VALIDATION OF 30-DAY POSTOPERATIVE MORTALITY AFTER HIP FRACTURE REPAIR AS A POTENTIAL QUALITY MEASURE FOR PUBLIC REPORTING

Prepared for the Office of Statewide Health Planning and Development, Healthcare Outcomes Center

By
Patrick S. Romano, M.D., M.P.H.
University of California-Davis
Health System

Spring 2014

TABLE OF CONTENTS

SECTION 1: EXECUTIVE SUMMARY 4
Validation Questions and Key Findings 4

SECTION 2: METHODOLOGY 9
General 9
Sampling hospitals and patients 9
Medical record abstraction 10
Data Analysis 11

SECTION 3: RESULTS 23

SECTION 4: CONCLUSIONS AND RECOMMENDATIONS 39

Section 1: Executive Summary

The Healthcare Outcomes Center at the Office of Statewide Health Planning and Development (OSHPD) develops public reports comparing hospital outcomes for selected medical conditions and surgical procedures for patients treated in hospitals throughout California. Over the last decade, OSHPD has regularly reported risk-adjusted hospital mortality rates for heart attack, community-acquired pneumonia, and heart bypass surgery.

This technical report, prepared for OSHPD, summarizes the validation of 30-day postoperative mortality after hip fracture as a potential quality measure for public reporting to Californians. The Hip Fracture Mortality Validation Study was designed by the UC Davis research team in collaboration with OSHPD staff, the AB 524 Technical Advisory Committee, and a Clinical Advisory Panel created specifically for this project. It was designed to address a variety of concerns, specified in detail later in this report, about the validity of using hospital-reported International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes in the California Patient Discharge Data Set to report publicly on hospital performance.

The original methodology for estimating risk-adjusted 30-day postoperative hip fracture mortality was developed in 1997-1998 by Grace Carter and colleagues at RAND, using 1992-1994 data from the OSHPD's Patient Discharge Data Set. This developmental work is fully described in a separate report.[1] We were subsequently commissioned by OSHPD to update the analyses performed by RAND and to validate their methods, using 1995-1996 data.

Validation Questions and Key Findings

The primary purpose of the validation study was to evaluate the impact of errors in hospitals' reporting of risk factors and unmeasured risk factors on risk-adjusted outcomes estimated using OSHPD's hospital discharge data. This is known as an evaluation of criterion validity, because data of uncertain accuracy are compared against data that are known or believed to be more accurate. These latter data represent a "criterion standard," which is uniformly applied across all hospitals, even though it may not be perfect.

The validation study was also designed to provide some information about whether measurable differences in the process of care explain, in part, observed differences in risk-adjusted outcomes. This is known as an evaluation of construct validity, because it is based on the construct (or conceptual model) that better processes of care should lead to better outcomes, and hence worse outcomes should result from worse processes of care. To the extent that there are strong associations between process measures and risk-adjusted outcome measures, we become more confident that both sets of measures describe true quality of care.

[1] Carter GM. *A Risk Adjusted Model for Mortality Following Surgery for Hip Fracture*. Los Angeles, CA: RAND; 1998. http://www.rand.org/pubs/drafts/2008/DRU1871.pdf ; accessed September 1, 2013.

The basic design was a retrospective cohort study of elderly patients who were admitted to acute care nonfederal hospitals in California for hip fracture in 1995 or 1996, and who underwent surgical repair during that hospitalization. These patients were followed for 30 days after admission to ascertain all deaths, using both OSHPD's Patient Discharge Data Set and linked Vital Statistics death records. Hospitals, and cases within hospitals, were randomly sampled using a two-stage, stratified cluster method. To ensure sufficient statistical power to answer the questions listed below, we oversampled hospitals with lower or higher than expected death rates, and patients who died. After obtaining a complete photocopy of each sampled record, we collected data through detailed, independent review by a coding professional and a research nurse, with physician back-up as needed.

The remainder of this Section outlines the specific research questions addressed by this study, and summarizes the key findings.

Question 1. How well do the data reported to OSHPD capture whether a hip fracture occurred and whether the fracture was surgically repaired?

We received 1,007 of the 1,047 records that we requested from participating hospitals (96.2%). Only six of these records (0.6%) did not meet inclusion criteria, either because of incorrect or incomplete coding or because of the timing of the hip fracture. Three patients had a fracture of another bone (e.g., femoral shaft) without a hip fracture, one patient's hip fracture occurred after admission to the hospital (e.g., not present on admission), and two patients' fractures were periprosthetic. We conclude that OSHPD's Patient Discharge Data may be used to identify acute hip fractures among elderly patients.

Question 2. What is the statewide reporting accuracy for risk factors in the risk adjustment models, using California patient discharge data?

Previous OSHPD validation studies have demonstrated that the validity of risk factor data varies significantly, depending on the severity and importance of the risk factor. If certain risk factors are widely underreported or overreported by hospitals, then risk-adjustment models that include those factors may be biased. For example, if hospitals only report the most severe cases of a risk factor, then that risk factor may appear to have a much greater impact on patient outcomes than it actually does. To identify risk factors that may be coded too poorly to include in risk-adjustment models, professional coding experts completed a "blind" recoding of each discharge in the validation study. We compared their findings with the information that hospitals reported to OSHPD, on a patient-by-patient basis.

Overall, 24.4% of hip fractures statewide had one or more omitted (false negative) risk factors, from among the 10 risk factors in the 30-day mortality model. Conversely, 16.1% of hip fractures had one or more unconfirmed (false positive) risk factors. With three exceptions (i.e., chronic renal failure, fall from bed or chair, high impact injury), two of which relate to external

cause-of-injury [E] codes, coding of risk factors included in the 30-day mortality model was found to be moderately sensitive (>73%) and predictive (>72%). With two uncommon exceptions (i.e., complicated diabetes, high-risk cancer), present-on-admission coding of risk factors included in the 30-day mortality model was also valid, with at least 90% agreement between original and recoded data. Serious coding validity problems were found to affect several important risk factors that were not included in the risk-adjustment model. This finding affirmed the decisions of researchers at RAND and UC Davis who developed the model.

Question 3. Is there a significant difference in the coding of important risk factors when comparing hospitals with significantly fewer deaths than expected, hospitals with significantly more deaths than expected, and hospitals that aren't outliers?

The ICD-9-CM coding of all key risk factors was examined by hospital outlier status, to determine whether variation in coding practices may explain why some hospitals appear to have better-than-expected outcomes and others appear to have worse-than-expected outcomes. In other words, hospitals that underreport risk factors may have relatively low expected death rates, and thus relatively high risk-adjusted death rates, because their patients do not appear as sick as they actually are. To address this question, we compared the accuracy of risk factor reporting across hospitals stratified by whether they had fewer deaths than expected, more deaths than expected, or neither.

We found several statistically significant differences in reporting of risk factors between hospitals with more deaths than expected and hospitals with fewer deaths than expected, but no consistent pattern. For example, the percentage of cases with one or more omitted (false negative) risk factors, from among the 10 risk factors in the 30-day mortality model, was 29.9% at high mortality outlier hospitals, 23.9% at non-outlier hospitals, and 29.4% at low-mortality outlier hospitals. Conversely, the percentage of cases with one or more unconfirmed (false positive) risk factors was 14.6% at high-mortality hospitals, 16.3% at non-outlier hospitals, and 13.8% at low-mortality outlier hospitals.

Substituting recoded data for administrative data in our multivariate model to estimate mortality risk had modest impact. For example, hospitals that had 50% more deaths than expected using administrative data still had 33% more deaths than expected using recoded data, and hospitals that had 66% fewer deaths than expected using administrative data still had 67% fewer deaths than expected using recoded data. We conclude that differential reporting of risk factors in the OSHPD's patient discharge data accounts for little of the observed variation in risk-adjusted death rates across hospitals. There is minimal evidence of systematic underreporting or overreporting of coded risk factors across risk-adjusted mortality strata.

Question 4. How does the risk-adjustment model change when additional clinical variables are used as risk factors?

Administrative datasets provide limited data, based on ICD-9-CM codes, to characterize patients' risk of death. To assess how much risk-adjustment models using the OSHPD's data could be improved through supplementation with more detailed clinical data, our nurse abstractors collected key clinical data elements from all sampled medical records. These additional clinical variables were identified through a literature review and through discussions with a Clinical Advisory Panel. This information was then used, in part, to determine whether more complete clinical information would improve the validity of our risk-adjustment model for 30-day mortality.

Adding several clinical data elements abstracted from medical records modestly improved the predictive performance of risk-adjustment models for 30-day mortality, at both the patient and hospital levels. These clinical data elements include tachycardia (high heart rate) at admission, hypothermia (low temperature) at admission, hypokalemia (low serum potassium), leukocytosis (high white blood cell count), and compromised ambulation prior to admission. Past history of diabetes mellitus was also an independent predictor of postpartum death, even though this information should already have been captured using ICD-9-CM diagnosis codes for diabetes, presumably because coders were unable to identify all patients who actually had diabetes.

