Thirty-Day Postoperative Mortality Risk Estimates and 1-Year Survival in Veterans Health Administration Surgery Patients

Tracy Smith, MS; Xinli Li, PhD; William Nylander, MD, MBA; William Gunnar, MD, JD

IMPORTANCE For more than 2 decades, the Veterans Health Administration (VHA) has relied on risk-adjusted, postoperative, 30-day mortality data as a measure of surgical quality of care. Recently, the use of 30-day mortality data has been criticized based on a theory that health care professionals manage patient care to meet the metric and that other outcome metrics are available.

OBJECTIVES To determine whether postoperative mortality data identify a delay in care to meet a 30-day mortality metric and to evaluate whether 30-day mortality risk score groups stratify survival patterns up to 365 days after surgery in surgical procedures assessed by the Veterans Affairs Surgical Quality Improvement Program (VASQIP).

DESIGN, SETTING, AND PARTICIPANTS Patients undergoing VASQIP-assessed surgical procedures within the VHA from October 1, 2011, to September 30, 2013, were evaluated. Data on 365-day survival follow-up of 212,733 surgical cases using VHA Vital Status and admission records were obtained with 10,947 mortality events. Data analysis was conducted from September 3, 2014, to November 9, 2015.

MAIN OUTCOMES AND MEASURES Survival up to 365 days after surgery for the overall cohort divided into 10 equal groups (deciles).

RESULTS There were 10,947 mortality events identified in a cohort of 212,733 surgical patients. The mean probability of death was 1.03% (95% CI, 1.01%-1.04%). Risk estimate groups in the 212,733 surgical cases analyzed showed significantly different postoperative survival, with consistency beyond the time frame for which they were developed. The lowest risk decile had the highest 365-day survival probability (99.74%; 95% CI, 99.66%-99.80%); the highest risk decile had the lowest 365-day survival probability (72.04%; 95% CI, 71.43%-72.64%). The 9 lowest risk deciles had linear survival curves from 0 to 365 postoperative days, with the highest risk decile having early survival risk and becoming more linear after the first 180 days. Survival curves between 25 and 35 days were consistent for all risk deciles and showed no evidence that mortality rates were affected in the immediate period beyond day 30. The setting of mortality varied by postoperative day ranges, with index hospitalization events declining and deaths outside of the hospital increasing up to 365 days.

CONCLUSIONS AND RELEVANCE Deciles of 30-day mortality estimates are associated with significantly different survival outcomes at 365 days even after removing patients who died within the first 30 postoperative days. No evidence of delays in patient care and treatment to meet a 30-day metric were identified. These findings reinforce the usefulness of 30-day mortality risk stratification as a surrogate for long-term outcomes.

Published online January 6, 2016.
The nexus between quality of care and outcomes as measured by mortality has been inferred for more than a century. In the 1850s, Florence Nightingale improved the care and treatment of wounded soldiers in the Crimean War through enhanced sanitation resulting in decreased mortality.1 Sixty years later, Codman promoted a hospital comparison of in-hospital mortality with that of the 4 best general hospitals in the United States.2 It has been almost 40 years since Donabedian advocated for the use of specific outcomes measures only when a valid association between quality of care and the outcome measured can be proved.3

Proving the association between survival as an outcome and quality of care has been a long coming. Between 1986 and 1992, the Health Care Financing Administration, the predecessor of the Center for Medicare & Medicaid Services, released annual in-hospital mortality statistics for deaths within 30 days of hospital admission. Despite discontinuation of this practice by the Health Care Financing Administration in the face of criticism regarding administrative data validity, studies4 showed an association between hospitals with higher-than-expected mortality rates and problems with quality of care. In 1990, New York began public reporting of in-hospital mortality rates for 28 hospitals for coronary artery bypass graft procedures, and time-series studies5 showed that hospitals with high mortality levels were able to improve quality through enhancements in structure and process.

For the past 2 decades, the Veterans Health Administration (VHA) has used risk-adjusted outcomes models to monitor and track 30-day mortality for cardiac and noncardiac surgical procedures.6,7 Any VHA surgical program with significantly higher-than-expected 30-day mortality underwent mortality case review and a selective site visit process by subject matter experts to evaluate the structure and process associated with surgical care delivery. Subsequent reports6 showed improvement in quality of care for individual VHA surgery programs with a level of concern and a decrease in 30-day mortality for all VHA surgical procedures.

