of medical products.⁵ Wu *et al.*⁶ recently published a checklist to help clinicians interpret research studies that use PROs. In fact, PROs have become so widespread that multiple major journals have published viewpoint articles suggesting that PROs be collected regularly in all clinical settings.^{7,8}

This article is featured in "This Month in Anesthesiology," page 1A. Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are available in both the HTML and PDF versions of this article. Links to the digital files are provided in the HTML text of this article on the Journal's Web site (www.anesthesiology.org). This article has a video abstract. This is a substudy of the SATISFY-SOS surgical outcomes registry (NCT02032030).

Submitted for publication September 20, 2015. Accepted for publication February 26, 2016. From the Department of Anesthesiology, Washington University School of Medicine, St. Louis, Missouri.

Copyright © 2016, the American Society of Anesthesiologists, Inc. Wolters Kluwer Health, Inc. All Rights Reserved. Anesthesiology 2016; 124:1265-76

Although PROs are widely used to measure largely subjective outcomes, such as quality of life, functionality, and pain, they have not commonly been used to measure more objective medical outcomes such as myocardial infarction. Obtaining such outcomes by patient report rather than by chart review is desirable because chart review is unable to capture events that happen at home or events for which patients do not seek care at affiliated medical facilities. When compared to the medical record, nonsurgical patients self-report elements of their medical history with high negative agreement (range, 64 to greater than 99%) but with low positive agreement (range, 1 to 85%).⁹⁻¹⁶ This means that one modality almost always reports a symptom as absent when the other modality reports that symptom as absent, but the two modalities frequently disagree when one modality reports a symptom as present.¹⁷ Similar results were obtained in a study examining patient report of medical complications after bone marrow transplant.¹⁸ The accuracy of patient reports of medical complications has not been investigated as thoroughly among surgical patients. Existing studies have been limited to either a single type of surgery¹⁹⁻²¹ or to a single postoperative complication.²²⁻²⁴ Additional characterization of PROs in a diverse surgical population is necessary if they are to be used to identify postoperative complications for clinical care, research, or quality improvement. The purpose of this study was to examine the convergent validity of PROs and medical record review in detecting postoperative complications in the first 30 days after a wide variety of surgical procedures.

Materials and Methods

Patient Population

The Human Research Protection Office at Washington University approved this study. This analysis is a part of the Systematic Assessment and Targeted Improvement of Services Following Yearlong Surgical Outcomes Surveys (SATISFY-SOS) project (NCT02032030). All patients who completed surveys provided written, informed consent when visiting the Center for Preoperative Assessment and Planning. A waiver of consent was obtained to extract data from the medical record of surgical patients who did not visit the preoperative clinic or did not enroll in SATISFY-SOS during their preoperative clinic visit. Patients aged 18 yr or older were eligible if they were receiving anesthesia services for any procedure at Barnes-Jewish Hospital (St. Louis, Missouri) or an affiliated hospital between July 2012 and June 2013. This includes patients receiving general anesthesia, regional anesthesia, monitored anesthesia care, and procedural sedation.

Postprocedure Surveys

Patients received a questionnaire approximately 30 days after their procedure. In addition to previously validated instruments such as the Veterans RAND 12-item short form²⁵ and the Barthel Index,²⁶ the survey included a question asking what complications patients experienced while in the hospital recovering from their procedure. Patients could select

from a list of 18 complications, write in another complication, indicate they prefer not to answer, or indicate they experienced no complications (table 1). Patients could select more than one complication. The survey contained a separate question about complications after hospital discharge, which was not included as part of this investigation. The survey was conducted in English only.

The survey was sent by e-mail if the patient provided an e-mail address at the time of consent. If the patient did not provide an e-mail address or did not respond to the e-mail within 14 days, then a paper copy of the questionnaire was sent through the mail. If the patient did not return the mailed survey within 21 days, a second copy was mailed. If the patient did not return the second mailed survey within 14 days, then the patient was telephoned and asked the questions over the phone. If there was initially no answer at the telephone number, then the patient was telephoned at least once per week, up to five times. Surveys returned and electronically processed by June 2013 were included in the analysis.

Automated Chart Review

An automated computer algorithm was used to review the patient's electronic medical record from the time of the procedure until hospital discharge. Researchers were blinded to survey responses during development of the algorithm. The algorithm is described in full in the supplemental appendix (Supplemental Digital Content 1, <http://links.lww.com/ALN/B268>). Most of the complications were identified using

Table 1. Survey Question Regarding Postoperative Complications

While you were in the hospital recovering from your procedure did you have or did a doctor or nurse tell you that you had any of the following complications? (Fill in all that apply)
Heart attack
Your heart stopped beating (cardiac arrest)
Heart failure (congestive heart failure)
Abnormal heart rhythm such as atrial fibrillation
Severe pain coming from your heart (angina)
Stroke (e.g., weakness on one side of the body or difficulty speaking)
Blood clot in your leg
Blood clot in your lung
Infection in the surgical wound
You stopped breathing (respiratory arrest)
You were placed on a breathing machine because you were struggling to breathe on your own (respiratory failure)
An infection in your lungs (pneumonia)
Kidney failure and you needed dialysis
Nerve injury related to your procedure
Internal bleeding from your stomach or intestine (gastrointestinal bleed)
Stomach or intestinal ulcer
Severe pain lasting for > 1 day
Severe nausea and vomiting lasting for > 1 day
Prefer not to answer
None
Other: _____

International Classification of Diseases, 9th edition, codes present at the time of discharge, after the exclusion of admitting diagnoses. Arrhythmias were identified using new rhythm abnormalities on electrocardiogram reports, a new prescription for amiodarone, or the performance of a cardioversion. Angina was defined as pain with documented location of “substernal” and severity of at least 7/10. Respiratory failure was defined using procedure codes for mechanical ventilation. Severe pain was defined as a pain score of at least 7/10 on postoperative day 1 or later. Severe nausea and vomiting was defined as the administration of an antiemetic (ondansetron, metoclopramide, scopolamine, prochlorperazine, or dexamethasone) on postoperative day 1 or later. The automated chart review was performed for all patients who returned the survey.