Overall, adding information from clinical abstraction of medical records improved the discrimination of predictive models for 30-day death from c=0.716 to c=0.746. In other words, clinical data increased the percentage of randomly selected, paired hip fracture deaths and survivors who would be correctly "predicted" or classified by the model from 71.6% to 74.6%. The model's ability to explain hospital mortality rates, expressed as a partial R^2 statistic, increased from 9.7% using ICD-9-CM diagnoses reported to OSHPD, and 6.9% using recoded data, to 17.1% using both coded diagnoses and clinical data. However, models based on ICD-9-CM coded data generated risk-adjusted mortality rates that were highly correlated (Spearman r=0.837 using OSHPD-reported diagnoses; r=0.876 using recoded diagnoses) with those generated using a "gold standard" model based on both diagnoses and clinical data. Hospitals that had 33% more deaths than expected using recoded administrative data had only 13% more deaths than expected using additional clinical data, whereas hospitals that had 67% fewer deaths than expected using recoded administrative data still had 68% fewer deaths than expected using additional clinical data.

We conclude that adding clinical risk factors would significantly improve the performance of risk-adjustment models for 30-day mortality after hip fracture, and would approximately double the explained variation in hospital-level outcomes, using either the ICD-9-CM diagnoses reported to OSHPD or recoded diagnoses. Adding these clinical risk factors and substituting recoded for OSHPD-reported ICD-9-CM codes, helped to explain higher risk-adjusted mortality at some, but not all, high-mortality outlier hospitals. However, the OSHPD model based on current administrative data does accurately identify low-mortality outlier hospitals.

Question 5. Are there meaningful differences in the process of care between hospitals with fewer deaths than expected and hospitals with more deaths than expected?

If a hospital's risk-adjusted 30-day postoperative hip fracture death rate is a valid quality indicator, then hospitals with low rates should manage these patients in ways that maximize the likelihood of successful outcomes. These management practices are also known as processes of care, because they describe the process by which nurses, physicians, and other health professionals provide care at the bedside. To address this question, we collected detailed information about medical and nursing management, although many important components of perioperative and postoperative care could not be ascertained from medical records. For example, we hypothesized that several physician and nurse behaviors known to increase the risk of postoperative infection would be more prevalent at hospitals with high risk-adjusted death rates than at hospitals with low risk-adjusted death rates. We added these process factors into our validated risk-adjustment model for hip fracture mortality, to assess the extent to which they explain the observed variation in mortality, even after adjusting for patient risk.

We found that some processes of care differed across risk-adjusted mortality strata, and nearly all of these differences were in the expected direction, but none was statistically significant. For example, prophylactic antibiotics were administered within the optimal time window (0-2 hours before start of surgery) to 54% of eligible hip fracture patients at high-mortality outlier hospitals, 65% of eligible hip fracture patients at non-outlier hospitals, and 67% of eligible patients at low-mortality outlier hospitals. Adding these measurable process factors to the risk-adjustment model for 30-day mortality improved model performance (e.g., from c=0.746 to c=0.788) at the patient level, but did not significantly improve the model's ability to explain hospital outcomes. In other words, the observed differences in processes of care, such as use and timing of prophylactic antibiotics, across hospital mortality strata were too small to explain differences in risk-adjusted mortality. We conclude that measurable, documented differences in processes of care account for little of the observed performance variation between low-mortality outlier and high-mortality outlier hospitals. However, the results do suggest some interventions that may help high-mortality outlier hospitals to improve patient outcomes.

Section 2: Methodology

General

This was a retrospective cohort study based on a two-stage, stratified cluster sample of hospitalizations for hip fractures found in the Office of Statewide Health Planning and Development's Patient Discharge Data Set. The sample was designed to support estimation of diagnosis coding accuracy, mean severity of illness, and clinical treatment practices for the entire population of patients who were treated for hip fracture in medium-to-large California hospitals during 1995 and 1996, as well as for the subpopulations of patients who received care at hospitals with higher than expected and lower than expected risk-adjusted mortality. In the following sections, the sampling and data collection methodologies are described in detail. We adapted these methods from previous work performed by Carter et al. at the RAND Corporation, under a previous contract with OSHPD. [1]

Sampling hospitals and patients

The initial sampling frame included all patients in the Patient Discharge Data Set who were discharged from a nonfederal acute care hospital between January 1, 1995, and December 1, 1996, were at least 65 years of age at admission, had a reparative procedure for a hip fracture (ICD-9-CM procedure codes 78.55, 79.05, 79.15, 79.25, 79.35, 79.45, 79.55, 81.51, 81.52), and had a principal diagnosis of hip fracture (ICD-9 CM 820.xx) or a related injury or condition (e.g., osteoporosis) with a secondary diagnosis of hip fracture. Patients without social security numbers were excluded due to incomplete follow-up after discharge, while patients admitted in 1994 were excluded due to incomplete follow-back. We also excluded patients with primary or metastatic cancer potentially involving the femur, late effects of prior hip fracture, aseptic necrosis, malunion or nonunion of prior hip fracture, fracture of the femoral shaft, or revision of prior arthroplasty. Finally, we excluded patients with significant head, chest, or abdominal trauma in the same episode of care. If a patient was transferred from one hospital to another, he or she was assigned uniquely to the first hospital in which reparative surgery was done. This sampling frame, after exclusions, included 38,939 hip fracture cases from 377 hospitals (after linkage of transfer records).

We then used multivariate logistic regression to estimate the probability of 30-day mortality for each patient in the sampling frame, adjusting for age (using 5-year strata to account for the nonlinear relationship between age and mortality), admission from a skilled nursing facility, number of hospitalizations in the previous 12 months, mechanism of injury (e.g., fall from bed or chair, fall while standing or walking, other activity), fracture site (e.g., femoral neck, intertrochanteric, subtrochanteric), 11 comorbidities (e.g., atrial fibrillation, chronic renal failure, congestive heart failure, chronic obstructive lung disease, dementia or delirium, diabetes with complications, valvular heart disease, concurrent fracture of other bone [most often wrist], high-risk cancer [without bony metastasis], and high impact injury), and statistically significant interactions involving age and gender. We selected these risk factors based on clinical literature reviews, discussions with a clinical advisory panel, and empirical analyses. This model was

validated through a split-sample approach, and showed adequate calibration and discrimination in the validation sample (c=0.729).

In the first stage of sampling, we excluded 127 low-volume hospitals with fewer than 50 eligible patients, leaving 250 hospitals eligible for sampling. Low-volume hospitals were excluded because their risk-adjusted mortality rates are inherently unreliable, and the primary aim of the study was to evaluate differences in diagnosis coding, severity of illness, and treatment across hospital strata defined by risk-adjusted mortality. Although 127 of 377 hospitals were excluded for this reason, these hospitals only had 3,190 hip fractures (8.2% of the statewide total during the study years). We classified each of the remaining hospitals as better than expected, worse than expected, or neither, based on the exact probability ($p<0.05$) that it experienced the observed number of deaths (or a more extreme number) by chance, given the expected number of deaths at that hospital. All 14 better-than-expected hospitals (low-mortality stratum), all nine worse-than-expected hospitals (high-mortality stratum), and 30 randomly selected hospitals with neither lower nor higher than expected mortality (intermediate-mortality stratum) were included in the study. The last set of hospitals was sampled with probability proportional to size to maximize the reliability of population-based estimates.

The second stage of sampling involved randomly sampling 390 patients from each of the three hospital strata, including 13 patients from each of 30 hospitals with neither lower nor higher than expected mortality. Among the hospitals in the high and low mortality strata, we sampled all 30-day deaths and a random sample of survivors. Among the hospitals in the third stratum, we oversampled deaths to achieve a target death rate of 21.4% with 13 cases per hospital. The administrator and director of medical records at each sampled hospital were contacted and invited to join the study. Despite being guaranteed anonymity, four hospitals refused to participate and were not replaced. Once a hospital consented to participate, its director of medical records was asked to provide complete photocopies of all sampled records. Overall, 1007 of 1047 requested records were received (96.2%).

Medical record abstraction

To inform the design of a medical record abstraction tool, we systematically reviewed the published English-language literature on risk factors for hip fracture mortality. We used PubMed to search for prospective and retrospective cohort studies (including randomized controlled trials, if appropriate) of surgically repaired patients in developed nations that were published between 1986 and 2001. Only studies that reported or allowed estimation of adjusted relative risks or odds ratios were included. The findings of this review are summarized in Table 1.

Based on this review, and with input from an interdisciplinary Clinical Advisory Panel, we then developed an abstraction and recording instrument that was programmed for direct computer data entry. We wrote detailed guidelines to accompany the abstraction instrument. The draft instrument and guidelines were pretested and all identified problems were addressed. The parameters assessed included age, gender, type of fracture, implants used, surgical approach, type of anesthesia, date and time of each procedure, antibiotic and DVT prophylaxis, use and timing

of physical and occupational therapy, and mortality. The data entry program included limited response options and logic checks to minimize data entry errors (e.g., age younger than 65 years, out-of-range dates).

The clinical abstractors were either registered nurses or health professional students working under the direct supervision of a registered nurse. All abstractors had prior experience in medical record review for utilization management or quality improvement, and were trained to review all components of the record to obtain the most accurate possible information about each hospitalization. When needed, original records were reviewed with internists experienced in perioperative consultation or orthopedic surgeons with expertise in hip fracture treatment. When key documents (e.g., operative note, radiology report, admission history and physical, discharge summary) were missing, we re-solicited hospitals to obtain these documents, but were unable to resolve a few cases with missing information. Clinical abstractors were tested and carefully monitored; at least 5% of records were overread by two abstractors to ensure more than 95% agreement throughout the project.

