Published in final edited form as:

Ann Surg. 2021 October 01; 274(4): 637–645. doi:10.1097/SLA.0000000000005068.

The Correlation Between Case Total Work Relative Value Unit, Operative Stress and Patient Frailty: Retrospective Cohort Study

Katherine M Reitz, MD MSc[1], **Patrick R Varley, MD MSc**[1], **Nathan L Liang, MD MSc**[1,2], **Ada Youk, PhD**[3,4], **Elizabeth L George, MD MSc**[5], **Myrick C Shinall Jr, MD PhD**[6], **Paula K Shireman, MD MS MBA**[7,8,9], **Shipra Arya, MD SM**[5], **Edith Tzeng, MD**[1,2,10], **Daniel E Hall, MD MDiv MHSc**[1,2,3,11]

[1]Department of Surgery, University of Pittsburgh, Pittsburgh, PA

[2]Division of Vascular Surgery, University of Pittsburgh, Pittsburgh, PA

[3]Center for Health Equity Research and Promotion, Veterans Affairs Pittsburgh Healthcare System, Pittsburgh, PA

[4]Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA

[5]Division of Vascular Surgery, Stanford University School of Medicine, Stanford, CA

[6]Department of Surgery, Vanderbilt University Medical Center, Nashville, TN

[7]Department of Surgery, University of Texas Health San Antonio, San Antonio, TX

[8]Department of Surgery, South Texas Veterans Health Care System, San Antonio, TX

Current Address and affiliation of corresponding and reprint request author Daniel Hall, MD MDiv MHSc, 200 Lothrop Street, UPMC Presbyterian, Pittsburgh, PA 15213, hallde@upmc.edu.

Katherine M Reitz MD MSc – Key contributions to concept and design with Dr. Hall, data acquisition, statistical analysis and interpretation of data; drafted manuscript to completion; generated and reviewed the final documents.

Patrick R Varley MD MSc -- Key contributions to concept and design, data acquisition – especially Center for Medicare and Medicaid Data, and invaluable to interpretation of data; critically evaluated the order and presentation of data; reviewed the final documents.

Nathan L Liang MD MSc -- Key contributions to data analysis including the creation of the conceptualization and R code for Figure 4; critically evaluated discussed the message generated throughout the manuscript writing process; reviewed the final documents.

Ada Youk PhD -- Key contributions to concept and design, data acquisition, and statistical analysis plan; critically evaluated the presentation of data especially in regards to data and figures; reviewed the final documents.

Elizabeth L George MD MSc -- Key contributions to data acquisition and provided stata code for key conceptual analysis; flow of content within the manuscript; reviewed the final documents.

Myrick C Shinall Jr MD PhD -- Key contributions to acquisition of data including the Operative Stress Score and interpretation of data including but not limited to the box and whisker plots; critically evaluated the presentation of data and conceptualization of data presentation throughout results; reviewed the final documents.

Paula K Shireman MD MS MBA -- Key contributions to interpretation of data and underlying meaning of our findings; edited and evaluated manuscript to completion including adding multiple conceptual features to the discission and important references; reviewed the final documents.

Shipra Arya MD SM -- Key contributions to data acquisition including the Operative Stress Score and Risk Analysis Index; critically evaluated the presentation and encouraged conceptual depth of understanding throughout the results and discussion; reviewed the final documents.

Edith Tzeng MD – Key contributions to concept and design as well as presentation of data and conceptual framework for results; multiple rounds of review of manuscript to completion to help refine the message of the content; reviewed the final documents.

Daniel E Hall MD MDiv MHSc -- Key contributions to primary concept and design with Dr. Reitz, data acquisition – all sources used, and interpretation of data; preliminary discussion and multiple rounds of edits for the drafted manuscript to completion; reviewed the final documents.

No authors report disclosures, conflict of interest or relevant financial interests related to the content of the manuscript. The opinions expressed here are those of the authors and do not necessarily reflect the position of the Department of Veterans Affairs or the US government.

