UCSF

UC San Francisco Previously Published Works

Title

Accuracy of do not resuscitate (DNR) in administrative data.

Permalink

<https://escholarship.org/uc/item/92m0r2zk>

Journal

Medical care research and review : MCRR, 70(1)

ISSN

1077-5587

Authors

Goldman, L Elizabeth
Chu, Philip W
Osmond, Dennis
et al.

Publication Date

2013-02-01

DOI

10.1177/1077558712458455

Peer reviewed

Accuracy of Do Not Resuscitate (DNR) in Administrative Data

Medical Care Research and Review

70(1) 98–112

© The Author(s) 2013

Reprints and permission:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/1077558712458455

http://mcr.sagepub.com

**L. Elizabeth Goldman¹, Philip W. Chu¹,
Dennis Osmond¹, and Andrew Bindman¹**

Abstract

This article evaluates the accuracy of reporting do not resuscitate (DNR) orders in administrative data for use in risk-adjusted hospital assessments. We compared DNR reporting by 48 California hospitals in 2005 patient discharge data (PDD) with gold-standard assessments made by registered nurses (RNs) who reabstracted 1,673 records of patients with myocardial infarction, pneumonia, or heart failure. The PDD agreed with the RN reabstraction in 1,411 (84.3%) cases. The administrative data did not reflect a DNR order in 71 of 512 records where the RN indicated there was (14% false negative rates), and reflected a DNR order in 191 of 1,161 records where the RN indicated there was not (16% false positive rate). The accuracy of DNR was more problematic for patients who died, suggesting that hospital-reported DNR is problematic for capturing patient preferences for resuscitation that can be used for risk-adjusted outcomes assessments.

Keywords

do not resuscitate, administrative data, accuracy

The Center for Medicare and Medicaid Services (CMS) and many states use hospital discharge administrative data generated for billing purposes to assess quality of care based on risk-adjusted mortality (Agency for Healthcare Research and Quality, 2001; Hospital Outcomes Center, 2004; Iezzoni, 1997; Iezzoni et al., 1996). One concern is

This article, submitted to *Medical Care Research and Review* on September 15, 2011, was revised and accepted for publication on July 26, 2012.

¹University of California–San Francisco, San Francisco, CA, USA

Corresponding Author:

L. Elizabeth Goldman, Department of Medicine, University of California–San Francisco, 1001 Potrero Avenue, San Francisco, CA 94110, USA
Email: legoldman@medsfgh.ucsf.edu

whether judgments based on administrative data account for a patient's preference regarding cardiovascular resuscitation. The presumption is that if a hospital has a greater proportion of patients for whom physicians have recorded do not resuscitate (DNR) orders, the hospital's observed mortality rate could increase, and be misinterpreted as a quality deficit (Marrie et al., 2002).

Recording whether a DNR order was written during a hospitalization is not a routine element of most hospital administrative discharge data. Including this element could potentially improve risk-adjusted hospital assessments, but only if it is accurately recorded, and reflects a patient's preference prior to receiving hospital-based treatment (Chen, Sosnov, Lessard, & Goldberg, 2008; Shepardson, Gordon, Ibrahim, Harper, & Rosenthal, 1999; Tabak, Johannes, Silber, & Kurtz, 2005; Teno et al., 1995). Otherwise, the DNR order may reflect events that occur during the hospitalization.

Since 1999, the California Office of Statewide Health Planning and Development has taken a unique step in its administrative patient discharge data (PDD), and required all licensed acute care hospitals in the state to record a patient's DNR status at the time of admission (Healthcare Outcomes Center, 2006). California instructs hospitals to record a patient as having a DNR order only if it is signed by a physician within 24 hours of admission. California uses this time window to serve as a proxy to distinguish cases in which the DNR reflects a patient's preference regarding resuscitation on admission from those in which the order results from deterioration due to complications of the hospitalization. One potential use of this measure is to improve the calculation of risk-adjusted hospital assessments for public reporting.

Background

California instructs hospitals to record DNR orders within the first 24 hours of an admission based on the assumption that this time window reflects patient preferences regarding cardiac resuscitation prior to hospitalization and not any hospital complications of care. The attempt to capture this information in a valid and reliable way in hospital administrative data requires that physicians accurately document DNR orders that reflect their communications and assessments of patients in a timely manner. In addition, personnel employed in hospital medical record departments who abstract charts for reporting of administrative data must make an accurate assessment of the presence of a DNR order, and determine whether or not it was recorded within the first 24 hours of an admission.