Data analysis

The univariate distribution of each variable was examined; implausible and missing values were verified or corrected (by registered nurse or physician review) whenever necessary. All analyses were weighted to account for our oversampling of high-mortality and low-mortality hospitals, and deaths within those hospitals. Weights were constructed as the inverse of the probability of sampling each record, adjusted for hospital withdrawals and nonresponse. Consequently, the weighted results provided unbiased estimates of patient characteristics and processes of care for all hip fractures at medium-to-high volume hospitals in California. We examined weighted proportions for variables of interest at hospitals grouped according to their risk-adjusted mortality classification. We compared differences in proportions using the first order Taylor series linear approximation for estimation of variance. This approximation was required because subjects were clustered within hospitals and therefore could not be considered independent observations. To account for the sampling design and the clustering of observations within hospitals, we applied the second-order Rao and Scott correction of the Pearson chi square statistic. This corrected chi square statistic was then converted into an F statistic by dividing it by its degrees of freedom. These procedures were implemented using the survey research options in Stata Release 6 (Stata Corporation, 1999, College Station, Texas).

Table 1. Risk factors for short-term mortality after hip fracture repair in developed nations, based on review of published retrospective and prospective cohort studies

Risk Factor	Time Frame	Risk Estimate**	Reference***
Demographics			
Age (increase per year)	In-hospital	1.04[a]	Myers et al., 1991
	90 days	1.07[a]	Lu-Yao et al., 1994a
		1.07[b]	Todd et al., 1995
		1.06	Cree et al., 2000[1]
70-79	30 days	1.33	O'Hara et al., 2000[2]
75-79 vs. 65-74		1.62[a]	Roos et al., 1996[3]
80-89		2.17[a]	O'Hara et al., 2000[2]
80-84 vs. 65-74		2.09[a]	Roos et al., 1996[3]
90+		4.44[a]	O'Hara et al., 2000[2]
90+ vs. 65-74		3.20[a]	Roos et al., 1996[3]
Sex (male vs. female)	In-hospital	1.6[a]	Myers et al., 1991
	30 days	1.6	Kiel et al., 1994
		0.46[a]	O'Hara et al., 2000[2]
		2.17[a]	Roos et al., 1996[3]
	90 days	1.4	Magaziner et al., 1989
		2.21[a]	Lu-Yao et al., 1994a
		2.88[b]	Todd et al., 1995
		4.00[a]	Cree et al., 2000
	6 mos.	1.5	Magaziner et al., 1989
		4.96[a]	Robbins, 1989

** Unless otherwise indicated, these figures represent estimates of the relative risk or odds ratio among those with the risk factor compared to those without the risk factor. Some variation may exist due to differences in adjustment for other risk factors or control variables across studies.

*** Complete list of references, pages 17-21.

[a] $p < 0.05$

[b] $p < 0.01$

[1] Population-based prospective inception cohort study of hip fractures (patient age >64 years without previous hip fracture in past 5 years, or underlying pathological condition to which the hip fracture could be attributed).

[2] Retrospective cohort study of consecutive hip fracture patients (age 60 and over) who underwent surgical repair at US hospitals between 1983 and 1993.

[3] Population-based retrospective study of Manitoba (1979-1992, n=10007) and New England (1984-1985, n=16206) residents (age 65 and over) who underwent surgical repair of a femoral neck fracture.

Risk Factor	Time Frame	Risk Estimate**	Reference***
	12 mos.	1.9 [a]	Magaziner et al., 1989
		1.90 [a]	Zuckerman et al., 1995
		2.3 [a]	Ray et al., 1990
		2.59 [a]	Sembo & Johnell, 1993
Race (black vs. white)	In-hospital	1.1	Myers et al., 1991
	90 days	1.01	Lu-Yao et al., 1994a
		1.5	Magaziner et al., 1989
	l2 mos.	1.2	Ray et al., 1990
		1.8	Magaziner et al., 1989
Race (white vs. nonwhite)	30 days	0.82	O'Hara et al., 2000[2]
Occupational prestige	90 days	0.92	Cree et al., 2000
Level of education	90 days	1.11	Cree et al., 2000
Prior residence (nursing home vs. non-nursing home)	30 days	1.94 [a]	O'Hara et al., 2000[2]
	90 days	1.39 [a]	Lu-Yao et al., 1994a
		0.95	Cree et al., 2000
		3.38	Holmberg et al., 1986
	12 mos.	2.88	Holmberg et al., 1986
Living with someone	30 days	1.3	Kiel et al., 1994
Region (Manitoba vs. New England)	30 days	1.35 [a]	Roos et al., 1996[3]
Comorbidities			
Malnutrition	In-hospital	6.01 [b]	Incaizi et al., 1994
	30 days	2.39 [a]	O'Hara et al., 2000[2]
Diabetes	30 days	1.33 [a]	Roos et al., 1996[3]
Diabetes Mellitus (treated)	30 days	1.41 [a]	O'Hara et al., 2000[2]
Venous disease	In-hospital	2.04 [a]	Incaizi et al., 1994
Other digestive system diseases (ICD 570-579)	In-hospital	3.6 [a]	Myers et al., 1991
Cardiovascular disease	In-hospital	1.7 [a]	Myers et al., 1991
	30 days	2.03 [a]	O'Hara et al., 2000[2]
	90 days	2.13 [b]	Todd et al.,1995

	6 mos.	3.15 [a]	Robbins, 1989
Atrial fibrillation	30 days	2.08 [a]	O'Hara et al., 2000[2]
Valvular disease	30 days	1.50 [a]	O'Hara et al., 2000[2]
Hypertensive disease	In-hospital	0.41 [a]	Myers et al., 1991
	30 days	0.96	O'Hara et al., 2000[2]
Stroke or transient ischemic attack	30 days	1.50 [a]	O'Hara et al., 2000[2]
Neoplasm	In-hospital	2.4 [a]	Myers et al., 1991
	6 mos.	3.54 [a]	Robbins, 1989
Disease of the musculoskeletal system and connective tissue	In-hospital	0.32 [a]	Myers et al., 1991
Diseases of the nervous system	In-hospital	0.67 [a]	Myers et al., 1991
	30 days	0.41 [a]	Kiel et al., 1994
Lung findings on physical exam	6 mos.	3.0 [a]	Robbins, 1989
Mental disorders (ICD 290-319)	In-hospital	0.45 [a]	Myers et al., 1991
	30 days	0.78	Kiel et al., 1994
Dementia	6 mos.	0.8	Magaziner et al., 1989
Disorientation/coma	30 days	3.77 [a]	Kiel et al., 1994
		2.37 [a]	O'Hara et al., 2000[2]
Delirium (exclusive of dementia)	30 days	3.1	Marcantonio et al., 2000[4]
	90 days	3.2 [a]	Magaziner et al., 1989
	6 mos.	3.5 [a]	Magaziner et al., 1989
		1.1	Marcantonio et al., 2000[3]
	12 mos.	3.1 [a]	Magaziner et al., 1989
Congestive heart failure	In-hospital	2.1 [a]	Myers et al., 1991
	30 days	32.3 [a]	Nettleman et al., 1996[5]
	30 days	2.88 [a]	Roos et al., 1996[3]
Angina	30 days	25.7 [a]	Nettleman et al., 1996[4]
COPD	30 days	11.1 [a]	Nettleman et al., 1996[4]

[4] Prospective cohort study of emergent admissions (patients age > 65 years) for surgical repair of hip fracture; delirium was diagnosed using the Confusion Assessment Methods algorithm.
[5] Retrospective study of Medicare hip fractures between 1/1/92 and 6/30/94 at Iowa hospitals with at least 25 hip fracture admissions annually and representing a range of crude mortality rates (mean 30 day mortality rate was 8.5%).

		1.46 [a]	O'Hara et al., 2000[2]
		1.44 [a]	Roos et al., 1996[3]
Cancer[+]	30 days	2.93 [a]	Roos et al., 1996[3]
Cerebrovascular disease and procedures	30 days	1.59 [a]	Roos et al., 1996[3]
Other comorbidities[±]	30 days	1.27 [a]	Roos et al., 1996[3]
2 of following pre-existing medical conditions: diabetes mellitus, congestive heart failure, cardiac arrhythmia, ischemic heart or obstructive pulmonary disease, ongoing anti-coagulation therapy	180 days	1.95 [a]	Zuckerman, et al., 1995
3 pre-existing medical conditions		2.57 [a]	Zuckerman, et al., 1995
Charlson comorbidity index (any vs. none)	30 days	2.77 [a]	O'Hara et al., 2000[2]
Sickness at admission 3.01 - 5.4	30 days	1.39	O'Hara et al., 2000[2]
Sickness at admission 7.5-8.4		2.66 [a]	O'Hara et al., 2000[2]
Sickness at admission >8.5		5.64 [a]	O'Hara et al., 2000[2]
ASA physical status III (vs. I or II)	30 days	2.54 [a]	O'Hara et al., 2000[2]
ASA physical status IV or V (vs. I or II)		6.64 [a]	O'Hara et al., 2000[2]
Prior hospital stay within 1 month of admit	30 days	2.41 [a]	O'Hara et al., 2000[2]
Clinical Findings			
Anemia	In-hospital	0.36 [a]	Myers et al., 1991
High blood urea nitrogen (>20 mg%)	6 mos.	2.74 [a]	Robbins, 1989
Low albumin (<3 mg%)	6 mos.	2.61 [a]	Robbins, 1989
	12 mos.	1.58 [a]	Sembo & Johnell, 1993
Medications (on admission to SNF)			
Cardiac medications	30 days	2.33 [a]	Kiel et al., 1994
Antidepressants	30 days	3.12 [a]	Kiel et al., 1994
Narcotics	30 days	1.86 [a]	Kiel et al., 1994
Functional Status			
Able to walk alone or with help	30 days	0.76	Kiel et al., 1994
ADL score (unit increase, where a higher score represents lower function)	90 days	1.10 [b]	Todd et al., 1995
Any ADL dependency[6]	30 days	2.23 [a]	Kiet et al., 1994

[+] Cancer includes ICD-9-CM Codes 140, 171.99, 174-208.99, 273.0, 273.3, V10.46, 60.5(P), 62.4-62.41(P).
[±] Other comorbidities includes peripheral vascular disease, renal disease, liver disease, hemiplegia or paraplegia.
[6] No dependency of any ADL (Activities of Daily Living; such as bathing, toileting, eating/feeding) prior to surgery.