To allow for assessment of selection bias, the automated chart review was also performed for a matched cohort of patients who did not enroll in this study. To allow for assessment of follow-up bias, the automated chart review was also performed for a matched cohort of patients who consented to the study but did not return the survey. Patients were matched using the propensity score matching technique based on age, sex, race, body mass index, American Society of Anesthesiologists physical status, and comorbidities (coronary artery disease, previous myocardial infarction, congestive heart failure, heart valve disease, atrial fibrillation, history of pacemaker or defibrillator, previous cerebrovascular accident, previous aortic disease, history of venous thromboembolism, diabetes, pulmonary hypertension, end-stage renal disease requiring dialysis, chronic obstructive pulmonary disease, obstructive sleep apnea, liver cirrhosis, previous or current cancer, gastroesophageal reflux disease, previous anemia, previous thrombocytopenia, previous positive Coombs test, and dementia). Propensity scores were calculated using logistic regression in which the dependent variable was enrollment in the study (for the first matched cohort) or return of the survey (for the second matched cohort). Matching was performed using the nearest neighbor method with replacement using a caliper of 0.10.

Manual Chart Review

A subset of the patients who returned the survey was selected for manual chart review. All patients who reported a complication on the survey, as well as a random sample of the remaining survey respondents, were included in this subset. Each chart was systematically reviewed by two independent, trained researchers. The reviewer used the contents of the medical record to determine whether each complication had been diagnosed while the patient was in the hospital (*i.e.*, if a progress note or discharge summary referred to the complication or if diagnostic tests supported the diagnosis). Each researcher discussed potential ambiguities with a third independent researcher. Discrepancies between these two manual data extraction processes were resolved by manual chart review undertaken by a fourth independent researcher. The records reviewed included anesthesiologist and recovery

room notes, admission notes, progress notes, discharge summaries, laboratory values, and reports from imaging studies and diagnostic tests. Chart reviewers were blinded to survey results. The two independent manual chart reviews were reliable, achieving near-perfect agreement for eight complications (Cohen's κ values ranged from 0.5 to 1.0).

Statistical Methods

The characteristics of the patient population were described using frequency and percentage for categorical variables and using median and interquartile range for nonnormally distributed continuous variables. The incidence of complications found by each modality was described using frequency and percentage. Agreement between the survey results and the automated chart review was quantified using positive agreement, negative agreement, and Cohen's κ statistic. By definition, the κ statistic was undefined if either modality identified zero cases of a particular complication. The same methods were used to quantify agreement between the survey results and the manual chart review and between the two chart review methods. Cases were excluded from the agreement statistics if the patient marked “prefer not to answer” on the survey or left the question blank. As a sensitivity analysis, the agreement statistics were repeated stratifying patients by survey modality, sex, and age (dichotomized as less than 60 yr and 60 yr or more). To assess for selection bias at the time of enrollment, McNemar test was used to compare the incidence of complications found by automated chart review between survey respondents and a matched cohort of patients who did not enroll in the study. To assess for follow-up bias, McNemar test was used to compare the incidence of complications found by automated chart review between survey respondents and matched nonrespondents. To explore the effects of combining complication detection modalities, the survey results and automated chart review were combined into a composite complication assay. The composite assay was positive if either the survey or the automated chart review indicated a complication. The survey and automated chart review were chosen for the composite assay because these modes of detecting complications are more feasible than manual chart review for scaling to large patient populations. Agreement between the composite assay and the manual chart review was quantified using positive agreement, negative agreement, and Cohen's κ statistic. All analyses were performed using SAS version 9.3 (SAS Institute, Inc., USA). *P* values less than 0.05 were considered statistically significant.

Telephone Calls

To qualitatively explore the reasons for discordance between survey responses and chart review, one researcher phoned 70 randomly selected patients for whom the survey and the manual chart review yielded discordant results. A result was discordant if any complication was positive by one method and negative by the other method. Patient reports

of angina, nerve injury, severe pain, and severe nausea/vomiting were not included, as patients might experience these complications without notifying the medical staff. The researcher asked patients to describe their experiences and also conducted an unblinded review of the patient's entire chart, including past admissions. The reason for each discordant report was clarified where possible. Naturally occurring themes among the reasons for discordance were identified.

Results

Patient Population

Of the approximately 30,000 patients who visited the Center for Preoperative Assessment and Planning between July 2012 and June 2013, 8,792 (30%) were enrolled during their preoperative clinic visit and 5,133 patients (58% of those enrolled) returned the postoperative survey (fig. 1). This analysis includes the 1,578 surveys that had been returned to our department and electronically processed by June 2013. As shown in table 2, these patients underwent a variety of surgical procedures. Patients were enrolled an average of 10 days before surgery (fig. 2).

Incidence of Postoperative Complications

Of those who returned the survey, 446 patients reported at least one complication, 1,010 patients reported no complication, 5 patients indicated that they prefer not to answer, and 117 patients left the question blank. The most commonly reported complication was severe pain lasting more than 1 day (188 patients), followed by abnormal heart rhythm (86 patients) and severe nausea/vomiting lasting more than 1 day (63 patients). The frequency count of each complication is shown in table 3. Pain, abnormal heart rhythm, and nausea/vomiting were also the three most common complications found by each of the chart review modalities.

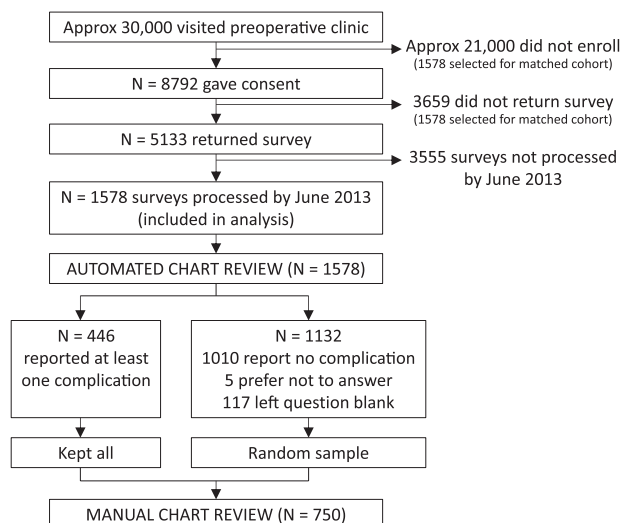


Fig. 1. Patients included at each phase of the analysis.