Based on this understanding of the process, we hypothesized (a) that there would be variation in the accuracy of DNR reporting across hospitals, (b) that there would be inaccuracy based on the timing of when the DNR order was written, and (c) that accuracy would vary with patient characteristics. We were particularly interested in the accuracy of DNR in patients who died since the reporting of DNR in these patients would be more influential in risk-adjusted mortality assessments. We thought hospitals might be less likely to record DNR orders accurately in their administrative data in cases of death as physicians may not have time or motivation to complete documentation. Alternatively, hospitals may be more thorough in recording DNR orders in

administrative data for patients who ultimately died as a way to deflect a potential liability claim.

New Contribution

This study is the first large-scale evaluation of the accuracy of DNR reporting in administrative data. It addresses the question of whether California's approach to DNR reporting reflects the presence of a physician's order of DNR with sufficient accuracy to be valid in risk-adjusted mortality assessments. No prior study has described whether there is variability across hospitals in DNR reporting accuracy that could bias hospital mortality assessments, or assessed whether patient characteristics are associated with the accuracy of DNR reporting.

Method

As part of a larger study examining the accuracy of administrative data collected in California, we audited DNR reporting in the 2005 California Patient Discharge Data (PDD; Goldman, Chu, Osmond, & Bindman, 2011). We compared the accuracy of DNR reporting against a gold standard created from blinded abstraction performed by registered nurses (RNs) of corresponding medical records. The PDD includes patient demographic, diagnostic, DNR reporting, and disposition codes for approximately 3.7 million hospitalizations per year from all nonfederal California acute care hospitals. From a file that included hospital and patient identifiers, we selected a probability sample of records for review. For efficiency, we randomly selected hospitals from within the state in proportion to the number of eligible records with prespecified diagnoses, and then we sampled records using International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes corresponding to one of three principal diagnoses: acute myocardial infarction, community-acquired pneumonia, or congestive heart failure. We selected these three principal diagnoses because they occur frequently and are associated with relatively high mortality. California publicly reports hospital assessments for two of these conditions: acute myocardial infarction (Romano & Luft, 1996) and community-acquired pneumonia (Healthcare Outcomes Center, 2006; Hospital Outcomes Center, 2004).

To meet separate study objectives related to coding accuracy of other data elements in the PDD, we oversampled cases that had one of two prespecified acute secondary diagnoses specific to each of the three principal diagnoses. Among cases whose principal diagnosis was acute myocardial infarction, we oversampled cases that had shock or pulmonary edema listed as a secondary diagnosis. For community-acquired pneumonia, we oversampled cases with secondary diagnoses of septicemia or respiratory failure. For congestive heart failure cases, we oversampled cases with secondary diagnoses of acute myocardial infarction or acute renal failure. To prevent any single hospital from disproportionately influencing our sample, we capped the number of records

with each of the principal diagnoses and risk factors at 10 for a given hospital in the sample.

Data Extraction

To replicate the conditions under which California hospitals report the DNR data element, we adopted California's definition of whether a patient had a DNR order on admission:

a directive from a physician in a patient's current inpatient medical record instructing that the patient is not to be resuscitated in the event of a cardiac or pulmonary arrest. In the event of a cardiac or pulmonary arrest, resuscitative measures include, but are not limited to, the following: cardiopulmonary resuscitation (CPR), intubation, defibrillation, cardioactive drugs, and assisted ventilation.

Patients can request to have a "partial DNR" order, which consists of an order to restrict any of the aforementioned resuscitative measures. For example, a patient may decide that she would want to be intubated if she was in respiratory distress, but not receive defibrillation if her heart stopped. For the purposes of this study, a "partial DNR" order was considered a DNR order. The DNR order needed to be written (signed and dated) within the first 24 hours of the patient's admission (Haas, 2000). The RNs who reabstracted the information documented the date and time the physician signed the order. If the order was dated and signed within 24 hours of the admission, we considered the patient as having a DNR order that should have been recorded in the PDD. A patient was considered admitted (and thus we started the 24-hour clock) if there was a field on the face page of the medical record identifying the date and time that a patient was considered admitted or a dated and timed admission note from a physician (whichever was earlier). We obtained the following patient characteristics from the PDD: age in years, gender, death during the hospitalization, and number of specified comorbidities (Elixhauser, Steiner, Harris, & Coffey, 1998). We used number of comorbidities as a proxy for severity of illness.