Difficulty with shopping	12 mos.	3.02 [a]	Sembo & Johnell, 1993
Low mental status (in hospital MMSE)	90 days	0.97 [a]	Cree, et al., 2000
Prefracture social support	90 days	0.64	Cree, et al., 2000
Prefracture function	90 days	0.99	Cree, et al., 2000
Perception of health (SF-12)	90 days	0.75	Cree, et al., 2000
Hip Fracture Characteristics			
Site (intertrochanteric or petrochanteric or extracapsular vs. neck or intracapsular)	90 days	1.2	Magaziner et al., 1989
		1.2 [a]	Lu-Yao et al., 1994a
	6 mos.	1.2	Magaziner et al., 1989
	l2mos	1.5 [a]	Todd et al., 1995
		1.2	Magaziner et al., 1989
		1.1	Magaziner et al., 1989
		1.33	Keene et al., 1993
Petrochanteric	30 days	1.24 [a]	Roos et al., 1996[3]
Intertrochanteric vs. subtrochanteric	30 days	1.25 [a]	O'Hara et al., 2000[2]
Intertrochanteric vs. femoral neck	6 mos.	1.23	Fox et al., 1999[7]
Intertrochanteric vs. femoral neck	1 year	1.02	Fox et al., 1999
Process of Care			
Hemiarthroplasty vs. osteosynthesis	In-hospital	1.4	Holt et al., 1994
	30 days	1.53	Parker, 1992
		1.4	Lu-Yao et al., 1994b
		2.24 [a]	O'Hara et al., 2000[2]
	90 days	1.2 [a]	Lu-Yao et al., 1994a
		1.0	Lu-Yao et al., 1994b
	6 mos.	1.0	Lu-Yao et al., 1994b
		1.03	Parker, 1992
	12 mos.	1.0	Lu-Yao et al., 1994b
Other care[8] vs. osteosynthesis	90 days	3.2 [a]	Lu-Yao et al., 1994a
Uncemented vs. cemented hemiarthroplasty	In-hospital	0.63	Holt et al., 1994

[7] Prospective study of 923 elderly patients admitted for intertrochanteric and femoral neck hip fractures between 1984 and 1986

[8] Other care is defined as any treatment other than arthroplasty or osteosynthesis.

Anterior vs. posterior hemiarthroplasty	45 days	0.4 [a]	Lu-Yao et al., 1994b
Operative delay >3 days	In-hospital	8.25 [b]	Rogers et al., 1995[9]
	30 days	1.26 [a]	Roos et al., 1996[3]
	12 mos.	1.76[c]	Zuckerman et al., 1995
Operative delay for medical reasons[10]	30 days	2.04 [a]	O'Hara et al., 2000[2]
Internal fixation, other than pinning	30 days	2.47 [a]	O'Hara et al., 2000[2]
Internal fixation vs. arthroplasty[11]	6 mos	1.52	Fox et al., 1999[7]
	1 year	1.16	Fox et al., 1999[7]
Surgeon with > 5 yrs. experience vs. 1-5 yrs.	In-hospital	1.4	Holt et al., 1994
Preoperative transfusion	30 days	2.45 [a]	O'Hara et al., 2000[2]
Postoperative use of aspirin	30 days	0.24 [a]	Nettleman et al., 1996[4]
Hip replacement vs. another procedure	6 mos	1.275[d]	Burns, et al., 1999[12]

[9] Retrospective review of isolated low-impact hip fractures (patients age > 65 years) stratified into early (<24hrs), intermediate (24 to 72 hrs), and late (>72 hours) operative fixation.
[c] $p = 0.05$
[10] Length of time constituting an "operative delay" is undefined.
[11] Internal fixation vs. arthroplasty among femoral neck hip fracture patients only.
[d] $p > 0.10$
[12] Prospective cohort study of femoral neck fracture patients discharged alive, treated by hip replacement versus another procedure.

Table 1 reference list:

Ahmad LA, Eckhoff DG, and AM Kramer. "Outcome studies of hip fractures: A functional viewpoint." *Orthopaedic Review* 1994;23 (1):19-24.

Barnes B, and K Dunovan. "Functional outcomes after hip fracture." *Physical Therapy* 1987;67(1*):1675-9.*

Barrett-Connor E. "The economic and human costs of osteoportic hip fracture." *The American Journal of Preventive Medicine* 1995;98(supl2A):2A3 5-8S.

Bonar SK, ME Tinetti, Speechley M, et al. "Factors associated with short- versus long-term skilled nursing facility placement among community living-hip fracture patients." *Journal of the American Geriatrics Society* 1990;38:1139-44.

Burns RB, et al. "Do hip replacements improve outcomes for hip fracture patients?" *Medical Care* 1999;37:285-294.

Ceder L, Ekelund L, Inerot S, et al. "Rehabilitation after hip fracture in the elderly." *Acta Orthopaedica Scandinavica* 1979;50:681-8.

Cobey JC, JH Cobey, L Conant, et al. "Indicators of recovery from fractures of the hip." *Clinical Orthopaedics and Related Research* 1976;*117:258-63.*

Craik R. "Disability following hip fracture." *Physical Therapy 1994;74(5):387-98.*

Cree M, et al. "Mortality and institutionalization following hip fracture." *Journal of the American Geriatrics Society* 2000;48:283-288.

Dolk T. "Operation in hip fracture patients -analysis of the time factor." *Inquiry: The British Journal of Accident Surgery* 1990;20 :369-72.

Dolk T. "Hip fractures-treatment and early complications." *Upsala Journal of Medical Sciences* 1989a;94;195-207.

Dolk T. "Influence of treatment factors on the outcome after hip fractures." *Upsala Journal of Medical Sciences* 1989;94:209-221.

Dorr LD, R Glousman, ALS Hoy, et al. "Treatment of femoral neck fractures with total hip replacement versus cemented and noncemented hemiarthioplasty." *Journal of Arthroplasty* 1986;(1):21-8.

Fisher ES, JA Baron, DJ Malenka, et al. "Hip fracture incidence and mortality in New England." *Epidemiology* 1994;2 116-22.

Fitzgerald JF, PS Moore, and RS Ditus. "The care of elderly patients with hip fracture: Changes since implementation of the prospective payment system." *New England Journal of Medicine* 1988;319(21):1392-7.

Folman Y, R Gepstein, A Assaraf, et al. "Functional recovery after operative treatment of femoral neck fractures in an institutionalized elderly population." *Archives of Physical Medicine and Rehabilitation* 1994;75*:454-6.*

Ford AB. "Reducing the threat of hip fracture." *Am J Public Health* 1989;79(3):269-70.

Fox KM, Magaziner J, Hebel R, Kenzora JE, Kashner TM. "Intertrochanteric versus femoral neck hip fractures: differential characteristics, treatment and sequelae." *Journal of Gerontology, Series A.* 1999;54:M635-40.

Furstenberg AL, and MD Mezey. "Differences in outcome between black and white elderly hip fracture patients." *Journal of Chronic Diseases* 1-987;40(1):931-8.

Gebhard JS, HC Amstutz, DM Zinar, et al. "A comparison of total hip arthroplasty and hemiarthroplasty for treatment of acute fracture of the femoral neck." *Clinical Orthopaedics and Related Research* 1991;282:123-31.

Holmberg S, P Conradi, R Kalen, et al. "Mortality after cervical hip fracture: 3002 patients followed for 6 years." *Acta Orthopaedica Scandinavica* 1986;57:8-11.

Holmberg S, R Kalen, and K-G Thorugren. "Treatment and outcome of femoral neck fractures." *Clinical Orthopaedics and Related Research* 1987;218:42-52.

Holt EM, RA Evans, CJ Hindley, et al. "1000 femoral neck fractures: The effect of pre-injury mobility and surgical experience on outcome." *Injury* 1994;25(2):91-5.

Incalzi RA, 0 Capparella, A Gemma, et al. "Predicting in-hospital mortality after hip fracture in elderly patients." *The Journal of Trauma* 1994;36(1):79-82.

Jalovaara P, and H Virkkunen. "Quality of life after primary hemiarthroplasty for femoral neck fracture: A 6-year follow-up of 185 patients." *Acta Orthopaedica Scandinavica* 1991;62(3):208-17.