Table 2. Characteristics of the Cohort (N = 1,578)

Characteristics	Median or Frequency	Interquartile Range or Percentage
Age	59	51–68
Body mass index	29	25–34
Female sex	883	56
Race		
White	1,403	89
Black American	159	10
Other/did not provide	16	1
History of ever smoking	750	48
American Society of Anesthesiologists physical status		
1	93	6
2	787	50
3	641	40
4	57	4
Surgical specialty		
Cardiac	38	2
Electrophysiology	38	2
General	252	16
Gynecology	107	7
Neurosurgery	82	5
Ophthalmology	62	4
Orthopedic	288	18
Other	39	2
Otolaryngology	82	5
Plastic	46	3
Thoracic	47	3
Urology	107	7
Vascular	34	2
Specialty data missing	356	22*
Hypertension	620	39
Coronary artery disease	156	10
Previous myocardial infarction	74	5
Congestive heart failure	66	4
Atrial fibrillation	79	5
Cerebrovascular disease	67	4
Previous venous thromboembolism	69	5
Diabetes mellitus	203	13
Chronic kidney disease	84	5
Chronic obstructive pulmonary disease or asthma	177	11
Obstructive sleep apnea	166	11
Gastroesophageal reflux disease	409	26
Current or previous cancer	323	20

*Percentages in this cell do not add to 100% due to rounding.

The automated chart review was performed for patients who returned the survey, for a matched cohort of patients who consented to the study but did not return the survey by June 2013, and (via waiver of consent) for a matched cohort of patients who did not consent to the study. As measured by the automated chart review, patients who returned the survey were less likely to experience a complication (355 of 1,578, 22%) compared to those who did not return the survey (421 of 1,578, 27%): McNemar chi-square test (degree of freedom = 1) = 7.75, P = 0.006. Patients who returned the

Downloaded from http://aas2.silverchair.com/anesthesiology/article-pdf/124/6/1265/487222/20160600_0-00020.pdf by guest on 01 February 2023

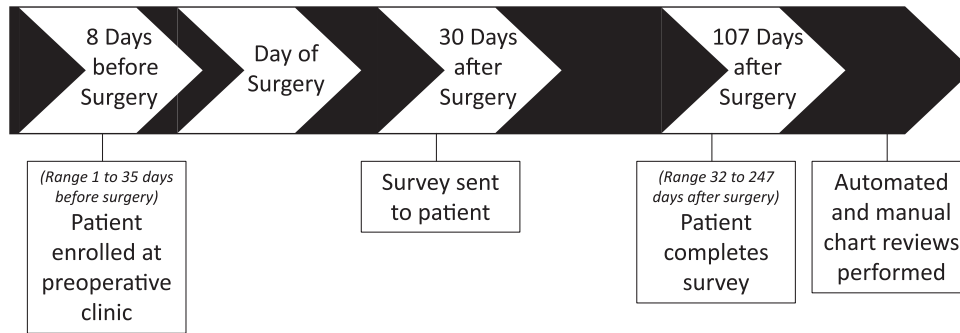


Fig. 2. Timeline of events for a typical patient enrolled in the study. Numbers in the timeline reflect the median number of days before or after surgery each event occurred.

survey were equally likely to experience a complication (355 of 1,578, 22%) compared to those who did not consent to the study (332 of 1,578, 21%); McNemar chi-square test (degree of freedom = 1) = 0.93, $P = 0.36$.

Agreement between Surveys and Chart Reviews

The survey and the automated chart review (N = 1,578) showed poor-to-moderate positive agreement (range, 0 to 41%) and excellent negative agreement (range, 93 to greater than 99%; table 3). Kappa values ranged from 0 for cardiac arrest and angina to 0.43 for pulmonary embolism. Likewise, the survey and the manual chart review (N = 750) showed poor positive agreement (range, 0 to 55%) and excellent negative agreement (range, 82 to 100%; table 4). Kappa values ranged from 0.00 for stomach or intestinal ulcer to 0.53 for stroke. These results did not change substantially when patients were stratified based on survey modality, sex, or age.

Agreement statistics between the automated chart review and the manual chart review are shown in table 5. The two chart review methods also showed poor-to-moderate positive agreement (range, 0 to 58%) and excellent negative agreement (range, 87 to 100%). Kappa values ranged from 0.00 for angina, gastrointestinal bleed, and severe pain to 0.55 for arrhythmia.

Agreement statistics between the composite assay (which was positive if the survey or the automated chart review was positive) and the manual chart review are shown in table 6. The two methods showed poor-to-moderate positive agreement (range, 0 to 59%) and excellent negative agreement (range, 82 to 100%). Kappa values ranged from 0.00 for stomach or intestinal ulcer to 0.48 for severe nausea and vomiting.

Reasons for Discordant Reports of Complications

Of the 70 randomly selected patients whom we attempted to contact by telephone, we successfully reached 42 patients.