Recently, an ethical argument has developed in the literature questioning the use of 30-day mortality as a measure of quality of care.9 The implication is that processes associated with intensive care management and end-of-life care are modified against the patient’s best interest to achieve a specific outcome: the delay in death beyond postoperative day 30. Reporting of mortality statistics beyond 30 days has been suggested9 as a possible solution. The intent of the present study was to compare 30-, 90-, 180-, and 365-day mortality statistics with preoperative patient risk and evaluate whether 30-day mortality is associated with delays in care.

Methods

Data Sources and Study Population
The Veterans Affairs Surgical Quality Improvement Program (VASQIP) collects data on surgical cases with measurable risk of mortality performed in VHA operating rooms according to Current Procedural Terminology code eligibility criteria.6,7,10 The VHA surgical programs performing VASQIP-eligible operations are required to submit the first 36 eligible cases in each 8-day sampling frame. The preoperative risk, intraoperative descriptive, and postoperative occurrence data are collected using national definitions by Surgical Quality Nurses.6,7,10 The National Surgery Office performs data completeness and validation checks on the data submitted for quarterly reporting of results and continual local and regional process improvement. Per VHA policy, Handbook 1058.05,11 the information presented in this article represents a VHA operations activity and therefore does not require informed consent or institutional review board approval. This analysis of VASQIP data was performed retrospectively.

30-Day Mortality Risk Prediction
The VASQIP data are used to annually derive logistic regression models predictive of 30-day, all-cause mortality using 3 years of combined data. Data are not deidentified. Missing values in the VASQIP database are limited; however, the IVEware Statistical Analysis System (SAS) macro is used to impute missing information.12 Potential predictors and confounders were identified using the purposeful selection macro.13 Multiple models were developed: a noncardiac overall model and specialty-specific models for general, vascular, orthopedic, urologic, thoracic (noncardiac), and cardiac (coronary artery bypass graft and valve/other) surgical specialties.

Survival Outcomes
Mortality outcomes in the VASQIP database were confirmed using the VHA Vital Status file, which combines Center for Medicare & Medicaid Services, Social Security Administration, and VHA internal utilization data to determine a best source, including flagging of records that indicate a death date followed by use of VA services.14,15 Records flagged as use after death were not considered deaths in this analysis. The VHA National Surgery Office regularly refreshes vital status information on VASQIP database patients for analysis of long-term outcomes. The VASQIP records were also matched to admission records in the VHA Corporate Data Warehouse to ascertain admission status to a VHA hospital or nursing home on the date of death.

Statistical Analysis
The VASQIP estimates for surgeries performed from October 1, 2011, to September 30, 2013, were restricted to the risk estimate for the patient’s first VASQIP-assessed surgery. This cohort was ranked and divided into 10 equal groups (deciles) for the overall population and separately for each of the surgical specialties. Deciles of risk defined the strata for survival analysis, with Kaplan-Meier plotting and log-rank tests of strata differences. Pairwise comparisons of strata were made in the Sidak multiple comparison adjustment (P < .05). These analyses were performed using survival from postoperative day 0 to 365 days and, in a sensitivity analysis, from postoperative day 30 to 365 days for patients who survived the first 30 postoperative days. Survival estimates for each decile were tested for linearity vs postoperative day using linear regression. For the overall cohort, linear regression splines with a node at 30 days were used to test for a difference in slope before and after postoperative
day 30. Kaplan-Meier plots, log-rank tests, and linear regressions were repeated for deciles within surgical specialties. All prediction and survival analyses were performed using SAS, version 9.3 (SAS Institute Inc); Kaplan-Meier curves were plotted using R, version 3.1.3 (R Foundation for Statistical Computing). Data analysis was conducted from September 3, 2014, to November 9, 2015.

Results

Risk Deciles and Survival Estimates
The VASQIP risk estimates were available for 236,125 surgical cases from October 1, 2011, to September 30, 2013, for all specialties combined. Keeping the first estimate record for each patient, the data set contained 212,733 records with 10,947 mortality events. The mean probability of death in the population was 1.03% (range, 0.005%-96.0%; 95% CI, 1.01%-1.04%). The lowest risk decile had a mean risk estimate of 0.02% (range, 0.005%-0.04%; 95% CI, 0.022%-0.023%). The fifth decile’s probability mean was 0.21% (range, 0.17%-0.24%; 95% CI, 0.205%-0.206%). The tenth risk decile had a mean probability of 7.18% (range, 1.82%-95.98%; 95% CI, 7.05%-7.31%).