Analysis

We calculated sensitivity, specificity, positive predictive value, and negative predictive value of reported DNR orders at 24 hours in the PDD using the RN reabstracts as our gold standard, and described the averages, weighted averages, and ranges of these test characteristics across hospitals in our sample. We defined sensitivity as the percentage of cases in which the PDD correctly recorded a DNR order at 24 hours among those that the RN confirmed a DNR order was recorded in the medical record within 24 hours of admission. Specificity was defined as the percentage of cases in the PDD identified as not having a DNR order recorded within 24 hours of admission

among those that the RN reported as not having a DNR order recorded in the medical record within 24 hours of admission. Positive predictive value was defined as the percentage of cases that the RN identified a DNR order in the medical record within 24 hours of admission among records reported in the PDD as having a DNR order within 24 hours of admission. The negative predictive value was defined as the percentage of cases in which the RN did not identify a DNR order in the medical record within 24 hours of admission among records that the PDD did not identify as having a DNR order within 24 hours.

As our main outcomes, we categorized the accuracy of the PDD's reporting of DNR using false negatives and false positives. We defined the false negative rate as the percentage of cases the PDD indicated that there was no DNR order within 24 hours of admission among the medical records in which the RN indicated a DNR was truly present (in other words: $1 - \text{sensitivity}$). We defined the false positive rate as the percentage of cases the PDD indicated that there was a DNR order present within 24 hours of admission among the medical records in which the RN indicated there was not a DNR order within 24 hours of admission (in other words: $1 - \text{specificity}$). The RN report of DNR at 24 hours was considered the gold standard of true positives, and those records where the RN did not find a DNR at 24 hours were considered true negatives.

Next we evaluated the effect of the timing of the DNR order on our findings, recognizing that the exact time and date of the admission can be subject to interpretation and may vary by hospital practice patterns. Therefore, we calculated false negative and false positive rates using a modified definition of DNR that expanded the time window that the RN could identify a DNR order. In other words, the gold standard was recalibrated as an RN-identified DNR order written and signed within 48 hours. Subsequent sensitivity analyses then expanded the time window to 72, 96, 120 hours, and last, to include DNR orders written at any time during the hospitalization.

We conducted descriptive bivariate analyses of the associations between each patient characteristic and the accuracy of the DNR order in the PDD. We used hierarchical logistic regression (SAS PROC GLIMMIX) to perform both bivariate and multivariate testing of the association of patient characteristics with DNR false negatives and false positives at the hospital level. We included a hospital random effect to control for potential clustering of reporting practices within a single hospital in addition to an indicator variable for the patient's principal diagnosis. We included hospital and chart weights that reflected the sampling strategy and trimmed the weights to two standard deviations to limit the effects of outlier weights (Cochran, 1977).

We tested the association between patient characteristics and DNR order reporting accuracy using the modified definitions of a true DNR that expanded the time window for determining whether DNR was truly present (RN abstraction) to see how these would affect predictors of DNR reporting accuracy. We expected that false negative rates would increase overall as the time window was expanded because the RN would be identifying DNR orders written later than 24 hours. However, because patients who die are more likely to have a DNR order (Shepardson, Youngner, Speroff, & Rosenthal,

1999), we expected to see an even higher rate of false negative DNR orders in patients who died as the time window of the RN assessment was increased but the hospital determination reflected in the PDD was fixed at 24 hours. By the same token, we hypothesized that false positive rates would decline overall as the time window expanded as the RN would identify DNR orders written later in the hospitalization. This would be especially true in those patients who die because DNR status would need to be documented in patients who die or resuscitation would be attempted. DNR orders written later in the hospital often reflect those with a clinical deterioration (Zingmond & Wenger, 2005). As false positives in the PDD can result from reporting true DNR orders written later in the hospitalization, we would expect that as the time window lengthens in which the RN could report a DNR order, the false positive rates would decrease especially for those patients who die. However, were we to find little decrease in the false positive rates even after expanding the time window to anytime during the hospitalization, then this would suggest that there are additional factors, for instance incomplete documentation, that account for false positives.

Results

Our sample consisted of 1,673 cases from 48 acute care hospitals in California (Table 1). This represented approximately 0.7% of the 243,253 admissions for acute myocardial infarction, congestive heart failure, and community-acquired pneumonia in California in 2005. The sampled patients were elderly (mean age 74.6 years) with a high mortality rate (29.0% died during the hospitalization). Among the 1,673 cases, 632 (37.8%) were recorded as having a DNR order by the hospitals, compared with 10.5% among the overall population of admissions for these medical conditions.