Table 3. Agreement between Survey and Automated Chart Review (N = 1,578)

Complication	Frequency				Agreement		
	Survey Negative; Chart Review Negative	Survey Negative; Chart Review Positive	Survey Positive; Chart Review Negative	Survey Positive; Chart Review Positive	Positive Agreement (%)	Negative Agreement (%)	Kappa (95% CI)
Myocardial infarction	1,566	5	6	1	15	> 99	0.15 (-0.12 to 0.42)
Cardiac arrest	1,571	2	5	0	0	> 99	0.00 (0.00 to 0.00)
Congestive heart failure	1,545	10	16	7	35	99	0.34 (0.15 to 0.53)
Arrhythmia	1,461	31	58	28	39	97	0.36 (0.25 to 0.46)
Angina	1,568	2	8	0	0	> 99	0.00 (0.00 to 0.00)
Stroke	1,555	12	5	6	41	99	0.41 (0.18 to 0.64)
Deep venous thrombosis	1,558	0	20	0	0	99	—
Pulmonary embolism	1,560	10	3	5	43	> 99	0.43 (0.18 to 0.69)
Wound infection	1,526	16	32	4	14	98	0.13 (0.00 to 0.25)
Respiratory arrest	1,571	0	7	0	0	> 99	—
Respiratory failure	1,556	0	19	3	24	99	0.24 (0.02 to 0.46)
Pneumonia	1,531	26	11	10	35	99	0.34 (0.18 to 0.50)
Kidney failure requiring dialysis	1,561	1	13	3	30	> 99	0.30 (0.04 to 0.56)
Nerve injury	1,478	62	36	2	4	97	0.01 (-0.04 to 0.06)
Gastrointestinal bleed	1,546	22	8	2	12	99	0.11 (-0.04 to 0.26)
Peptic ulcer	1,571	0	7	0	0	> 99	—
Severe pain > 1 day	1,387	3	187	1	1	94	0.01 (-0.01 to 0.02)
Severe nausea/vomiting > 1 day	1,365	150	35	28	23	94	0.18 (0.11 to 0.26)

Kappa is undefined if either the survey or the chart review identifies zero cases of a particular complication.

Table 4. Agreement between Survey and Manual Chart Review (N = 750)

Complication	Frequency				Agreement		
	Survey Negative; Chart Review Negative	Survey Negative; Chart Review Positive	Survey Positive; Chart Review Negative	Survey Positive; Chart Review Positive	Positive Agreement (%)	Negative Agreement (%)	Kappa (95% CI)
Myocardial infarction	743	0	7	0	0	> 99	—
Cardiac arrest	745	0	5	0	0	> 99	—
Congestive heart failure	726	1	22	1	8	98	0.08 (−0.07 to 0.22)
Arrhythmia	651	13	56	30	47	95	0.42 (0.31 to 0.53)
Angina	740	2	7	1	18	99	0.18 (−0.13 to 0.48)
Stroke	739	0	7	4	53	> 99	0.53 (0.22 to 0.84)
Deep venous thrombosis	728	2	15	5	37	99	0.36 (0.13 to 0.59)
Pulmonary embolism	742	0	5	3	55	> 99	0.54 (0.19 to 0.90)
Wound infection	706	8	35	1	4	97	0.02 (−0.06 to 0.11)
Respiratory arrest	743	0	7	0	0	100	—
Respiratory failure	728	0	20	2	17	99	0.16 (−0.04 to 0.36)
Pneumonia	729	0	17	4	32	99	0.31 (0.08 to 0.55)
Kidney failure requiring dialysis	734	0	14	2	22	99	0.22 (−0.03 to 0.47)
Nerve injury	706	6	32	6	24	97	0.22 (0.06 to 0.38)
Gastrointestinal bleed	739	1	9	1	17	99	0.16 (−0.12 to 0.45)
Peptic ulcer	742	1	7	0	0	99	0.00 (−0.01 to 0.00)
Severe pain > 1 day	464	98	110	78	43	82	0.25 (0.17 to 0.32)
Severe nausea/vomiting > 1 day	550	137	28	35	30	87	0.20 (0.12 to 0.28)

Note that these 750 patients are a subset of the 1,578 patients shown in table 3. Kappa is undefined if either the survey or the chart review identifies zero cases of a particular complication.

Table 5. Agreement between Automated and Manual Chart Review (N = 750)

Complication	Frequency				Agreement		
	Automated Review Negative; Manual Review Negative	Automated Review Negative; Manual Review Positive	Automated Review Positive; Manual Review Negative	Automated Review Positive; Manual Review Positive	Positive Agreement (%)	Negative Agreement (%)	Kappa (95% CI)
Myocardial infarction	746	0	4	0	0	> 99	—
Cardiac arrest	749	0	1	0	0	> 99	—
Congestive heart failure	736	1	12	1	13	99	0.13 (−0.10 to 0.36)
Arrhythmia	686	17	21	26	58	97	0.55 (0.42 to 0.68)
Angina	746	3	1	0	0	> 99	0.00 (0.00 to 0.00)
Stroke	738	0	8	4	50	> 99	0.50 (0.19 to 0.80)
Deep venous thrombosis	743	7	0	0	0	> 99	—
Pulmonary embolism	738	0	9	3	40	> 99	0.40 (0.08 to 0.71)
Wound infection	729	8	12	1	9	98	0.08 (−0.09 to 0.24)
Respiratory arrest	750	0	0	0	0	100	—
Respiratory failure	746	1	2	1	40	> 99	0.40 (−0.15 to 0.94)
Pneumonia	724	2	22	2	14	98	0.13 (−0.04 to 0.31)
Kidney failure requiring dialysis	745	1	3	1	33	> 99	0.33 (−0.16 to 0.82)
Nerve injury	713	11	25	1	5	97	0.03 (−0.07 to 0.13)
Gastrointestinal bleed	733	2	15	0	0	99	0.00 (−0.01 to 0.00)
Peptic ulcer	749	1	0	0	0	> 99	—
Severe pain > 1 day	571	175	3	1	1	87	0.00 (−0.02 to 0.02)
Severe nausea/vomiting > 1 day	550	90	28	82	58	90	0.49 (0.41 to 0.57)

Note that these 750 patients are a subset of the 1,578 patients shown in table 3. Kappa is undefined if either the automated or the manual chart review identifies zero cases of a particular complication.