[9]University Health System, San Antonio, TX

[10]Department of Vascular Surgery, Veterans Affairs Pittsburgh Healthcare System, Pittsburgh, PA

[11]Wolff Center, UPMC, Pittsburgh, PA

Abstract

Objective: Assess the relationships between case total work Relative Value Units (wRVU), patient frailty and the physiologic stress of surgical interventions.

Summary Background Data: Surgeon reimbursement is frequently apportioned by wRVU. These subjective, procedure-specific valuations generated by physician survey estimate the intensity and time for typical patient care services. We hypothesized wRVU would not adequately account for patient-specific factors, such as frailty, that modify the required physician work, regardless of procedural complexity.

Methods: Using National and Veterans Affairs Surgical Quality Improvement Programs (2015–2018), we evaluated the correlation between case total wRVU, patient frailty [Risk Analysis Index (RAI)] and physiologic surgical stress [Operative Stress Score (OSS)].

Results: Of 4,111,371 (86%) cases, the correlation between total wRVU and operative stress was moderate [ρ_s=0.587 (95%CI, 0.586–0.587)], but negligible with frailty [ρ=0.177 (95%CI, 0.176–0.178)]. Very high operative stress procedures [n=34,047 (1%)] generated a mean total wRVU of 55.1 (SD, 12.9), comprising 7%, 2%, and 1% of thoracic, vascular, and general surgical cases, respectively. Very frail patients [n=152,535 (4%)] accounted for 9% of thoracic, 9% of vascular, 4% of general, 5% of urologic, and 4% of neurologic surgical cases, generating 21.0 (SD, 12.4) mean total wRVU. Some non-frail patients undergoing low operative stress procedures [n=60,128 (2%)] nonetheless generated the highest quintile wRVU; these comprised >15% of plastic, gynecologic, and urologic surgical cases.

Conclusions: Surgeon reimbursement correlates with operative stress but not patient frailty. The total wRVU does not adequately reflect patient-specific factors that increase physician workload required to render optimal care to complex patients.

Mini-Abstract:

Of 4,111,371 cases, the correlation between total work Revenue Value Unit and the physiologic stress of a surgical intervention was moderate, but negligible with patient frailty. The work Revenue Value Unit does not reflect patient-specific factors that increase the physician workload required to optimally care for complex patients.

Keywords

Frailty; operative stress; financial reimbursement; relative value scales; reimbursement; physician services

INTRODUCTION

In 1988, a Harvard University multidisciplinary team systematically investigated factors contributing to physician work and developed a Resource-Based Relative Value Scale for

services, overhauling of physician reimbursement.[1] The factors, assessed through national surveys, encompassed “time spent before, during, and after the service and the intensity with which that time is spent,” in addition to “the opportunity cost of training, which represents income forgone.”[1] The results quantified a numerical work Relative Value Unit (wRVU) for each *Current Procedural Terminology* (CPT) code.

Centers for Medicare and Medicaid Services (CMS) began using the CPT-specific wRVU to reimburse physicians in 1992, with annual updates based upon recommendations from the Relative Value Scale Unit Committee (RUC). Mirroring the original process, updated recommendations are guided by a minimum of 30 surveyed physicians estimating the time and intensity of work required for patient care based on clinical vignettes, describing a typical patient.[2,3] Critics have demonstrated that this subjective methodology results in an association between wRVU and operative time, but not the overall time required for patient care.[4–6]

Physician work depends on the time required to perform procedures, complexity of each patient, and patients’ physiologic response to the stress of the surgical intervention.[7,8] Frail patients experience higher rates of postoperative complications, length of stay, failure to rescue, and mortality, especially after higher stress procedures that exhaust their limited physiologic reserve.[7,9–16] Optimal perioperative care for such medically complex patients requires increased physician work; however, the wRVU quantifying surgeon effort lack consideration for the complexity of individual patients. For instance, the wRVU assigned to a laparoscopic appendectomy performed on an otherwise healthy 20-year-old is identical to that assigned to an 80-year-old with multiple comorbidities. To the extent that operative stress typically increases with operative time and intensity, we hypothesize that wRVU will correlate with operative stress, but not correlate with patient frailty, indicating a potentially unappreciated aspect of physician time and effort that could be used to adjust compensation schemes to better reflect actual physician work.