Using the RN reabstraction as a gold standard, the average unweighted sensitivity across hospitals was 83% (range 0%-100%). The average unweighted specificity of reporting DNR was 77% (range 0% to 100%). The average unweighted positive predictive value was 66% (range 0% to 100%), and the average unweighted negative predictive value was 91% (range 50% to 100%) across all hospitals. The unweighted and weighted averages were similar: 86% sensitivity, 84% specificity, 70% positive predictive value, and 93% negative predictive value.

The RN reabstraction of the 1,673 cases identified 512 (30.6%) with a physician DNR order written within 24 hours of admission. As expected, having a DNR order within 24 hours of an admission as determined by the RN abstractor was associated with an increased probability of death (33.8% of patients died with a DNR order as compared with 27.0% of patients without a DNR order; chi-square $p < .001$). The PDD agreed with the RN reabstraction in 1,411 (84.3%) cases. When compared with the gold standard RN reabstraction, the PDD did not report a DNR order in 71 of 512 cases where the RN identified a DNR order (14% false negative rates) and the PDD reported DNR orders in 191 of 1,161 cases where the RN did not identify a DNR order (16% false positive rate). Of the 191 false positive DNR orders, 89 (47%) had a DNR order written after 24 hours.

Table 1. Patient Characteristics by DNR Reporting Accuracy

Patient Characteristics	Sample Size	False Positives; n (%)	False Negatives; n (%)
Total	N = 1,673	191 (11.4)	71 (4.2)
Age (years)			
19-49	100	10 (10.0)	0 (0.0)
50-59	172	10 (5.8)	1 (0.6)
60-69	234	13 (5.6)	7 (3.0)
70-79	437	45 (10.3)	18 (4.1)
80-89	510	70 (13.7)	31 (6.1)
90+	220	43 (19.5)	14 (6.4)
Female	874	95 (10.9)	31 (3.5)
Male	799	96 (12.0)	40 (5.0)
No. of comorbidities			
0	67	10 (14.9)	2 (3.0)
1	196	26 (13.3)	9 (4.6)
2	283	35 (12.4)	13 (4.6)
3	355	49 (13.8)	20 (5.6)
4	292	29 (9.9)	11 (3.8)
5	218	21 (9.6)	5 (2.3)
6+	262	21 (8.0)	11 (4.2)
Alive at discharge	1187	121 (10.2)	39 (3.3)
Died in hospital	486	70 (14.4)	32 (6.6)
Principal diagnosis			
AMI	587	85 (14.5)	31 (5.3)
CAP	510	50 (9.8)	14 (2.7)
CHF	576	56 (9.7)	26 (4.5)

Note: DNR = do not resuscitate; AMI = acute myocardial infarction; CAP = community-acquired pneumonia; CHF = congestive heart failure.

Increasing the time window during a hospitalization that an RN could consider a DNR order as present increased the false negative rate as expected: from 17%, including DNR orders identified by the RN by 48 hours, to 32%, including DNR orders written at any time during hospitalization (Figure 1). In contrast, the false positive rates decreased when we expanded the time window that the RN could identify a DNR order; the false positive rate decreased from 14% at 48 hours to 11%, including DNR orders identified by the RN at any time during the hospitalization.

Death in the hospital was associated with both false negative and false positive DNR orders. Unadjusted, death increased the odds of having a false negative DNR order (odds ratio [OR] = 1.81; 95% CI [confidence interval] = 1.08-3.03; Table 2) and a false positive DNR order (OR = 1.69; 95% CI = 1.22-2.35; Table 3). These patterns persisted in the adjusted analysis; death in the hospital approached statistical significance for

Figure 1. Effect of expanding the time window for RN (registered nurse) identification of DNR (do not resuscitate) orders on false negative and false positive rates

false negatives (OR = 1.70; 95% CI = 0.95-3.04) and was a statistically significant risk of a false positive DNR order (OR = 1.77; 95% CI = 1.21-2.58). As anticipated, the association between death and false negatives increased when we expanded the time window in which the RN could identify a DNR order (Table 2) as did the association between death and false positives (Table 3).