Jette AM, BA Harris, PD Cleary, et al. "Functional recovery after hip fracture." *Archives of Physical Medicine and Rehabilitation* 1987;86:73 *5-50.*

Katz S, AB Ford, RW Moskowitz, et al. "Studies of illness in the aged-the index of ADL: A

standardized measure of biological and physiological flinction." *Journal of the American Medical Association* 1963;185:914.

Keating JF, Robinson CM, Court-Brown CM, et al. "The effect of complications after hip fracture on rehabilitation." *Journal of Bone and Joint Surgery.* 1993;75(6):976.
Keeler EB, KL Kahn, and SS Bentow. "Assessing quality of care for hospitalized Medicare patients with hip fracture using coded diagnoses from the Medicare provider analysis and review files." *A Rand Note.* 1992.

Keene GS, MJ Parker, and GA Pryor. "Mortality and morbidity after hip fractures." *British Medical Journal* 1993;307:1248-50.

Kiel DP, A Eichorn, O Intrator, et al. "The outcomes of patients newly admitted to nursing homes after hip fracture." *American Journal of Public Health* 1994;84(8):1281-6.

Koval KJ, ML Skovron, GB Aharonoff, et al. "Ambulatory ability after hip fracture: A prospective study in geriatric patients." *Clinical Orthopaedics and Related Research* 1995;310:150-9.

Koval KJ, and JD Zuckerman. "Functional recovery after fracture of the hip." *Journal of Bone and Joint Surgery* 1994;76-A(5):75 1-8.

Kuokkanen HOM, and OL Korkala. "Factors affecting survival of patients with hip fractures." *Acta Orthopaedica Belgica* 1992;58(4):425-8.

Larsson S, S Friberg, and LI Hansson. "Trochanteric fractures: Mobility, complications, and mortality in 607 cases treated with the sliding screw technique." *Clinical Orthopaedics and Related Research* 1990;260:232-41.

Laskin RS, MA Gruber, AJ Zimmerman. "Intertrochanteric fractures of the hip in the elderly: A retrospective analysis of 236 cases." *Clinical Orthopaedics and Related Research* 1979;141:188-95.

Lu-Yao GL, JA Brown, JA Barrett, ES Fisher. "Treatment and survival among elderly Americans with hip fractures: a population-based study." *American Journal of Public Health* 1994a;84(8):1287-91.

Lu-Yao GL, Keller RB, Littenberg B, and JE Wennberg. "Outcomes after displaced fractures of the femoral neck: A meta-analysis of one hundred and six published reports." *Journal of Bone and Joint Surgery* 1994;76-A(1):1*5-25.*

Magaziner J, EM Simonsick, TM Kashner, et al. "Survival experience of aged hip fracture patients." *American Journal of Public Health* 1989;79(3):274-8.

Marcantonio ER, Flacker JM, Michaels M, Resnick NM. “Delirium is independently associated with poor functional recovery after hip fracture.” *Journal of the American Geriatrics Society* 2000;48:618-624.

Marottoli RA, LF Berkman, LM Cooney, *Jr.* "Decline in physical flinction following hip fracture." *Journal of the American Geriatrics Society* 1992;40:861-6.

Marottoli RA, LF Berkman, L Leo-Summers, LM Cooney Jr. "Predictors of mortality and institutionalization after hip fracture: The New Haven EPESE cohort." *American Journal of Public Health* 1994;84(l 1):1807-12.

Miller CW. "Survival and ambulation following hip fracture." *Journal of Bone and Joint Surgery* 1978;60-A(7):930-3.

Miller K, Atzenhofer K, Gerber G and M Reichel. "Risk prediction in operatively treated fractures of the hip." *Clinical Orthopaedics and Related Research* 1993;293:148-152.

Mossey JM, K Knott, and R Craik. "The effects of persistent depressive symptoms on hip fracture recovery." *Journal of Gerontology* 1990;45(5):M163-8.

Mossey JM, E Mutran, K Knott, et al. "Determinants of recovery 12 months after hip fracture: The importance of psychosocial factors.” *American Journal of Public Health* 1989;79(3):279-86.

Mullen JO, and NL Mullen. "Hip fracture mortality: A prospective, multifactorial study to predict and minimize death risk." *Clinical Orthopaedics and Related Research* 1992;280:2 14-22.

Myers AH, EG Robinson, ML Van Natta, et al. "Hip fractures among the elderly: Factors associated with in-hospital mortality." *American Journal of Epidemiology* 1991;134(10):1128-37.

Nettleman, MD, Alsip J, Scharder M, Schulte M. “Predictors of mortality after acute hip fracture.” *Journaal of General Internal Medicine* 1996;11:765-767.

Palmer RM, RM Saywell, TW Zollinger, et al. "The impact of the prospective payment system on the treatment of hip fractures in the elderly." *Archives of lnternal Medicine* 1989;149:2237-41.

Parker MJ. "Internal fixation or arthroplasty for displaced subcapital fractures in the elderly." *Inquiry: The British Journal of Accident Surgery* 1992;23(8):521-4.

Pitto RP. "The mortality and social prognosis of hip fractures." *International Orthopaedics* 1994;18:109-13.

Ray WA, MR Griffin, DK Baugh. "Mortality following hip fracture before and after implementation of the prospective payment system." *Archives of Internal Medicine* 1990;50:2109-14.

Robbins JA. "Hip fractures." *Medical Rounds* 1989;2(3):99-1 15.

Robinson CM, D Saran, and IH Annan. "Intracapsular hip fractures: Results of management adopting a treatment protocol." *Clinical Orthopaedics and Related Research* 1994;302:82-91.

Rogers FB, Shackford ST, Keller MS. Early fixation reduces morbidity and mortality in elderly patients with hip fractures from low-impact falls. *Journal of Trauma: Injury, Infection, and Critical Care.* 1995;39:261-265.

Sembo I, and 0 Jolmell. "Consequences of a hip fracture: A prospective study over 1 year." *Osteoporosis International* 1993;3:148-153.

Todd CJ, CJ Freeman, C Camilleri-Ferrante, et al. "Differences in mortality after fracture of hip: The East Anglian audit." *British Medical Journal* 1995;310:904-8.

United States Department of Health and Human Services (USDHHS), Public Health Service. Healthy People 2000. DHHS Publication No. (PHS) 91-50212, 1991.

Weatherall M. "One year follow up of patients with hip fracture of the proximal femur." *New Zealand Medical Journal* 1994; 107(983):308-9.

Wood DJ, GK Ions, JM Quinby, et al. "Factors which influence mortality after subcapital hip fracture." *Journal of Bone and Joint Surgery* 1992;74(2):199-202.

Zuckerman JD, ML Skovron, KJ Koval, et al. "Postoperative complications and mortality associated with operative delay in older patients who have a fracture of the hip." *Journal of Bone and Joint Surgery* 1995;*77-A(l0):1551-6.*

Section 3: Results

Question 1. How well do the data reported to OSHPD capture whether a hip fracture occurred and whether the fracture was surgically repaired?

We received 1,007 of the 1,047 records that we requested from participating hospitals (96.2%). Only six of these records (0.6%) did not meet inclusion criteria, either because of incorrect or incomplete coding or because of the timing of the hip fracture. Three patients had a fracture of another bone (e.g., ICD-9-CM 821.xx, fracture of other and unspecified parts of femur) without a hip fracture, one patient's hip fracture occurred after admission to the hospital (e.g., not present on admission), and two patients' fractures were periprosthetic. "Present on admission" status was not used in OSHPD's original definition of hip fractures, which predated the implementation of that data element in the Patient Discharge Data Set, but it should be used in future studies on this topic. Periprosthetic fractures are supposed to be coded using a different ICD-9-CM code (996.44, peri-prosthetic fracture around prosthetic joint) that is not included in OSHPD's definition of hip fractures. We conclude that OSHPD's Patient Discharge Data may be used to identify acute hip fractures among elderly patients.

The sample included 768 females (79%) and 233 males (21%). The average age of all patients was 82.7 years (range 65-99), with 71% of patients between the ages of 75 and 89 years. Of the 1,001 patients included in this analysis, 52% had sustained a femoral neck fracture (520 hips), 42% had sustained an intertrochanteric hip fracture (416 hips) and 6% had sustained a subtrochanteric hip fracture (65 hips). Of the femoral neck fractures, 85% were displaced and 15% were non-displaced. Operative treatment for these fractures included some type of hip arthroplasty in 68% and internal fixation in 32%. Among these arthroplasties, 73% were bipolar (of which 84% were cemented), 26.8% were unipolar (of which 57% were cemented), and 0.2% were total hip procedures. The vast majority of intertrochanteric fractures were treated with internal fixation (94%), whereas 5% were treated with hemiarthroplasty and 1% with a total hip arthroplasty. The fixation devices used for intertrochanteric hip fractures included compression screws (96%), an intramedullary rod (1.7%), or percutaneous pins (2.3%). Ninety two percent of subtrochanteric fractures were treated with internal fixation and 5% were treated with a hemiarthroplasty. The most common internal fixation device used was a compression screw (86%), followed by intramedullary rods (13%) and pins (1.4%). The results are summarized in Table 1.1.