Table 6. Agreement between Composite Assay (Combination of Survey and Automated Chart Review) and Manual Chart Review (N = 750)

Complication	Frequency				Agreement		
	Combined Review Negative; Manual Review Negative	Combined Review Negative; Manual Review Positive	Combined Review Positive; Manual Review Negative	Combined Review Positive; Manual Review Positive	Positive Agreement (%)	Negative Agreement (%)	Kappa (95% CI)
Myocardial infarction	740	0	10	0	0	> 99	—
Cardiac arrest	744	0	6	0	0	> 99	—
Congestive heart failure	720	1	28	1	7	98	0.06 (−0.06 to 0.18)
Arrhythmia	640	5	67	38	51	95	0.47 (0.37 to 0.57)
Angina	739	2	8	1	17	> 99	0.16 (−0.12 to 0.45)
Stroke	733	0	13	4	38	> 99	0.38 (0.11 to 0.64)
Deep venous thrombosis	728	2	15	5	37	99	0.36 (0.13 to 0.59)
Pulmonary embolism	735	0	12	3	33	> 99	0.33 (0.05 to 0.61)
Wound infection	698	7	43	2	7	97	0.06 (−0.04 to 0.15)
Respiratory arrest	743	0	7	0	0	100	—
Respiratory failure	728	0	20	2	17	99	0.16 (−0.03 to 0.36)
Pneumonia	715	0	31	4	21	98	0.20 (0.03 to 0.36)
Kidney failure requiring dialysis	733	0	15	2	21	99	0.21 (−0.04 to 0.45)
Nerve injury	682	6	56	6	16	96	0.14 (0.03 to 0.25)
Gastrointestinal bleed	726	1	22	1	8	98	0.08 (−0.07 to 0.22)
Peptic ulcer	742	1	7	0	0	> 99	0.00 (−0.01 to 0.00)
Severe pain > 1 day	462	97	112	79	43	82	0.25 (−0.17 to 0.32)
Severe nausea/vomiting > 1 day	526	79	52	93	59	89	0.48 (0.40 to 0.55)

Note that these 750 patients are a subset of the 1,578 patients shown in table 3. Kappa is undefined if either the composite assay or the manual chart review identifies zero cases of a particular complication.

Because some patients had more than one discrepancy between self-report and chart review, we were able to investigate 54 pairs of discordant complication reports. This included 49 complications reported on the survey but not found during manual chart review and 5 complications found during chart review but not reported on the survey. Naturally occurring themes among the reasons for discordance are outlined in table 7. The most common themes were patients accurately reporting events that occurred before surgery (n = 22), patients reporting a complication when they did not intend to do so (n = 11), and patients accurately reporting events that occurred after hospital discharge (n = 7). Some patients misinterpreted events, such as reporting ventilator use during general anesthesia as a complication. Among those patients with complications found during chart review that were not reported on the survey, some patients had no recall of the event discovered on chart review even when specifically asked, while other patients did remember the event after being prompted.

Discussion

In this cohort study, patient report of postprocedure complications showed low-to-moderate positive agreement and excellent negative agreement both when compared to an automated chart review and when compared to a manual chart review. Agreement was not improved when patient report and automated chart review were combined into a

composite assay. Much of the discordance occurred when a patient reported an event that happened before the procedure or after hospital discharge as a postoperative in-hospital complication. There was also evidence of a follow-up bias, as patients who did not return the survey were more likely to have complications detected by automated chart review than patients who returned the survey.

As in previous studies, our findings suggest that outcomes obtained from patients frequently do not match outcomes obtained from the medical record. Low positive agreement (0 to 71%) and high negative agreement (71 to greater than 99%) for patient-reported postoperative complications have been observed after hernia repair^{19,20} and after gynecologic oncology surgery.²¹ Some studies of orthopedic populations provide limited information because the medical record was only consulted if patient report of a postoperative complication was positive.^{27–29} The only complication that has been investigated in diverse surgical populations is wound infection, for which similar rates of positive agreement (47 to 83%) and negative agreement (95 to 98%) have been found.^{22–24} The high negative agreement is likely driven by the low incidence of the observed complications. This study extends the existing literature by examining a wide variety of complications in a broad population of surgical patients.

Several factors may lead to discordance between patient report and chart review. Patient report may not accurately reflect postprocedure events if healthcare professionals do not communicate effectively with patients in the hospital, if

Downloaded from http://aas2.silverchair.com/anesthesiology/article-pdf/124/6/1265/487222/20160600_0-00020.pdf by guest on 01 February 2023

Table 7. Reasons for Discordance between Survey and Manual Chart Review (N = 54 Complications)

	General Theme	Specific Reason for Discordance	Frequency	Example	
Survey positive; chart review negative	The patient accurately reports an event that occurred before surgery.	The patient accurately reports a chronic medical diagnosis.	9	55-yr-old male with a history of atrial fibrillation underwent radiofrequency ablation. He remained in sinus rhythm after the procedure. He reported "abnormal heart rhythm" on the survey.	
		The patient accurately reports a complication of a previous surgery.	9	34-yr-old male underwent laparoscopic cholecystectomy. He had a history of deep venous thrombosis 3 yr previously after resection of a bone tumor. He reported "blood clot in your legs" on the survey.	
		The patient accurately reports a past medical event that was unrelated to surgery.	4	76-yr-old male with a history of stroke many years before underwent radical cystoprostatectomy and sigmoid colectomy for bladder cancer. He reported "stroke" on the survey.	
	The patient accurately reports a complication that occurred in the hospital after surgery.	The patient accurately reports a complication that occurred in the hospital after surgery. (The chart review misses a complication that did occur.)	2	72-yr-old female with a history of atrial fibrillation underwent radiofrequency ablation. She was treated postoperatively with vancomycin and cefepime for probable pneumonia. She reported "infection in your lungs" on the survey, but this was not detected during chart review.	
		The patient accurately reports a complication that occurred after hospital discharge.	4	71-yr-old male underwent resection of a left temporal lobe mass. After discharge, he presented to his primary care physician with a cough. Workup revealed a pulmonary embolism, and he was prescribed warfarin. He reported "blood clot in your lungs" on the survey.	
	The patient inaccurately reports an event as a complication.	The patient accurately reports a complication that occurred during a subsequent hospitalization.	The patient accurately reports a complication that occurred during a subsequent hospitalization.	3	63-yr-old male underwent Whipple procedure for gallbladder carcinoma. He was discharged with a surgical drain in place. He was readmitted a month later with nausea/vomiting. His surgical drain output became purulent and culture was positive. He reported "infection of the wound" on the survey.
			The patient misinterprets an event that happened before surgery.	2	54-yr-old male underwent colonoscopy. Five months previously, he was admitted to the hospital for chest pain. Cardiac biomarkers were negative, and stress test was negative. He reported "heart attack" on the survey. On the telephone, he confirmed this was the event he referenced on the survey.
		The patient misinterprets an event that happened during hospitalization.	3	30-yr-old female underwent laser retinopexy for retinal detachment. She was extubated at the end of the surgery and discharged home. She reported "respiratory failure" on the survey because she was placed on a ventilator while under general anesthesia.	
	The patient does not intend to report a complication.	The patient misinterprets an event that happened after hospital discharge.	The patient misinterprets an event that happened after hospital discharge.	2	36-yr-old female underwent bilateral mastectomy for breast carcinoma. She was subsequently admitted for chest pain and fever. She underwent a pericardial window for a pericardial effusion. She reported this event as "abnormal heart rhythm" on the survey.
			The patient does not recall indicating a complication on the survey.	10	57-yr-old male underwent oropharyngeal cancer excision. He indicated "respiratory failure," "infection in your lungs," "kidney failure and you needed dialysis," "nerve injury," "internal bleeding," and "stomach or intestinal ulcer" on the survey. On the telephone, he had no recall of these survey responses and indicated that he experienced no complications.
There was a technical problem with the survey.		1	49-yr-old male with a history of end-stage renal disease underwent arteriovenous fistula creation for hemodialysis access. He marked "kidney failure and you needed dialysis" on the survey, but then he crossed out his answer. The computer still read this as a positive survey response.		