In addition to death in the hospital, several other patient characteristics were associated with false positive reporting of DNR orders (Table 3). In the unadjusted analyses, increasing age (OR = 1.74; 95% CI = 1.51-2.00), female gender (OR = 1.37; 95% CI = 1.01-1.86), and being admitted with an acute myocardial infarction (OR = 1.59; 95% CI = 1.10-2.30) were associated with an increased odds of a false positive DNR order while having more comorbidities (OR = 0.87; 95% CI = 0.81, 0.95) was associated with a decreased odds of a false positive DNR order. In adjusted analysis, increasing age (OR = 1.80; 95% CI = 1.54-2.10) and admission for an acute myocardial infarction (OR = 1.59; 95% CI = 1.21-2.86) remained a statistically significant predictor of a false positive DNR order.

Discussion

Our study is the first, to our knowledge, to evaluate the accuracy of a DNR data element in administrative hospital discharge data. We found that the PDD’s reporting of DNR orders, defined as those recorded within the first 24 hours of a hospital admission, are somewhat inaccurate for three common medical diagnoses. Importantly, for use of this variable in risk-adjusted hospital assessments, DNR reporting accuracy varies across hospitals and is more problematic among patients who died. Expanding

Table 2. Characteristics Associated With False Negative DNR Orders

Characteristics (N = 512)	Sensitivity Analyses						
	Unadjusted, 24 Hours	Adjusted, 24 Hours	48 Hours	72 Hours	96 Hours	120 Hours	At Any Time
Age (years)	0.90 (0.70, 1.16)	0.81 (0.61, 1.08)	0.76 (0.60, 0.96)	0.69 (0.56, 0.85)	0.64 (0.53, 0.78)	0.65 (0.54, 0.79)	0.63 (0.53, 0.74)
Female	0.96 (0.58, 1.60)	0.98 (0.57, 1.71)	0.89 (0.55, 1.42)	1.04 (0.66, 1.62)	0.98 (0.64, 1.51)	0.95 (0.63, 1.44)	0.84 (0.59, 1.20)
No. of comorbidities	1.00 (0.87, 1.15)	1.03 (0.88, 1.20)	1.05 (0.92, 1.20)	1.05 (0.93, 1.19)	1.02 (0.91, 1.15)	1.07 (0.95, 1.20)	1.14 (1.03, 1.25)
Principal diagnosis							
AMI	1.29 (0.73, 2.29)	1.10 (0.59, 2.07)	1.06 (0.61, 1.85)	1.15 (0.67, 1.95)	1.23 (0.73, 2.08)	1.26 (0.75, 2.09)	1.07 (0.68, 1.67)
CAP	0.63 (0.32, 1.27)	0.46 (0.21, 0.99)	0.54 (0.28, 1.03)	0.67 (0.37, 1.22)	0.73 (0.41, 1.31)	0.74 (0.42, 1.31)	0.82 (0.51, 1.31)
CHF (ref)							
Died (yes)	1.81 (1.08, 3.03)	1.70 (0.95, 3.04)	1.79 (1.10, 2.93)	1.93 (1.22, 3.04)	2.10 (1.35, 3.27)	2.38 (1.56, 3.64)	3.16 (2.19, 4.57)

Note: DNR = do not resuscitate; AMI = acute myocardial infarction; CAP = community-acquired pneumonia; CHF = congestive heart failure.

Table 3. Characteristics Associated With False Positive DNR Orders

Characteristics (N = 1,161)	Unadjusted, 24 Hours	Adjusted, 24 Hours	Sensitivity Analyses				
			48 Hours	72 Hours	96 Hours	120 Hours	At Any Time
Age (years)	1.74 (1.51, 2.00)	1.80 (1.54, 2.10)	1.70 (1.45, 2.00)	1.65 (1.40, 1.95)	1.63 (1.38, 1.93)	1.65 (1.38, 1.96)	1.59 (1.31, 1.93)
Female	1.37 (1.01, 1.86)	1.21 (0.85, 1.72)	1.12 (0.77, 1.64)	1.13 (0.76, 1.68)	1.07 (0.72, 1.61)	1.10 (0.73, 1.67)	1.01 (0.62, 1.64)
No. of comorbidities	0.87 (0.81, 0.95)	0.91 (0.83, 1.01)	0.90 (0.81, 1.00)	0.92 (0.82, 1.03)	0.91 (0.82, 1.03)	0.91 (0.81, 1.02)	0.89 (0.78, 1.02)
Principal diagnosis							
AMI	1.59 (1.10, 2.30)	1.86 (1.21, 2.86)	1.67 (1.05, 2.66)	1.68 (1.03, 2.72)	1.77 (1.08, 2.93)	1.73 (1.05, 2.86)	1.86 (1.04, 3.32)
CAP	0.98 (0.65, 1.47)	1.66 (1.03, 2.69)	1.51 (0.90, 2.53)	1.49 (0.87, 2.54)	1.64 (0.94, 2.85)	1.47 (0.83, 2.58)	1.52 (0.79, 2.95)
CHF (ref)							
Died (yes)	1.69 (1.22, 2.35)	1.77 (1.21, 2.58)	1.67 (1.10, 2.54)	1.60 (1.03, 2.49)	1.62 (1.03, 2.57)	1.80 (1.12, 2.87)	2.41 (1.35, 4.32)