Table 1.1: Characteristics of patients included in the OSHPD Hip Fracture Mortality Validation Study

Patient Characteristic	**Weighted value**
Age (years)	82.7
Female gender (%)	79
Mortality (%)	
30-day	5.2
180-day	10.4
Mean length of stay (days)	5.6
Timing of surgery	
Within 24hr of presentation (%)	68
24-48 hours after presentation (%)	24
> 48 hours after presentation (%)	8
Type of anesthesia	
General only (%)	57
Regional only (%)	40
Combined general and regional (%)	3
Surgical procedure	
Internal fixation (%)	62.1
Hemiarthroplasty (%)	37.5
Total hip arthroplasty (%)	0.4
Fracture type	
Subcapital and/or femoral neck (%)	51
Intertrochanteric only (%)	43
Subtrochanteric +/-Intertrochanteric (%)	6
Femoral head (%)	4
Displaced fracture (% among femoral neck fractures)	82

Question 2. What is the statewide reporting accuracy for risk factors in the risk adjustment models, using California patient discharge data?

Previous OSHPD validation studies have demonstrated that the validity of risk factor data varies significantly, depending on the severity and importance of the risk factor. If certain risk factors are widely underreported or overreported by hospitals, then risk-adjustment models that include those factors may be biased. For example, if hospitals only report the most severe cases of a risk factor, then that risk factor may appear to have a much greater impact on patient outcomes than it actually does. To identify risk factors that may be coded too poorly to include in risk-adjustment models, professional coding experts completed a "blind" recoding of each discharge in the validation study. We compared their findings with the information that hospitals reported to OSHPD, on a patient-by-patient basis (Tables 2.1-2.5).

Overall, 24.4% of hip fractures statewide had one or more omitted (false negative) risk factors, from among the 10 risk factors in the 30-day mortality model. Conversely, 16.1% of hip fractures had one or more unconfirmed (false positive) risk factors. With three exceptions (i.e., chronic renal failure, fall from bed or chair, high impact injury), two of which relate to external cause-of-injury [E] codes, coding of risk factors included in the 30-day mortality model was found to be moderately sensitive (>73%) and predictive (>72%). With two uncommon exceptions (i.e., complicated diabetes, high-risk cancer), present-on-admission coding of risk factors included in the 30-day mortality model was also valid, with at least 90% agreement between original and recoded data. Serious coding validity problems were found to affect several important risk factors that were not included in the risk-adjustment model. This finding affirmed the decisions of researchers at RAND and UC Davis who developed the model.

Table 2.1: Sensitivity, positive predictive value, and reliability of risk factors included in the risk-adjustment model, using independently recoded data as the gold standard

Risk Factor	**No. with risk factor by UCD recoding**	**Weighted sensitivity (%)**	**Weighted positive predictive value (%)**	**Weighted kappa**
Atrial fibrillation	137	94.4	82.9	0.867
Chronic renal failure	23	51.3	70.2	0.587
Congestive heart failure	181	73.4	81.1	0.736
COPD/bronchitis/emphysema	183	77.6	86.2	0.780
Dementia/delirium	249	79.7	93.0	0.818
Diabetes, complicated	20	84.0	52.5	0.642
Fall from bed/chair (E-code)	102	54.6	72.2	0.592
Heart valve disease	72	85.8	71.9	0.764
High risk cancer, not lung	11	99.5	94.7	0.970
Other bone fracture	91	77.8	76.1	0.755
Pertrochanteric facture (intertroch/subtroch)	491	94.6	98.5	0.928
Violent event (E-code)	98	54.6	80.2	0.613

Table 2.2: Sensitivity, positive predictive value, and reliability of risk factors not included in the risk-adjustment model, using independently recoded data as the gold standard

Risk Factor	**No. with risk factor by UCD recoding**	**Weighted sensitivity (%)**	**Weighted positive predictive value (%)**	**Weighted kappa**
Acute myocardial infarction	23	83.7	80.9	0.821
Acute renal failure	20	31.2	34.8	0.325
Alcohol use disorder	17	77.7	81.8	0.789
Cerebral degeneration	5	97.4	25.3	0.401
Cerebrovascular disease, acute	5	62.1	5.8	0.106
Cerebrovascular disease, chronic	16	67.5	62.0	0.641
Cerebrovascular disease, combined (acute or chronic)	20	70.8	49.4	0.574
Coma	7	74.8	37.3	0.496
Conduction/rhythm problem	132	59.9	45.4	0.447
Degenerative osteoarthritis	73	61.0	47.1	0.487
Depression	63	37.5	61.4	0.435
Diabetes, combined (complicated or uncomplicated)	124	82.3	95.6	0.873
Diabetes, uncomplicated	104	74.3	93.7	0.813
Hematologic-lymphatic cancer	2	100.0	100.0	1.000
Hip replacement, partial	397	91.9	99.0	0.927
Hip replacement, total	6	51.1	10.6	0.167
Open fracture	1	100.0	19.9	0.330
Open reduction of fracture/separation	201	96.4	36.8	0.349
Other cardiovascular disease	32	61.8	65.2	0.621
Parkinson’s disease	53	91.5	99.3	0.950
Peptic ulcer	15	46.5	28.2	0.343
Peripheral vascular disease	48	47.6	70.7	0.546
Seizure disorder	20	91.4	46.0	0.609
Tobacco use	13	52.8	20.4	0.271

Table 2.3: Sensitivity, positive predictive value, and reliability of selected risk factors and treatment-related variables, using nurses' clinical abstraction as the gold standard (showing impact of including comorbidity data from prior OSHPD records)

Risk Factor	**No. with risk factor by UCD reabstraction**	**Weighted sensitivity (%)**	**Weighted positive predictive value (%)**	**Kappa**
Cerebrovascular disease, combined				
Index record only	188	9.0	56.5	0.118
Index or prior record	188	20.6	60.1	0.248
Congestive heart failure				
Index record only	151	69.9	73.8	0.677
Index or prior record	151	70.4	72.0	0.668
COPD/bronchitis/emphysema				
Index record only	180	79.9	87.4	0.802
Index or prior record	180	82.0	81.1	0.777
Diabetes, combined				
Index record only	127	85.4	96.0	0.894
Index or prior record	127	86.4	91.8	0.879
Died in hospital	76	96.6	100.0	0.983
Hip replacement, partial	401	92.1	99.9	0.935
Hip replacement, total	6	87.9	10.6	0.184
Open reduction of fracture/epiphysis	597	81.2	100.0	0.762
Pertrochanteric fracture (Inter/Sub)	480	95.2	92.8	0.882

Table 2.4: Agreement on "present at admission" coding between OSHPD data and independently recoded data from UCD, for risk factors included in the risk-adjustment model (omitting risk factors based on E codes, for which "present at admission" status was not reported)

Risk Factor	**Present at admission**		**Number with "present at admission" reported in both data**	**Agreement on "present at admission"**	
	OSHPD data (%)	**Recoded data (%)**		**%**	**Kappa**
Atrial fibrillation	87%	82%	51	92%	0.67
Chronic renal failure	99%	100%	3	100%	
Congestive heart failure	88%	85%	58	93%	0.76
COPD/bronchitis/emphysema	98%	99%	58	100%	1.00
Dementia/delirium	99%	100%	85	98%	
Diabetes, complicated	96%	75%	3	67%	
Heart valve disease	99%	100%	21	100%	1.00
High risk cancer, not lung cancer	96%	100%	5	60%	
Other bone fracture	99%	93%	22	100%	1.00

Table 2.5: Agreement on "present at admission" coding between OSHPD data and independently recoded data from UCD, for risk factors not included in the risk-adjustment model (omitting risk factors based on E-codes, for which "present at admission" status was not reported)

Risk Factor	Present at admission		Number with "present at admission" reported in both data	Agreement on "present at admission"	
	OSHPD data (%)	Recoded data (%)		%	Kappa
Acute myocardial infarction	41%	22%	6	100%	1.00
Acute renal failure	49%	11%	7	100%	1.00
Alcohol use disorder	98%	100%	7	86%	.
Cerebral degeneration	98%	100%	2	100%	.
Cerebrovascular disease, acute	47%	.	0	.	.
Cerebrovascular disease, chronic	97%	100%	3	100%	.
Coma	46%	50%	2	50%	.
Conduction/rhythm problem	80%	57%	43	67%	0.33
Degenerative osteoarthritis	99%	100%	13	100%	1.00
Depression	97%	100%	11	100%	1.00
Diabetes, uncomplicated	98%	100%	34	100%	1.00
Hematologic-lymphatic cancer	100%	.	0	.	.
Open fracture	100%	.	0	.	.
Other cardiovascular disease	94%	100%	7	100%	1.00
Parkinson's disease	100%	100%	21	100%	1.00
Peptic ulcer	91%	83%	2	100%	.
Peripheral vascular disease	99%	100%	9	89%	.
Seizure disorder	97%	100%	6	100%	1.00
Tobacco use	99%	100%	4	100%	.

Question 3. Is there a significant difference in the coding of important risk factors when comparing hospitals with significantly fewer deaths than expected, hospitals with significantly more deaths than expected, and hospitals that aren't outliers?

The ICD-9-CM coding of all key risk factors was examined by hospital outlier status, to determine whether variation in coding practices may explain why some hospitals appear to have better-than-expected outcomes and others appear to have worse-than-expected outcomes. In other words, hospitals that underreport risk factors may have relatively low expected death rates, and thus relatively high risk-adjusted death rates, because their patients do not appear as sick as they actually are. To address this question, we compared the accuracy of risk factor reporting across hospitals stratified by whether they had fewer deaths than expected, more deaths than expected, or neither.