METHODS

We used data from the American College of Surgeons (ACS) National Surgical Quality Improvement Program (NSQIP) and Veterans Affairs Surgical Quality Improvement Program (VASQIP) to maximize the external validity of findings in this retrospective cohort study. These datasets provide preoperative patient demographic and risk factors; operative procedures and surgical specialties; and 30-day postoperative outcomes.[17,18] We also used the publicly available Medicare Physician Fee Schedule Data which provides a wRVU per CPT code, defines the summation of wRVU if multiple CPT codes are assigned per case, and identifies CPT codes subject to a global billing period.[19]

This study was reviewed by the VA Pittsburgh Healthcare System institutional review board which determined this deidentified, retrospective secondary data analysis to be exempt (PRO1385). All data analysis (February 26, 2020 to March 1, 2021) was completed with Stata 15.1 (StataCorp), Prism 7.0 (GraphPad), or Python (Python Software Foundation; Pandas and Seaborn packages).[20,21] All reporting was consistent with the STROBE statement.[22]

Patient population and measures

We included all cases between January 1, 2015 and December 31, 2018. We assessed patient frailty with the Risk Analysis Index (RAI), a validated pseudo-continuous (ranging 11–60) measure calculated retrospectively from variables available in these datasets including age, gender, comorbid conditions, and functional status (i.e., performance of activities for daily life) predicting postoperative complications, failure to rescue, short and long term mortality.[9,16,23] We determined the physiologic stress of the surgical intervention using the Operative Stress Score (OSS), a validated tool that classifies 565 common CPT codes according to 5 categories of physiological stress (1, very low; 5, very high stress). The OSS was developed using modified Delphi consensus methodology whereby a panel of surgical specialists and anesthesiologists iteratively rated the stress of each CPT code.[7,8] Increasing OSS scores both alone and synergistically with RAI were associated with increasing postoperative complications, failure to rescue from complications, and mortality across a broad range of non-cardiac surgery interventions, validated in both VASQIP as well as NSQIP.[7,8,24] The wRVU associated with each CPT code was identified from the Medicare Physician Fee Schedule Data.[19]

Cases in the datasets include a principal procedure (i.e., CPT code) and up to 10 other procedures completed by the primary surgeon. For cases with multiple procedures (i.e., multiple CPT codes), operative stress was defined by the principal procedure OSS. The case total wRVU was calculated by summing each CPT code(s) wRVU according to the CMS Physician Fee Schedule multiple procedure modifier (Supplemental Table 1).[19]

We included cases completed by general, orthopedic, vascular, gynecologic, urologic, neurologic, otolaryngologic, plastic, or thoracic specialties. Records missing principal procedure CPT codes, OSS, or any variables for RAI calculation RAI were excluded. All analyses were completed on an individual case level. Therefore, these data represent a sample of major surgery, which in both improvement projects deliberately under-sample common, high-volume procedures to comprehensively represent procedures performed nationally in the VA and private sector.[17,18]

Statistical Analysis

To understand patient characteristics by the level of attributed physician work, we categorized the total wRVU into quintiles and compared baseline demographics, preoperative risk factors, procedural data, and 30-day mortality. Continuous variables were expressed as means [standard deviation (SD)] or medians [interquartile range (range)] and categorical variables as frequencies (percent).

We examined the correlation [95% confidence intervals (CI)] between total wRVU as a continuous variable and both OSS and RAI.[23] These relationships were quantified by Pearson's (ρ) or Spearman's Rank Correlation (ρ_s) for ordinal (OSS) or continuous (RAI) variables and visually depicted with box and whisker plots.[25]

We depicted the relationship between total wRVU, operative stress, and patient frailty simultaneously with a single figure demonstrating both the total wRVU quintile and case frequency within each possible OSS and RAI category. Highlighting the relationship

between total wRVU, OSS and RAI, we separately quantified the total wRVU and frequency of surgical specialties caring for (a) patients undergoing very high operative stress procedures (OSS5), (b) very frail patients (RAI≥40), and (c) robust or normal patients (RAI≤29) undergoing very low or low operative stress procedures (OSS≤2) which generated a very high total wRVU (top quintile).