Note: DNR = do not resuscitate; AMI = acute myocardial infarction; CAP = community-acquired pneumonia; CHF = congestive heart failure.

the time window during which a hospital could be given credit for correctly reporting in the PDD that a patient had a DNR order during a hospitalization did not fully eliminate the observed errors in the PDD's reporting of DNR orders.

Overall, the average sensitivity and specificity across all hospitals is consistent with many other data elements in administrative data used in performance assessments, including many diagnostic codes (Romano, Chan, Schembri, & Rainwater, 2002; Romano, Yasmeen, Schembri, Keyzer, & Gilbert, 2005; Zrelak et al., 2011) and present-on-admission coding (De Coster, Li, & Quan, 2008; Goldman et al., 2011; Meddings, Saint, & McMahon, 2010). However, ICD-9 codes for costly procedures such as coronary artery bypass graft tend to have better test characteristics with positive predictive values of greater than 97% (Lee et al., 2011). We oversampled patients with high DNR rates, and in spite of this, our positive predictive value was relatively low; we would expect it to be even lower in hospitals with a lower prevalence of patients with DNR orders on admission.

In our study, there was substantial variation in DNR reporting across hospitals. While our study did not explore the potential mechanisms, we postulate that the errors may either be because of differences in how physicians at certain hospitals document DNR or because of the practices of their medical record personnel who abstract information used for administrative reporting to the state. Medical record abstractors interpreting physician documentation of DNR in notes instead of specific physician orders or reporting of incompletely documented DNR orders (i.e., not dated or signed) could lead to false positive reports of DNR orders. While California does use the PDD for creating public reports on hospital performance, it has not used the DNR variable. Therefore, it is unlikely that hospitals were purposely reporting either false negatives or false positives in the PDD to improve hospital performance assessments.

We also found that patients who died were more likely to have DNR coded inaccurately. Our finding that patients who died were more likely to have false negatives, and this association increased as we expanded the time window in which an RN could determine that a DNR order had been written, is consistent with the notion that the probability of a DNR order increases during a hospitalization in association with an increased probability of death (Hakim et al., 1996; Wenger et al., 1995). Our finding that false positives rates decrease as the time window expands, but that the expanded time window does not eliminate the association between false positives and death, suggests that the association is not simply because of reporting DNR orders written later in the hospitalization. Poor or incomplete documentation of DNR orders (i.e., not dated or signed) may be more likely to occur among patients who die. Patients participating in certain clinical programs such as palliative care might contribute toward an overinterpretation by hospital medical record abstractors reporting information to the PDD that a DNR order was recorded in the first 24 hours of admission when in fact a medical record review indicated that there was no such order.

The findings that DNR reporting accuracy was variable across hospitals raises questions about the validity of using DNR orders as reported in administrative data for calculating risk-adjusted hospital mortality rates. Hospital performance is typically judged by

the observed mortality rate relative to the expected mortality rate. Hospitals that have false positive DNR orders in patients who died would erroneously appear to have lower than expected mortality (better performance than they truly did). Conversely, those hospitals that have false negatives DNR orders among patients who died would erroneously appear to have a higher than expected mortality (worse performance than they truly did). While the intended effect of using a DNR reporting variable is to adjust for the potential impact of patient preference on hospital outcomes, the high rate of inaccuracy in this variable, especially among patients who die in the hospital, make it problematic for this purpose.

Extending the time window in which hospitals could report a DNR would worsen reporting biases. Moreover, it would also increase the likelihood that the DNR order reflects a deterioration in a patient's condition during the hospitalization as a result of a complication of care, rather than an accurate reflection of patient preferences on admission (Hakim et al., 1996). This would have the further unintended effect of assigning hospitals a higher expected death rate as a result of quality deficiencies in the hospital, wholly unrelated to patient preferences. Recognizing that inaccurate capture of the timing of a DNR order in administrative data is a common error, it may be possible to improve the accuracy of hospital reporting with a targeted education intervention focused on this issue with hospital medical record abstractors.