We found several statistically significant differences in reporting of risk factors between hospitals with more deaths than expected and hospitals with fewer deaths than expected, but no consistent pattern. For example, the percentage of cases with one or more omitted (false negative) risk factors, from among the 10 risk factors in the 30-day mortality model, was 29.9% at high mortality outlier hospitals, 23.9% at non-outlier hospitals, and 29.4% at low-mortality outlier hospitals. Conversely, the percentage of cases with one or more unconfirmed (false positive) risk factors was 14.6% at high-mortality hospitals, 16.3% at non-outlier hospitals, and 13.8% at low-mortality outlier hospitals (Table 3.1).

Substituting recoded data for administrative data in our multivariate model to estimate mortality risk had modest impact. For example, hospitals that had 50% more deaths than expected using administrative data still had 33% more deaths than expected using recoded data, and hospitals that had 66% fewer deaths than expected using administrative data still had 67% fewer deaths than expected using recoded data (Table 3.2). We conclude that differential reporting of risk factors in the OSHPD's patient discharge data accounts for little of the observed variation in risk-adjusted death rates across hospitals. There is minimal evidence of systematic underreporting or overreporting of coded risk factors across risk-adjusted mortality strata.

Table 3.1: Sensitivity of risk factors included in the risk-adjustment model, using independently recoded data as the gold standard, stratified by hospital mortality classification

Risk Factor	**Number with risk factor**			**Weighted sensitivity (%)**		
	High	**Neither**	**Low**	**High**	**Neither**	**Low**
Atrial fibrillation	52	46	39	81	96*	78
Chronic renal failure	8	8	7			
Congestive heart failure	76	58	47	58	75	64
COPD/bronchitis/emphysema	63	68	52	73	78	78
Dementia/delirium	97	88	64	81	80	74
Diabetes, complicated	11	3	6	80	95*	24
Fall from bed/chair (E code)	41	27	34	57	54	56
Heart valve disease	26	28	18	73	88	58
High risk cancer, not lung	4	4	3	100	100	90
Other bone fracture	41	22	28	76	79	66
Pertrochanteric facture (intertroch/subtroch)	182	167	142	90	95	96
Violent event (E code)	26	34	38	67	53	68

* Significantly different across strata, $p<0.05$

Table 3.2: Positive predictive value of risk factors included in the risk-adjustment model, using independently recoded data as the gold standard, stratified by hospital mortality classification

Risk Factor	**Number with risk factor**			**Weighted sensitivity (%)***		
	High	**Neither**	**Low**	**High**	**Neither**	**Low**
Atrial fibrillation	52	46	39	91	82	90
Chronic renal failure	8	8	7			
Congestive heart failure	76	58	47	86	81	86
COPD/bronchitis/emphysema	63	68	52	73	87	86
Dementia/delirium	97	88	64	94	93	92
Diabetes, complicated	11	3	6	80	50	49
Fall from bed/chair (E code)	41	27	34	77	71	89
Heart valve disease	26	28	18	72	72	65
High risk cancer, not lung	4	4	3	100	94	100
Other bone fracture	41	22	28	96	74	82
Pertrochanteric facture (intertroch/subtroch)	182	167	142	98	99	97
Violent event (E code)	26	34	38	79	81	78

* No significant differences across strata, $p<0.05$

Question 4. How does the risk-adjustment model change when additional clinical variables are used as risk factors?

Administrative datasets based on ICD-9-CM codes provide limited data to characterize patients' risk of death. To assess how much risk-adjustment models using the OSHPD's data could be improved through supplementation with more detailed clinical data, our nurse abstractors collected key clinical data elements from all sampled medical records. These additional clinical variables were identified through a literature review and through discussions with a Clinical Advisory Panel. This information was then used, in part, to determine whether more complete clinical information would improve the validity of our risk-adjustment model for 30-day mortality.

Adding several clinical data elements abstracted from medical records modestly improved the predictive performance of risk-adjustment models for 30-day mortality, at both the patient and hospital levels. These clinical data elements include tachycardia (high heart rate) at admission, hypothermia (low temperature) at admission, hypokalemia (low serum potassium), leukocytosis (high white blood cell count), and compromised ambulation prior to admission. Past history of diabetes mellitus was also an independent predictor of death, even though this information should already have been captured using ICD-9-CM diagnosis codes for diabetes, presumably because coders were unable to identify all patients who actually had diabetes.

Overall, adding information from clinical abstraction of medical records improved the discrimination of predictive models for 30-day death from c=0.716 to c=0.746 (Table 4.1). In other words, clinical data increased the percentage of randomly selected, paired hip fracture deaths and survivors who would be correctly "predicted" or classified by the model from 71.6% to 74.6%. The model's ability to explain hospital mortality rates, expressed as a partial R^2 statistic, increased from 9.7% using ICD-9-CM diagnoses reported to OSHPD, and 6.9% using recoded data, to 17.1% using both coded diagnoses and clinical data. However, models based on ICD-9-CM coded data generated risk-adjusted mortality rates that were highly correlated (Spearman r=0.837 using OSHPD-reported diagnoses; r=0.876 using recoded diagnoses) with those generated using a "gold standard" model based on both diagnoses and clinical data. Hospitals that had 33% more deaths than expected using recoded administrative data had only 13% more deaths than expected using additional clinical data, whereas hospitals that had 67% fewer deaths than expected using recoded administrative data still had 68% fewer deaths than expected using additional clinical data(Table 4.2)

We conclude that adding clinical risk factors would significantly improve the performance of risk-adjustment models for 30-day mortality after hip fracture, and would approximately double the explained variation in hospital-level outcomes, using either the ICD-9-CM diagnoses reported to OSHPD or recoded diagnoses. Adding these clinical risk factors, and substituting recoded for OSHPD-reported ICD-9-CM codes, helped to explain higher risk-adjusted mortality at some, but not all, high-mortality outlier hospitals. However, the OSHPD model based on current administrative data does accurately identify low-mortality outlier hospitals.

Table 4.1: Goodness of fit statistics for 30-day hip fracture mortality risk-adjustment models with original and reabstracted ICD-9-CM codes, as well as clinical variables

Risk-adjustment model	Weighted partial R-square between observed and expected mortality[1]	Weighted hospital-level correlation with best model[2]	Model characteristics
			C statistic[3]
OSHPD ICD-9-CM codes	0.097	0.837	0.716
Reabstracted ICD-9-CM codes	0.069	0.876	0.698
Clinical variables	0.141	Not estimated	0.732
Reabstracted ICD-9-CM codes plus clinical variables	0.171	1.000	0.746

[1] The weighted partial R^2 between observed and expected mortality is the proportion of variation in observed hospital mortality, across the 49 hospitals in the hip fracture validation study, that is explained by the risk-adjustment model's prediction of expected mortality (after factoring out hospital volume). Note that this estimation is based on as few as 13 patients per hospital, so it may exaggerate the role of random error. These proportions may be multiplied by 100 to obtain percentages.

[2] The weighted hospital-level Spearman correlation is the rank order correlation, across the 49 hospitals in the hip fracture validation study, between risk-adjusted mortality using a suboptimal model and risk-adjusted mortality using the "best" model, which includes both recoded ICD-9-CM and clinical data. Note that this estimation is based on as few as 13 patients per hospital, so it may exaggerate the role of random error.

[3] The C statistic is a measure of discrimination, or a model's ability to distinguish individuals who had a poor outcome from individuals who had a good outcome. It represents the proportion of all randomly selected pairs of observations with different outcomes in which the patient who died had a higher expected probability of death than the survivor. This statistic is equivalent to the area under a receiver operating characteristic curve, which plots sensitivity versus 1-specificity at various cutoff values for the expected probability.

Table 4.2: Weighted, indirectly standardized 30-day hip fracture mortality ratios[1] by hospital mortality category, using risk-adjustment models with original and reabstracted ICD-9-CM codes, as well as clinical variables

Risk-adjustment model	Hospital mortality category			Absolute difference in risk-adjusted mortality rates between better and worse hospitals
	Better	Neither	Worse	
OSHPD ICD-9-CM codes	0.34	1.02	1.50	1.78 - 7.76%
Reabstracted ICD-9-CM codes	0.33	1.03	1.33	1.71 - 6.84%
Clinical variables	0.34	1.07	1.29	1.68 - 6.41%
Reabstracted ICD-9-CM codes plus clinical variables	0.32	1.09	1.13	1.58 - 5.64%

[1] The indirectly standardized mortality ratio represents the ratio of observed to expected readmissions, based on a specific risk-adjustment model. Ratios greater than one indicate hospital strata where more than the expected number of patients were readmitted; ratios less than one indicate hospital strata where fewer than the expected number of patients were readmitted.

Question 5. Are there meaningful differences in the process of care between hospitals with fewer deaths than expected and hospitals with more deaths than expected?