Sensitivity Analysis

We evaluated the robustness of our results with sensitivity analyses. First, we assessed the correlation between total wRVU and alterative measures of operative stress including case duration, number of procedures per case, as well as logarithmically transformed hospital length of stay (LOS), increases in which are associated with frailty.[26] Second, we evaluated the correlation between total wRVU, OSS and RAI among cases with only one CPT code per case and an established global period in which the wRVU for procedures following the index case cannot be accrued (Supplemental Table 2).[19] Third, we examined the relationship between wRVU, OSS and RAI at the CPT code level. For each principal procedure CPT code with an OSS, we computed the mean RAI and assessed the correlation between the principal procedure wRVU, OSS and summarized RAI. Finally, to understand differences in billing practices when total wRVU is directly (i.e., NSQIP) or indirectly (i.e., VASQIP) representative of physician reimbursement in the VA and private sectors, we evaluated the relationship between total wRVU, OSS and RAI differed between the datasets. We estimated the difference in total wRVU per dataset, controlling for OSS and RAI using a multivariable linear regression with a robust variance estimator. We then estimated the total wRVU, stratified by OSS and RAI categories.

RESULTS

We identified 4,821,713 case-level records in the datasets (NSQIP, n=3,935,119; VASQIP, n=886,594); 710,342 (15%) were excluded, resulting in 4,111,371 cases (Figure 1). The mean age was 58 years (SD, 16), 2,178,215 (53%) were male, with a mean RAI of 24.9 (SD, 8.3), median OSS of 3 (range, 2–3), and mean total wRVU of 18.4 (SD, 10.0). 1,293,931 (31%) cases had multiple procedures, of which 992,738 (77%) had multiple procedures contributing to the primary team's total wRVU [mean 25.8 (SD, 12.7)], and accounted for greater than 40% of thoracic [n=24,678 (42%)], gynecologic [n=81,287 (41%)], and plastic [n=45,533 (53%)] surgical cases (Table 1).

Total wRVU per case were divided into quintiles: very low (<9.9), low (10.0–15.0), moderate (15.1–20.6), high (20.7–25.0), and very high (>25.0). The very low quintile was younger [mean, 52 years (SD, 18) vs 61 (SD, 14)]; more likely male (65% vs 56%), Hispanic (11% vs 7%), and robust (43% vs 21%); and with a lower 30-day postoperative mortality (0.3% vs 1.6%) when compared to the high wRVU quintiles (Table 2).

Overall, the correlation between total wRVU (as a continuous variable) and operative stress was moderately positive [ρ_S=0.587 (95%CI, 0.586–0.587)](Figure 2A). The correlation was high among urologic [ρ_S=0.833 (95%CI, 0.832–0.834)]; moderate among general [ρ_S=0.694 (95%CI, 0.694–0.695)], vascular [ρ_S=0.657 (95%CI, 0.655–0.659)], neurologic [ρ_S=0.591 (95%CI, 0.588–0.594)], otolaryngologic [ρ_S=0.609 (95%CI, 0.606–0.613)] and thoracic

[ρ_s=0.677 (95%CI, 0.673–0.681)]; low among orthopedic [ρ_s=0.375 (95%CI, 0.655–0.659)] and plastic [ρ_s=0.460 (95%CI, 0.454–0.465)]; and negligible among gynecologic [ρ_s=0.169 (95%CI, 0.165–0.174)] surgical cases (Figure 2B).

Overall, the correlation between total wRVU and patient frailty was negligible [ρ=0.177 (95%CI, 0.176–0.178)](Figure 3A). The correlations were also negligible (correlation range, ρ=−0.148 to 0.238) for each surgical specialty, except otolaryngology, constituting a low positive correlation [ρ=0.420 (95%CI, 0.415–0.426)] and plastic surgery [ρ=−0.001 (95%CI, −0.007–0.006)] for which the correlation was not significant (Figure 3B).