The findings from this study have implications beyond California, as the Centers for Medicare and Medicaid Services (CMS) released a new ICD-9 code, V49.86, in October 2010 to document DNR status in administrative data sets. CMS's new ICD-9 code that documents DNR orders throughout a given hospitalization does not have the same potential to distinguish whether a patient had a DNR order prior to hospitalization or as a result of the hospitalization. Prior research suggests that patients with DNR orders recorded at the beginning of the hospitalization and those at the end differ, with later DNR orders more likely to represent futility of care (Zingmond & Wenger, 2005). For this reason, as compared with California's definition of a DNR order being within 24 hours of admission, CMS's code of a DNR order anytime during the hospitalization does not appear to be a more reliable method of recording the presence of a hospital DNR order. The CMS code may be an even less valid way of capturing patient preference and patient morbidity independent of health complications that arise as a consequence of hospital care.

Our study has several limitations. First, it was conducted in a single state. However, California had been collecting DNR reporting in its administrative patient discharge data for more than a decade by the time we conducted our study so we do not believe that the error we measured is related to being early on a learning curve. Second, our sample was limited to three principal diagnoses and included a relatively sick cohort with multiple comorbidities and high inpatient mortality. Our findings may be less generalizable to other diagnoses or hospitals with a different case mix. Our study focused on the accuracy of reporting DNR in administrative data and did not address whether the DNR order accurately reflects patients' preferences to forego resuscitation. Finally, our data are from the only large-scale validation of DNR reporting in

administrative data of which we are aware. However, as it is from 2005, it is important to note that DNR reporting accuracy may have improved or changed at individual hospitals or programs in response to the adoption of the ICD-9 code for DNR. Future research is warranted to assess if the situation has changed over time.

There may be reasons beyond risk-adjusted hospital assessments to encourage administrative databases to include DNR reporting. Efforts are underway to try to improve opportunities for patients of different racial and ethnic backgrounds to express their preferences about DNR orders, which would allow for rates of use of DNR orders according to patient characteristics to be captured. However, the California experience suggests that reporting of DNR preferences on admission by hospitals is inaccurate, particularly for older patients and those who die during hospitalization. We therefore recommend caution in adopting this variable in risk-adjustment models until a strategy can be established for ensuring that these data are captured more accurately. Future work should explore ways to improve hospital reporting of this important information or investigate alternatives to reporting DNR in administrative data.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by an Agency for Healthcare Research and Quality K08 Mentored Clinical Scientist Development Award grant 1 K08 HS018090-01 and a NIH/NCRR/OD UCSF-CTSI grant KL2 RR024130.

References

- Agency for Healthcare Research and Quality. (2001). *Quality indicators—guide to prevention quality indicators: Hospital admissions for ambulatory care sensitive conditions*. Rockville, MD: Author.
- Chen, J. L., Sosnov, J., Lessard, D., & Goldberg, R. J. (2008). Impact of do-not-resuscitation orders on quality of care performance measures in patients hospitalized with acute heart failure. *American Heart Journal*, 156, 78-84.
- Cochran, W. (1977). *Sampling techniques* (3rd ed.). New York, NY: John Wiley.
- De Coster, C., Li, B., & Quan, H. (2008). Comparison and validity of procedures coded with ICD-9-CM and ICD-10-CA/CCI. *Medical Care*, 46, 627-634.
- Elixhauser, A., Steiner, C., Harris, D. R., & Coffey, R. M. (1998). Comorbidity measures for use with administrative data. *Medical Care*, 36, 8-27.
- Goldman, L., Chu, P., Osmond, D., & Bindman, A. (2011). The accuracy of present-on-admission reporting in administrative data. *Health Services Research*, 46(Pt. 1), 1946-1962.
- Haas, J. (2000). *Report for the California hospital outcomes project community-acquired pneumonia, 1996; model development and validation*. Sacramento, CA: Office of Statewide Health Planning and Development.