If a hospital's risk-adjusted 30-day postoperative hip fracture death rate is a valid quality indicator, then hospitals with low rates should manage these patients in ways that maximize the likelihood of successful outcomes. These management practices are also known as processes of care, because they describe the process by which nurses, physicians, and other health professionals provide care at the bedside. To address this question, we collected detailed information about medical and nursing management, although many important components of perioperative and postoperative care could not be ascertained from medical records. For example, we hypothesized that several physician and nurse behaviors known to increase the risk of postoperative infection would be more prevalent at hospitals with high risk-adjusted death rates than at hospitals with low risk-adjusted death rates. We added these process factors into our validated risk-adjustment model for hip fracture mortality, to assess the extent to which they explain the observed variation in mortality, even after adjusting for patient risk.

We found that some processes of care differed across risk-adjusted mortality strata, and nearly all of these differences were in the expected direction, but none was statistically significant (Table 5.1). For example, prophylactic antibiotics were administered within the optimal time window (0-2 hours before start of surgery) to 54% of eligible hip fracture patients at high-mortality outlier hospitals, 65% of eligible hip fracture patients at non-outlier hospitals, and 67% of eligible patients at low-mortality outlier hospitals. Adding these measurable process factors to the risk-adjustment model for 30-day mortality improved model performance (e.g., from c=0.746 to c=0.788) at the patient level, but did not significantly improve the model's ability to explain hospital outcomes. In other words, the observed differences in processes of care, such as use and timing of prophylactic antibiotics, across hospital mortality strata were too small to explain differences in risk-adjusted mortality. We conclude that measurable, documented differences in processes of care account for little of the observed performance variation between low-mortality outlier and high-mortality outlier hospitals. However, the results do suggest some interventions that may help high-mortality outlier hospitals to improve patient outcomes.

Table 5.1: Weighted process of care factors across hospitals

Process of Care Variable	High Mortality Stratum	Intermediate Mortality Stratum	Low Mortality Stratum
Timing of surgery			
Within 24 hours of presentation	67	68	70
24-48 hours after presentation	25	23	23
Internal fixation (if femoral neck fracture)	27	32	33
Cement used (if hemiarthroplasty)	62	79	63
Anesthesia (general)	53	58	49
Antibiotic prophylaxis			
Any preoperative use	71	76	80
0-2 hours before surgery	54	65	70
None given pre-operatively	29	24	20
Venous thromboembolism (VTE) prophylaxis			
None	40	30	31
Pneumatic compression (IPC) only	15	14	19
Pharmacologic therapy*			
IPC + pharmacologic therapy*	19	23	23
First VTE Prophylaxis $\leq$ 1 day after surgery	58	68	67
VTE prophylaxis ordered at discharge	15	18	27
Blood transfusion performed	44	46	40
Beta blocker use	9	10	11
Among patients with CAD	7	15	20
Internal medicine/cardiology consultation	64	50	54
Among patients with CAD/CHF	83	67	68
Time to first mobilization after surgery			
0-1 days	50	49	54
2-3 days	38	36	36
$\geq$ 4 days	5	5	5
Never	7	10	5
Time to first ambulation after surgery			
0-1 days	26	24	28
2-3 days	43	51	43
Never	31	25	29
Physical therapy after surgery	96	95	96
Within 2 days after surgery	91	87	88

P-values for differences across all groups were > 0.05 in all cases.
VTE: venous thromboembolism. IPC: intermittent pneumatic compression. CAD: coronary artery disease. CHF: congestive heart failure. PT: physical therapy.
* Pharmacologic therapy includes unfractionated heparin, low molecular weight heparin, warfarin, or any combination of these three. Does not include aspirin.

Section 4: Conclusions and Recommendations

The key findings from the OSHPD Hip Fracture Mortality Validation Study are as follows:

1. OSHPD's Patient Discharge Data may be used to identify acute hip fractures among elderly patients, with less than 1% error.

2. Less than a quarter of hip fracture records had omitted (24.4%) or unconfirmed (16.1%) risk factors, based on a 30-day risk-adjusted mortality model with 10 predictors. With three exceptions (i.e., chronic renal failure, fall from bed or chair, high impact injury), two of which relate to external cause-of-injury [E] codes, coding of these risk factors was found to be moderately sensitive (>73%) and predictive (>72%).

3. With two uncommon exceptions (i.e., complicated diabetes, high-risk cancer), present-on-admission coding of risk factors included in the 30-day mortality model was also valid, with at least 90% agreement between original and recoded data. Serious coding validity problems were found to affect several important risk factors that were not included in the risk-adjustment model.

4. Although there were several statistically significant differences in reporting of risk factors across hospital strata, based on risk-adjusted mortality, these differences followed no consistent pattern. For example, the percentage of cases with an omitted (false negative) risk factor was 29.9% at high mortality outlier hospitals, 23.9% at non-outlier hospitals, and 29.4% at low-mortality outlier hospitals. Conversely, the percentage of cases with an unconfirmed (false positive) risk factor was 14.6% at high-mortality hospitals, 16.3% at non-outlier hospitals, and 13.8% at low-mortality outlier hospitals.

5. Differential reporting of risk factors accounts for little of the observed difference in risk-adjusted death rates between low-mortality and high-mortality outlier hospitals. For example, hospitals that had 50% more deaths than expected using administrative data still had 33% more deaths than expected using recoded data, and hospitals that had 66% fewer deaths than expected using administrative data still had 67% fewer deaths than expected using recoded data.

6. Adding several clinical data elements abstracted from medical records modestly improved the predictive performance of risk-adjustment models for 30-day mortality at the patient level, and approximately doubled the explained variation at the hospital level. These clinical data elements include tachycardia (high heart rate) at admission, hypothermia (low temperature) at admission, hypokalemia (low serum potassium), leukocytosis (high white blood cell count), and compromised ambulation prior to admission. Past history of diabetes mellitus was also an

independent predictor, even though this information should already have been captured using ICD-9-CM diagnosis codes for diabetes.

7. Nevertheless, models based on ICD-9-CM coded data generated risk-adjusted mortality rates that were highly correlated (Spearman r=0.837 using OSHPD-reported diagnoses; r=0.876 using recoded diagnoses) with those generated using a "gold standard" model based on both recoded diagnoses and clinical data. Hospitals that had 33% more deaths than expected using recoded administrative data had only 13% more deaths than expected using additional clinical data, whereas hospitals that had 67% fewer deaths than expected using recoded administrative data still had 68% fewer deaths than expected using additional clinical data.

8. Some processes of care differed across risk-adjusted mortality strata, and nearly all of these differences were in the expected direction, but none was statistically significant. Adding these factors to the risk-adjustment model for 30-day mortality improved its performance at the patient level, but did not significantly improve the model's ability to explain hospital outcomes. In other words, measurable, documented differences in processes of care account for little of the observed performance variation between low-mortality and high-mortality outlier hospitals.

Our recommendations based on these findings include the following:

1. OSHPD should proceed with publication of periodic reports on 30-day mortality following hip fracture repair at acute care hospitals in California. The first of these reports should use the most recent available data and an updated version of the risk-adjustment model tested in this Validation Study. This updated model would use risk factors reported as present on admission (or unknown) for each case.

2. Because this Validation Study demonstrated that classification of low-mortality outlier hospitals is more robust than classification of high-mortality outlier hospitals, OSHPD and the California Hospital Assessment and Reporting Task Force (CHART) should consider alternate classification and labeling schemes, such as: (1) designating low-mortality outlier hospitals, but not high-mortality outlier hospitals; (2) using alternative vocabulary such as "centers of excellence" or "California's best hospitals" to place emphasis on low-mortality outliers; (3) using a longer period of data, such as three years, to improve the stability of risk-adjusted mortality estimates and minimize the possibility of "regression to the mean"; or (4) using a more extreme probability threshold for designating high-mortality outliers than for designating low-mortality outliers. It should be recognized that options (3) and (4) are somewhat *ad hoc* solutions to the problem; we cannot provide any guarantee that either of these approaches would actually make the classification of low-mortality outliers more robust.

3. Although these findings were not statistically significant, the Validation Study

results suggest that hospitals should be encouraged to fully adhere to The Joint Commission's Surgical Care Improvement Project (SCIP) standards. For example, low-mortality outlier hospitals consistently (but not statistically significantly) outperformed nonoutlier and high-mortality outlier hospitals on delivery of prophylactic antibiotics within 2 hours before surgical incision, timely initiation of effective venous thromboembolism (VTE) prophylaxis, continuation of VTE prophylaxis at discharge, and use of prophylactic beta blockers among patients with known coronary artery disease.

4. OSHPD should submit this measure of 30-day mortality following hip fracture repair to the National Quality Forum (NQF) for review and potential endorsement, when the appropriate opportunity arises. There is currently no NQF-endorsed measure of hip fracture mortality except for AHRQ Inpatient Quality Indicator 19, which is focused on inpatient mortality. The latter measure is clearly inferior to OSHPD's measure, because 30-day mortality among elderly patients with hip fracture is over 2.3 times greater than inpatient mortality (e.g., 8.3% versus 3.5% based on 1999 data).[2] Reports based on inpatient mortality alone substantially underestimate the impact of this injury, and treatment thereof, on public health.

5. Given recent improvements in diagnostic coding, additional risk factors may qualify for inclusion in future analyses of hip fracture mortality. The sensitivity and positive predictive value of coding for these additional risk factors should be reevaluated in future record abstraction studies.

[2]Clark DE, DeLorenzo MA, Lucas FL, Wennberg DE. "Epidemiology and short-term outcomes of injured Medicare patients." *Journal of the American Geriatrics Society* 2004;52(12):2023-30.