Figure 4A represents the 4-dimensional relationship between patient frailty (Y-axis), operative stress (X-axis), wRVU quintile (color gradient), and case frequency (circle area). 2,739,152 (67%) cases were performed on non-frail patients (RAI≤29) undergoing low to moderate stress procedures (OSS2–3). Procedures with very high operative stress (OSS5) accounted for 34,047 (1%) cases, generating a mean total wRVU of 55.1 (SD, 12.9); these very high operative stress procedures accounted for 7% (n=4,133/59,110) of thoracic, 2% (n=5,744/279,115) of vascular, and 1% (n=24,151/1,803,368) of general surgical cases, with frequencies <0.1% for all others (Figure 4B; Supplemental Tables 3–4). Very frail patients (RAI≥40) accounted for 152,535 (4%) cases, generating a mean total wRVU of 21.0 (SD, 12.4). The case frequency among very frail patients did not vary across wRVU quintiles and encompassed 9% (n=5,302/59,110) of thoracic, 9% (n=26,472/279,115) of vascular, and 3–5% of general (n=62,762/1,803,368), urologic (n=14,747/289,841), and neurologic (n=7,957/229,632) surgical cases (Figure 4C; Supplemental Table 3). 1,584,388 (39%) cases were performed on non-frail patients (RAI≤29) with a very low or low operative stress (OSS1–2); nonetheless, 4% generated very high total wRVU (n=60,128). These very high total wRVU but low-stress and low-frailty cases were predominantly plastic, gynecologic, and urologic surgical cases, accounting for 10% (n=8,897/86,338), 3% (n=6,050/200,620), 4% (n=12,983/289,841) of each specialty's volume (Figure 4D; Supplemental Tables 3 and 5). The most common principal procedures performed on non-frail patients undergoing low stress procedures yielding a very high wRVU are available in Table 3.

Sensitivity Analysis

Alternative measures of operative stress, including case duration [median, 90 minutes (range, 54–145)], procedures per case [median, 1 procedure (range, 1–2)], and postoperative hospital LOS [median, 1 day (range, 0–2)] had a moderate-to-low positive correlation [ρ=0.608 (95%CI, 0.607–0.609), ρ_s=0.376 (95%CI, 0.375–0.377), and ρ=0.245 (95%CI,0.244–0.246)] with total wRVU.

Among cases with a single procedure [N=3,118,633 (76%)] and cases including procedures with an associated global period [n=3,941,642 (96%)], the OSS demonstrated a moderate [single procedure, ρ_s=0.561 (95%CI, 0.560–0.562); global period, ρ_s=0.657 (95%CI, 0.566–0.567)] and RAI a negligible [single procedure, ρ=0.220 (95%CI, 0.560–0.562); global period, ρ=0.207 (95%CI, 0.206–0.207)] correlation with total wRVU (Supplemental Figure 1–2).

A total of 562 principal procedure CPT codes had a OSS rating median 2 (IQR, 2–3)] with an overall mean RAI of 26 (SD, 5). At the CPT code level, again OSS demonstrated a moderate [ρ_s=0.502 (95%CI, 0.438–0.562)] and RAI a negligible [ρ=0.278 (95%CI, 0.200–0.353)] correlation with wRVU.

NSQIP cases (n=3,366,417) had a median OSS of 2 (range, 2–3), patients had a mean RAI of 23.1 (SD, 7.9), and generated a mean total wRVU of 18.4 (SD, 10.2). VASQIP cases (n=744,954) had a median OSS of 3 (range, 2–3), patients had a mean RAI of 31.2 (SD, 6.9), and generated a mean total wRVU of 18.4 (SD, 8.8). In both datasets, the correlation between total wRVU and OSS was moderate [NSQIP, ρ_s=0.568 (95%CI, 0.567–0.569); VASQIP, ρ_s=0.574 (95%CI, 0.573–0.576)] and RAI was negligible [NSQIP, ρ=0.214 (95%CI, 0.213–0.215); VASQIP, ρ=0.08 (95%CI, 0.025–0.030)]. When controlling for OSS and RAI, NSQIP generated 1.38 (95%CI, 1.36–1.41) greater total wRVU compared to VASQIP (Supplemental Table 6). This pattern continued when total predicted wRVU was stratified by OSS and RAI categories (Supplemental Figure 3).