The Kaplan-Meier plot for the overall population’s risk decile strata (Figure 1) demonstrates the separation of each of the deciles’ survival patterns at 365 postoperative days, which were confirmed by a significant log-rank test of strata equality (P < .001). The lowest risk decile had the highest 365-day survival probability (99.74%; 95% CI, 99.66%-99.80%); the highest risk decile had the lowest 365-day survival probability (72.04%; 95% CI, 71.43%-72.64%). Some multiple-comparison, Sidak-adjusted pairwise comparisons of deciles (1 vs 2, 2 vs 3, and 3 vs 4) were not significantly different (P = .95, P = .15, and P = .10, respectively); the rest of the comparisons of sequentially ordered strata were statistically significant (P < .01).

Risk deciles displayed linear slopes of the survival estimate vs days postoperative from 0 to 365 (each P < .001). Analysis of the residuals of the highest risk decile’s linear function indicated that a node placed at 180 postoperative days would be appropriate. Linear regression of the survival estimates of the highest risk decile into 2 periods from 0 to 180 postoperative days and 180 to 365 postoperative days showed that the 180-day survival estimates decreased at a rate of 0.098% per postoperative day and that the rate of decrease from postoperative day 181 to 365 was 0.041% (each P < .001).

The repeated specialty analysis of survival is displayed in the eFigure in the Supplement. The specialty sample sizes ranged from 4822 patients for plastic surgery to 68,913 for general surgery. Each specialty graph showed a distinct survival pattern for the highest risk decile with separation from all other deciles, and 365-day survival estimates varied by specialty.

When patients who died within 30 days of surgery were removed from the overall cohort and the survival curves were reevaluated, the survival curves showed similar separation (log-rank test, P < .001) and linear patterns with the exception of the highest risk decile, which showed increased mortality rates in the initial 30- to 180-day period compared with the 181- to 365-day period (Figure 2).

Investigation of Events Near 30 Postoperative Days
A visual investigation of the 5 days before and after the 30-day boundary revealed no excess mortality on postoperative
days 31 to 35 for any of the deciles (Figure 3). This finding was confirmed by linear spline analysis with a node at 30 days ($P > .05$).

**Setting of Mortality Events**

The overall cohort demonstrated declining proportions of deaths occurring during the index surgical hospitalization (Table), from 60.2% within the first 30 days of surgery to only 0.3% of deaths from postoperative day 181 to 365. Deaths occurring during a subsequent hospitalization after initial discharge or following an outpatient procedure represented 6.7% of the deaths within 30 days of surgery and 15.1% of the deaths overall. Nursing home and domiciliary admissions were the setting of 8.1% of the deaths within 30 days of surgery and represented 12.8% of all deaths within 365 days. The remaining patients who died did not have a matching VA admission record and are presumed to have died at home or in a non-VA facility. This category represented 25.0% of deaths occurring within the first 30 days of surgery and 57.4% of the deaths overall.

**Discussion**

Measuring and reporting surgical mortality outcomes at 2-day, 16,17 7-day, 18-20 30-day, 21 in-hospital, 22 in-hospital or 30-day, 23 and 60- and 90-day 24 intervals have been used to reflect the quality of care provided. The World Health Organization Safe Surgery Saves Lives initiative has adopted perioperative mortality as a quality-of-care estimate, 25 and a recently published Commission on Global Surgery report 26 in *The Lancet* settled on inpatient mortality statistics for tracking surgical outcomes. The results of our study support the use of a standardized point between 25 and 365 days to measure surgical outcomes. A linear association between preoperative risk assessment and postoperative days 25 and 365 was identified in all but the patients at highest risk. This association was reproducible in the following surgical specialties for VASQIP-assessed procedures: cardiac, general, neurosurgery, orthopedic, otolaryngology, plastic, thoracic (noncardiac), urology, and vascular. The survival estimate at 365 days varied by surgical specialty (eFigure in the Supplement).