(Continued)

Table 7. (Continued)

	General Theme	Specific Reason for Discordance	Frequency	Example
Survey negative; chart review positive	The patient does not recall the event discovered on chart review.	The patient has no recall of the event discovered on chart review.	2	58-yr-old male underwent left and right ventricular assist device placement for heart failure. Postoperatively the patient experienced a gastrointestinal bleed. The medical record indicated the patient underwent upper endoscopy, during which a gastric ulcer was clipped. The patient reported "internal bleeding" on the survey but not "stomach or intestinal ulcer." On the telephone, the patient denied having an ulcer even when specifically asked.
		The patient recalls the event discovered on chart review only when specifically prompted.	1	65-yr-old male underwent esophagectomy for esophageal cancer. Postoperatively, he experienced a leak at the anastomosis and infection requiring antibiotics. The patient did not report "infection of the wound" on the survey. On the telephone, the patient did recall this event when specifically asked.
		The patient reports the event as one complication, and the chart review identifies the same event as a different complication.	1	70-yr-old male underwent right upper lobe sleeve resection for broncholithiasis. Postoperatively, he experienced a bronchial stump leak and empyema. He reported "infection in your lung (pneumonia)" on the survey, while the chart review reported "infection of the wound."
		The patient reports the complication discovered on chart review occurred after discharge.	1	78-yr-old male underwent Maze procedure for atrial fibrillation. Chart review revealed shortness of breath and pulmonary edema requiring diuresis. The patient did not report "congestive heart failure" on the survey. On the telephone, he reported going to another hospital after discharge for treatment of shortness of breath.

patients cannot accurately recall postprocedure events when completing the survey, or if patients do not understand the survey question. Chart review may not accurately reflect postprocedure events if complications are inadequately documented in the medical record or if there are flaws in the chart review algorithm. However, a well-refined chart review algorithm can demonstrate excellent sensitivity and specificity.³⁰

The qualitative analysis (table 7) revealed that the most frequent reason for discordance between patient report and chart review was because the patient reported an event that happened either before the procedure occurred or after the patient was discharged from the hospital. This suggests that patient report may be unreliable in precisely describing the temporal relationship between events. Patients may have found the question stem confusing or read the question too quickly, not noticing the instruction to limit responses to complications that occurred in the hospital while they were recovering from their procedures. We potentially could have placed even greater emphasis on the time period by using special formatting such as italicized or underlined text. However, part of the problem could stem from patients' inability to accurately recall the timing of events, rather than lack of understanding the intent of the survey question. Intensive care patients who experience a postoperative complication

generally have no memories from the time of surgery until intensive care unit discharge.³¹ Recent work at our institution confirms that patients have limited memories of the perioperative period.³² Patients who do not have memories of the perioperative period may be unable to accurately describe the temporal relationship among perioperative events.

As indicated in table 7, some discordance between patient report and chart review was caused by patient misinterpretation of hospital events. Nearly half of patients in the United States demonstrate limited literacy, as measured by the National Adult Literacy Survey.³³ These individuals are likely to have limited health literacy as well, and healthcare providers face heightened challenges when trying to communicate with patients of limited health literacy.³³ Efforts to convey information to patients are not always successful. In fact, only 42% of medical patients at a public hospital were able to state their diagnosis at the time of hospital discharge.³⁴ For this reason, physician-patient communication has been identified as a key domain for improving the discharge process.³⁵ Methods to enhance information transfer to patients with limited health literacy include frequent visits with less information per encounter and use of the "teach-back" approach to actively assess patient understanding.³⁶ Improved communication may prevent patients from reporting routine in-hospital events (e.g., prophylactic

antibiotics) as complications and help patients report complications when they do occur.

Some patients unintentionally reported a complication (table 7), which may be caused by survey fatigue. The question analyzed in this study was one item on a 44-item survey. Furthermore, the question included a long list of answer choices (table 1), which can be especially overwhelming when presented over the telephone. When respondents lack motivation to focus on the survey, they begin to answer without carefully optimizing their decision-making—a process known as satisficing.³⁷ Common manifestations of satisficing include increased use of the “prefer not to answer” option, selection of the first reasonable option without consideration of later options, consistent agreement with the interviewer, consistent selection of the midpoint of a Likert scale, and random selection from among multiple choice answers.³⁷ Minimizing fatigue should help improve the quality of survey responses.