DISCUSSION

In this retrospective evaluation of four million cases from both the ACS and VA Surgical Quality Improvement Programs, total wRVU positively correlated with the physiologic stress of the surgical intervention but negligibly correlated with preoperative patient frailty. These findings suggest that total wRVU accounts for procedure-related complexity and stress, but not the variations in patient-related complexity that require additional physician effort to provide optimal patient care. Further, some low-stress procedures performed in low-frailty populations garnered high total wRVU, suggesting an opportunity to adjust total wRVU for high-complexity patients to account for the increased physician work required to care for these patients based on patient-specific factors, such as frailty.

Objective measures of intraoperative and postoperative physician work captures up to 80% of the overall variability in procedure based wRVU.[4] In the subjective and heavily scrutinized wRVU quantifications of perioperative physician work, intraoperative time explained the vast majority of wRVU variability and, as we have demonstrated here, also positively correlated with the physiologic stress of a surgical intervention.[5,6] Across surgical specialties, CPT-specific wRVU valuations were determined by physician surveys estimating the time and intensity required for procedures based on a clinical vignette describing a typical patient. In 2006, the Society of Thoracic Surgeons, RUC, and CMS began to incorporate objective measures of physician work into the wRVU valuation for cardiothoracic procedures.[27] However, the patterns we observed extended across all surgical specialties, including thoracic surgery. Therefore, available evidence suggests that the wRVU adequately accounts for the intraoperative work of specific surgical procedures, but not the work required for all perioperative care, especially for frail patients.

The 2006 Society of Thoracic Surgeons CPT code revaluation aimed to reduce error in both over- and underestimates in physician work and accurately reflect care for typical cardiothoracic patients who are aging with increasing comorbidity.[27] However, frailty is more complex than simply age and comorbidities. Frailty is a global syndrome of

decreased physiological reserve associated with but independent of age, driven by a complex accumulation of medical, functional, and social deficits.[11,28] Frail patients have a 10-fold increased risk of postoperative complications, increased rates of return to the operating room, longer hospitalizations, and a 50% increased rate of unplanned readmissions.[7,11–15] These outcomes directly and objectively translate into increased physician work, applying to frail and medically complex patients undergoing both low and high operative stress procedures.[7,11] In addition, recent work suggests identifying, assessing, and prehabilitating frail patients may improve their preoperative physiologic reserve, response to the stress of surgery, and postoperative recovery.[29–32] These efforts may improve patient outcomes, but also increase the physician work required to optimize complex patients. However, physician work quantified either objectively using the Society of Thoracic Surgeons National Database or subjectively with physician surveys translates into reimbursement for care provided to a typical patient. For example, assigned physician reimbursement for a laparoscopic appendectomy and additional any services provided in the global period (i.e., 1 day prior to and 90 days following the procedure) is 9.45 wRVU.[19] The optimal perioperative care for robust and frail patients undergoing a laparoscopic appendectomy are not equivalent in terms of physician work, yet they are defined by the same CPT code.[7,10,33]

In our analysis, thoracic, vascular, and general surgeons cared for the highest proportion of frail patients undergoing the highest level of operative stress, resulting in a synergistic effect on the risk for postoperative morbidity and mortality.[7] On the opposite end of the spectrum, the highest mean total wRVU per case were completed by thoracic, neurologic, and plastic surgeons. Plastic surgical interventions generated very high total wRVU for low stress procedures performed on robust patients. Our data notably differ from Childers et. al., who described a higher wRVU for neurosurgery and lower wRVU for plastic surgery when compared to general surgical interventions.[4] However, their analysis was completed at the level of the CPT code and not the case, limiting their analysis only to cases with a single principal procedure and excluding nearly 30% of cases for which multiple procedures were completed. Such an approach can yield meaningful findings;[4] yet it does not adequately represent the complexity of real-world practice or the variability between surgical specialties. By contrast, our approach follows the CMS Fee Schedule to assign a total case wRVU representing all procedures performed in that case, demonstrating that (a) multiple procedure cases generated an average of 26 wRVU compared to the 16 wRVU generated by single procedure cases, and (b) the proportions of single procedure cases vary according to surgical specialty. These data demonstrate that cases are undervalued if the wRVU generated by multiple procedures is ignored. They also suggest that surgeons who care for complex, frail patients are potentially undercompensated for the increased work required to optimize patient outcomes.