- Hakim, R. B., Teno, J. M., Harrell, F. E., Jr., Knaus, W. A., Wenger, N., Phillips, R. S., . . . Lynn, J. (1996). Factors associated with do-not-resuscitate orders: Patients' preferences, prognoses, and physicians' judgments. SUPPORT Investigators. Study to understand prognoses and preferences for outcomes and risks of treatment. *Annals of Internal Medicine*, 125, 284-293.
- Healthcare Outcomes Center. (2006). *Community-acquired pneumonia: Hospital outcomes in California, 2002-2004*. Sacramento: California Office of Statewide Health Planning and Development.
- Hospital Outcomes Center. (2004). *Report on hospital outcomes for community-acquired pneumonia in California, 1999-2001*. Sacramento: California Office of Statewide Health Planning and Development.
- Iezzoni, L. I. (1997). Assessing quality using administrative data. *Annals of Internal Medicine*, 127(8 Pt. 2), 666-674.
- Iezzoni, L. I., Ash, A. S., Shwartz, M., Daley, J., Hughes, J. S., & Mackiernan, Y. D. (1996). Judging hospitals by severity-adjusted mortality rates: The influence of the severity-adjustment method. *American Journal of Public Health*, 86, 1379-1387.
- Lee, D. S., Stitt, A., Wang, X., Yu, J. S., Gurevich, Y., Kingsbury, K. J., . . . Tu, J. V. (2011). Administrative hospitalization database validation of cardiac procedure codes. *Medical Care*. Advance online publication. doi:10.1097/MLR.0b013e3182329778
- Marrie, T. J., Fine, M. J., Kapoor, W. N., Coley, C. M., Singer, D. E., & Obrosky, D. S. (2002). Community-acquired pneumonia and do not resuscitate orders. *Journal of the American Geriatric Society*, 50, 290-299.
- Meddings, J., Saint, S., & McMahon, L. F., Jr. (2010). Hospital-acquired catheter-associated urinary tract infection: Documentation and coding issues may reduce financial impact of Medicare's new payment policy. *Infection Control & Hospital Epidemiology*, 31, 627-633.
- Romano P, R. L., & Luft, H. (1996). *Report on heart attack outcomes in California 1994-1996, Volume 3: Detailed statistical results*. Sacramento: California Office of Statewide Health Planning and Development.
- Romano, P. S., Chan, B. K., Schembri, M. E., & Rainwater, J. A. (2002). Can administrative data be used to compare postoperative complication rates across hospitals? *Medical Care*, 40, 856-867.
- Romano, P. S., Yasmeen, S., Schembri, M. E., Keyzer, J. M., & Gilbert, W. M. (2005). Coding of perineal lacerations and other complications of obstetric care in hospital discharge data. *Obstetrics & Gynecology*, 106, 717-725.
- Shepardson, L. B., Gordon, H. S., Ibrahim, S. A., Harper, D. L., & Rosenthal, G. E. (1999). Racial variation in the use of do-not-resuscitate orders. *Journal of General Internal Medicine*, 14, 15-20.
- Shepardson, L. B., Youngner, S. J., Speroff, T., & Rosenthal, G. E. (1999). Increased risk of death in patients with do-not-resuscitate orders. *Medical Care*, 37, 727-737.
- Tabak, Y. P., Johannes, R. S., Silber, J. H., & Kurtz, S. G. (2005). Should do-not-resuscitate status be included as a mortality risk adjustor? The impact of DNR variations on performance reporting. *Medical Care*, 43, 658-666.

- Teno, J. M., Hakim, R. B., Knaus, W. A., Wenger, N. S., Phillips, R. S., Wu, A. W., . . . Lynn, J. (1995). Preferences for cardiopulmonary resuscitation: Physician-patient agreement and hospital resource use. The SUPPORT investigators. *Journal of General Internal Medicine*, 10, 179-186.
- Wenger, N. S., Pearson, M. L., Desmond, K. A., Harrison, E. R., Rubenstein, L. V., Rogers, W. H., & Kahn, K. L. (1995). Epidemiology of do-not-resuscitate orders disparity by age, diagnosis, gender, race, and functional impairment. *Archives of Internal Medicine*, 155, 2056-2062.
- Zingmond, D. S., & Wenger, N. S. (2005). Regional and institutional variation in the initiation of early do-not-resuscitate orders. *Archives of Internal Medicine*, 165, 1705-1712.
- Zrelak, P. A., Sadeghi, B., Utter, G. H., Baron, R., Tancredi, D. J., Geppert, J. J., & Romano, P. S. (2011). Positive predictive value of the agency for healthcare research and quality patient safety indicator for central line-related bloodstream infection ("selected infections due to medical care"). *Journal of Healthcare Quality*, 33(2), 29-36.