Although patient report can be limited by temporal ambiguity, poor communication, or survey fatigue, chart review has limitations of its own. Subjective outcomes may not be documented adequately in the medical record, despite their importance to patients. For example, severe postoperative pain lasting greater than 1 day was only detected in four patients on automated chart review, but 188 patients reported this complication on the survey (table 3). Chart review is also unable to detect complications that occur after a patient leaves the hospital, unless the patient returns to the hospital or seeks care at a facility that shares a medical record with the hospital. Minor complications that require no treatment will be unrecognized. Patient report might be able to effectively fill these gaps. The shortcomings of automated chart review are of relevance to approaches that are currently adopted in anesthesiology research, including the Multicenter Perioperative Outcomes Group and the Anesthesia Quality Institute.

The reader should note that the authors do not intend to draw generalized conclusions about PROs. Although this article focuses on specific postoperative complications, many PROs ask patients about outcomes such as quality of life, return to work, or functional status. In fact, the survey circulated to patients after surgery as part of the SATISFY-SOS project includes the Veterans RAND 12-item short form²⁵ and the Barthel Index,²⁶ along with other common PROs. These PROs were not presented in this article because they have been previously validated and are well-accepted by the medical community.

The limitations of this study should be noted. For some complications, a small number of cases were observed. Small numbers can limit the precision of the agreement statistics, and κ could not be calculated for some complications because one of the chart review modalities detected zero cases. Even after matching on important characteristics, patients who did not return the survey had a higher incidence of postoperative complications detected by chart

review than patients who did return the survey. This could perhaps indicate a response bias. If patient report alone were used to identify complications, then the true incidence of complications would likely be underestimated. On the other hand, the similar incidence of complications between patients who enrolled in the study and patients who did not enroll in the study suggests the absence of a selection bias at the time of enrollment. Thus, the patients who enrolled in our study are likely representative of the patients who attend our preoperative clinic. However, because this study was conducted at a single academic medical center, the generalizability of the findings to other settings may be limited. The question utilized on the survey has not been previously validated, so it may have been misunderstood by patients. Changes to the questionnaire might have improved the convergent validity observed in this study. The manual chart review in this study was limited because several different researchers participated; however, the high level of agreement between the two independent reviews of each record suggests that the manual chart review algorithm was robust. Yet the poor-to-moderate agreement between the manual chart review and the automated chart review demonstrates that the results depend to a great extent upon the methodology used to perform the chart review.

As anesthesiologists focus increased attention on postoperative outcomes beyond the immediate postoperative period, it is essential to have practical tools to measure these outcomes. Patient report can provide information about subjective experiences and about events that happen at home, which cannot be detected by other data collection modalities. To collect data by patient report, it is imperative to communicate effectively with patients before hospital discharge, to clearly explain the time period of interest on the survey, and to minimize survey fatigue. However, patient report of objective complications frequently yields different results from those found by chart review. Thus patient report may be more appropriate for subjective experiences, while chart review may be preferable for objective complications.

Acknowledgments

The authors thank the staff of the Institute of Quality Improvement, Research, and Informatics at Department of Anesthesiology, Washington University School of Medicine, St. Louis, Missouri. The authors specially thank Will Godfrey, M.A. (Department of Anesthesiology, Washington University School of Medicine), for informatics support, and Beth Burnside, M.A. (now affiliated with Wright State Research Institute, Wright State University, Dayton, Ohio), for general project management.

Research reported in this publication was supported by the Washington University Institute of Clinical and Translational Sciences (St. Louis, Missouri; grant no. UL1TR000448) from the National Center for Advancing Translational Sciences of the National Institutes of Health (NIH, Bethesda, Maryland). Dr. Fritz was supported by the Washington

University Institute of Clinical and Translational Sciences (St. Louis, Missouri; grant nos. UL1TR000448 and TL1TR000449) from the National Center for Advancing Translational Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. SATISFY-SOS was funded by the Washington University Department of Anesthesiology and the Barnes-Jewish Hospital Foundation (St. Louis, Missouri; award reference no. 7937-77 to Dr. Avidan).

Competing Interests

The authors declare no competing interests.

Correspondence

Address correspondence to Dr. Avidan: Department of Anesthesiology, Washington University School of Medicine, Campus Box 8054, 660 S. Euclid Avenue, St. Louis, Missouri 63110. avidanm@anest.wustl.edu. This article may be accessed for personal use at no charge through the Journal Web site, www.anesthesiology.org.