The effect of the surgical procedure on survival has been shown 27 to last well beyond 30 days. Our findings demonstrate an association between risk deciles and survival beyond 30 days; removing the patients who died within 30 days supports this finding. Other studies 28 have shown that measuring mortality farther from the index surgical procedure increased the likelihood that the associated morbidity and death will occur after discharge from the initial hospitalization either at home, in a skilled nursing facility, or in the hospital following readmission. Our findings are consistent, identifying that 39.8% of 30-day mortality occurred after discharge from the index VA hospitalization or after outpatient surgery and increased to 82.3%, 93.6%, 98.8%, and 99.7% at days 60, 90, 180, and 365, respectively, as reported in the Table.

The VHA National Surgery Office relies on quarterly VASQIP risk assessment and 30-day mortality surgical outcomes data to identify VHA Surgical Programs with high mortality outlier statistics and a level of concern. 29 Any implication that 30-day mortality outcomes data are manipulated or gamed as a scheme to lower mortality outcomes statistics at the expense of patient care is unfounded and contradicted by our findings when examining the macro data for all participating VHA

![Figure 2. Survival Curves for Risk Deciles, Excluding Patient Mortalities Prior to Postoperative Day 30](image-url)
Surgery Programs. Specifically, the examination of surgical mortality between postoperative days 25 and 35 by risk decile showed no evidence of delayed mortality when the highest proportion of patients was hospitalized.

This study is strengthened by the examination of all VASQIP-assessed surgical procedures in the VHA, a large integrated health care system with a common electronic health record, manual data entry at each VHA surgery program, and verification of mortality events against a national registry.

The limitations of this study are that it was retrospective; examined an older, predominantly male population; and relied on national rather than facility-level data. The unethical behavior of any one health care professional or group of health care professionals at a single facility resulting in a delay in care and treatment to meet a 30-day mortality metric would not be discerned.

**Figure 3. Survival Curves for Risk Deciles, Postoperative Days 25-35**

<table>
<thead>
<tr>
<th>No. at risk</th>
<th>Decile 1</th>
<th>Decile 2</th>
<th>Decile 3</th>
<th>Decile 4</th>
<th>Decile 5</th>
<th>Decile 6</th>
<th>Decile 7</th>
<th>Decile 8</th>
<th>Decile 9</th>
<th>Decile 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postoperative Day</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
</tr>
<tr>
<td>Survival Estimate, %</td>
<td>92</td>
<td>93</td>
<td>94</td>
<td>95</td>
<td>96</td>
<td>97</td>
<td>98</td>
<td>99</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

| Overall Veterans Affairs Surgical Quality Improvement Program population. |

| Table. Setting of Mortality Events in the VASQIP Surgical Population |
|--------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| No. Mortality Events | No. Patients at Risk | Index VA Hospitalization | Subsequent VA Hospitalization | Nursing Home, Domiciliary | Other |
|---------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| POD | | | | | | | | | | |
| 1-30 | 2136 | 212733 | 1285 (60.2) | 143 (6.7) | 174 (8.1) | 534 (25.0) |
| 31-60 | 1231 | 210597 | 218 (17.7) | 236 (19.2) | 190 (15.4) | 587 (47.7) |
| 61-90 | 1006 | 209366 | 64 (6.8) | 181 (18.0) | 149 (14.8) | 612 (68.0) |
| 91-180 | 2363 | 208360 | 28 (1.2) | 427 (18.1) | 319 (13.5) | 1589 (67.2) |
| 181-365 | 4211 | 205997 | 12 (0.3) | 668 (15.9) | 565 (13.4) | 2966 (70.4) |
| Overall | 10947 | 212733 | 1607 (14.7) | 1655 (15.1) | 1397 (12.8) | 6288 (57.4) |

Abbreviations: POD, postoperative day; VA, Veterans Affairs; VASQIP, Veterans Affairs Surgical Quality Improvement Program.

*Readmission.

*bIncludes home, non-VA facility.
Conclusions

This study found a linear association between preoperative risk and postoperative mortality that extends beyond postoperative day 30 in all but the highest risk decile as measured at 90, 180, and 365 days for VASQIP-assessed surgical procedures. Veterans in the highest risk decile experienced a greater mortality risk within the first 180 days compared with days 181 to 365. The association between risk and survival was identified in all surgical specialties and in veterans who survived beyond postoperative day 30. We found no evidence that veteran care and treatment was compromised to meet a 30-day mortality metric. The findings of this study support the VHA's continued use of 30-day mortality as the basis for monitoring and tracking surgical outcomes and the quality of surgical care.