References

- Vetter TR, Goeddel LA, Boudreaux AM, Hunt TR, Jones KA, Pittet JF: The Perioperative Surgical Home: How can it make the case so everyone wins? *BMC Anesthesiol* 2013; 13:6
- Kain ZN, Vakharia S, Garson L, Engwall S, Schwarzkopf R, Gupta R, Cannesson M: The perioperative surgical home as a future perioperative practice model. *Anesth Analg* 2014; 118:1126-30
- Guyatt GH, Feeny DH, Patrick DL: Measuring health-related quality of life. *Ann Intern Med* 1993; 118:622-9
- Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD; CONSORT PRO Group: Reporting of patient-reported outcomes in randomized trials: The CONSORT PRO extension. *JAMA* 2013; 309:814-22
- Food and Drug Administration: Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labeling claims. *Federal Register* 2009; 74:65132-3
- Wu AW, Bradford AN, Velanovich V, Sprangers MA, Brundage M, Snyder C: Clinician's checklist for reading and using an article about patient-reported outcomes. *Mayo Clin Proc* 2014; 89:653-61
- Basch E, Torda P, Adams K: Standards for patient-reported outcome-based performance measures. *JAMA* 2013; 310:139-40
- Manary MP, Boulding W, Staelin R, Glickman SW: The patient experience and health outcomes. *N Engl J Med* 2013; 368:201-3
- Barbara AM, Loeb M, Dolovich L, Brazil K, Russell M: Agreement between self-report and medical records on signs and symptoms of respiratory illness. *Prim Care Respir J* 2012; 21:145-52
- Bergmann MM, Calle EE, Mervis CA, Miracle-McMahill HL, Thun MJ, Heath CW: Validity of self-reported cancers in a prospective cohort study in comparison with data from state cancer registries. *Am J Epidemiol* 1998; 147:556-62
- Haapanen N, Miilunpalo S, Pasanen M, Oja P, Vuori I: Agreement between questionnaire data and medical records of chronic diseases in middle-aged and elderly Finnish men and women. *Am J Epidemiol* 1997; 145:762-9
- Lampe FC, Walker M, Lennon LT, Whincup PH, Ebrahim S: Validity of a self-reported history of doctor-diagnosed angina. *J Clin Epidemiol* 1999; 52:73-81
- Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ: Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *J Clin Epidemiol* 2004; 57:1096-103
- Smith B, Chu LK, Smith TC, Amoroso PJ, Boyko EJ, Hooper TI, Gackstetter GD, Ryan MA; Millennium Cohort Study Team: Challenges of self-reported medical conditions and electronic medical records among members of a large military cohort. *BMC Med Res Methodol* 2008; 8:37
- St Sauver JL, Hagen PT, Cha SS, Bagniewski SM, Mandrekar JN, Curoe AM, Rodeheffer RJ, Roger VL, Jacobsen SJ: Agreement between patient reports of cardiovascular disease and patient medical records. *Mayo Clin Proc* 2005; 80:203-10
- Walker MK, Whincup PH, Shaper AG, Lennon LT, Thomson AG: Validation of patient recall of doctor-diagnosed heart attack and stroke: A postal questionnaire and record review comparison. *Am J Epidemiol* 1998; 148:355-61
- Barbara AM, Loeb M, Dolovich L, Brazil K, Russell ML: Patient self-report and medical records: Measuring agreement for binary data. *Can Fam Physician* 2011; 57:737-8
- Louie AD, Robison LL, Bogue M, Hyde S, Forman SJ, Bhatia S: Validation of self-reported complications by bone marrow transplantation survivors. *Bone Marrow Transplant* 2000; 25:1191-6
- Fränneby U, Gunnarsson U, Wollert S, Sandblom G: Discordance between the patient's and surgeon's perception of complications following hernia surgery. *Hernia* 2005; 9:145-9
- Haapaniemi S, Nilsson E: Recurrence and pain three years after groin hernia repair. Validation of postal questionnaire and selective physical examination as a method of follow-up. *Eur J Surg* 2002; 168:22-8
- Iyer R, Gentry-Maharaj A, Nordin A, Liston R, Burnell M, Das N, Desai R, Gornall R, Beardmore-Gray A, Hillaby K, Leeson S, Linder A, Lopes A, Meechan D, Mould T, Nevin J, Olaitan A, Rufford B, Ryan A, Shanbhag S, Thackeray A, Wood N, Reynolds K, Menon U: Patient-reporting improves estimates of postoperative complication rates: A prospective cohort study in gynaecological oncology. *Br J Cancer* 2013; 109:623-32
- Mitchell DH, Swift G, Gilbert GL: Surgical wound infection surveillance: The importance of infections that develop after hospital discharge. *Aust N Z J Surg* 1999; 69:117-20
- Sands K, Vineyard G, Platt R: Surgical site infections occurring after hospital discharge. *J Infect Dis* 1996; 173:963-70
- Whitby M, McLaws ML, Doidge S, Collopy B: Post-discharge surgical site surveillance: Does patient education improve reliability of diagnosis? *J Hosp Infect* 2007; 66:237-42
- Jones D, Kazis L, Lee A, Rogers W, Skinner K, Cassar L, Wilson N, Hendricks A: Health status assessments using the Veterans SF-12 and SF-36: Methods for evaluating outcomes in the Veterans Health Administration. *J Ambul Care Manage* 2001; 24:68-86
- Mahoney FI, Barthel DW: Functional evaluation: The Barthel Index. *Md State Med J* 1965; 14:61-5
- Alazzawi S, Bardakos NV, Hadfield SG, Butt U, Beer ZH, Field RE: Patient-reported complications after elective joint replacement surgery: Are they correct? *J Bone Joint Surg Br* 2012; 94:1120-5
- Dushey CH, Bornstein LJ, Alexiades MM, Westrich GH: Short-term coagulation complications following total knee arthroplasty: A comparison of patient-reported and surgeon-verified complication rates. *J Arthroplasty* 2011; 26:1338-42
- Greenbaum JN, Bornstein LJ, Lyman S, Alexiades MM, Westrich GH: The validity of self-report as a technique for measuring short-term complications after total hip arthroplasty in a joint replacement registry. *J Arthroplasty* 2012; 27:1310-5
- Murff HJ, FitzHenry F, Matheny ME, Gentry N, Kotter KL, Crimin K, Dittus RS, Rosen AK, Elkin PL, Brown SH, Speroff T: Automated identification of postoperative complications within an electronic medical record using natural language processing. *JAMA* 2011; 306:848-55

31. Lof L, Berggren L, Ahlström G: Severely ill ICU patients recall of factual events and unreal experiences of hospital admission and ICU stay—3 and 12 months after discharge. *Intensive Crit Care Nurs* 2006; 22:154–66
32. Chen Y, Cai A, Fritz BA, Dexter F, Pryor KO, Jacobsohn E, Glick DB, Willingham MD, Escallier KE, Winter AC, Avidan MS: Amnesia of the operating room in the B-Unaware and BAG-RECALL clinical trials. *Anesth Analg* 2016 [Epub ahead of print]
33. Institute of Medicine Committee on Health and Literacy: *Health Literacy: A Prescription to End Confusion*. Washington, National Academies Press, 2004
34. Makaryus AN, Friedman EA: Patients' understanding of their treatment plans and diagnosis at discharge. *Mayo Clin Proc* 2005; 80:991–4
35. Kripalani S, Jackson AT, Schnipper JL, Coleman EA: Promoting effective transitions of care at hospital discharge: A review of key issues for hospitalists. *J Hosp Med* 2007; 2:314–23
36. Paasche-Orlow M: Caring for patients with limited health literacy: A 76-year-old man with multiple medical problems. *JAMA* 2011; 306:1122–9
37. Krosnick JA: Survey research. *Annu Rev Psychol* 1999; 50:537–67