The quality improvement program datasets do not capture procedure code modifiers indicating when services provided are significantly greater (i.e., modifier-22) or less (i.e., modifier-52) than usual, and thus quantification of wRVU may be incomplete. Some surgeons may manage all aspects of perioperative care for a complex patient, resulting in a wide array of services in the global period; and thus, may be undercompensated for their work. Others may interpret the global period stringently, generating additional wRVU not captured in our analysis, including the application modifier-22 as well as evaluation and

management of conditions not typically associated with the surgical procedure (e.g., labile blood pressure after, but not directly related to, an appendectomy). Finally, other surgeons may mitigate the increased work of caring for frail patients by referring portions of care to other specialists (e.g., preoperative nutrition optimization). Future research may elucidate prevalent patterns, but our data suggest the current scheme of procedure-specific wRVU valuation does not have the capability to measure (or the flexibility to account for) individual patient-level factors, as evidenced by the negligible correlation between RAI score and total wRVU. Therefore, current wRVU attribution does not incentivize surgeons to perform additional work required to both prevent complications and optimize outcomes. Our finding are consistent with Medicare data showing that adding a frailty-based index to the standard Hierarchal Condition Category improved cost predictions,[34] advocating for Medicare risk adjustment to add such patient-level factors to address inequities in value-based payment programs.[35] We therefore suggest additional work focuses on objectively evaluating and incorporating graduated adjustment of the current wRVU structure to up-scale the care of frail and down-scale the care of robust patients.

Finally, exploratory analyses demonstrate potentially informative differences in total wRVU between cases in the VASQIP and NSQIP dataset. In each dataset, the CPT codes are assigned by billing coders after case completion and before data abstraction.[17,18] Accurate coding is incentivized in both sectors, yet much of VA care is rendered outside of the typical fee-for-service model. As such, incentives to maximize total wRVU may differ in each sector. Our exploratory analysis suggests that private-sector cases generated higher total wRVU than those performed in the VA after controlling for patient frailty and operative stress. It remains unclear if these findings represent an overestimation of work in NSQIP or an underestimation of work in VASQIP. If confirmed and better understood, this discrepancy may have significant policy implications for the VA which, pursuant to the MISSION act, is obligated to pay for an increasing amount of private-sector care that might otherwise be rendered within the VA.[36,37] These differences may be especially important given the characterization of the VA as a national safety net system due to caring for veterans with a higher proportion of catastrophic disabilities, mental illness and low incomes compared to the general population.[38]

Our provocative findings have limitations. First, case sampling for the quality improvement programs both overrepresent high frequency and underrepresent low frequency interventions and our analysis was limited to CPT codes with a corresponding OSS, curated in a Veteran patient sample.[17] Therefore, conclusions drawn based upon the procedure frequency, especially those more common in private than VA hospitals (i.e., gynecologic operations), must be interpreted with caution. Second, these data were abstracted by trained personnel retrospectively which may result in missingness not at random, potentially creating unmeasured confounding. Third, we included both the ACS and VA datasets to increase the external validity of our findings; however, exploratory comparative observations may not adequately control for patient-, provider-, and system-level differences. Fourth, CPT coding strategies vary regionally and this may account for variability in total wRVU not otherwise accounted for.[39]

We therefore conclude that physician work, as quantified by the wRVU, accounts for the complexity and stress of the surgical intervention but not patient complexity, a major driver of physician work intensity and time. Further studies are required to confirm and explore the specific patient factors driving the intensity and time of physician work. Upon identification, weighting or adjusting the wRVU may not only improve the accuracy of attributed physician work but also incentivize the identification and optimization of preoperative risk factors, improving patient outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding support:

This research was supported in part by the US Department of Veterans Affairs, the Veterans Health Administration, the Office of Research and Development; grants I21 HX-002345 and XVA 72-909 (Hall); grant 5R03AG050930 (Arya), grant K76AG068436 (Shinall), and L30 AG064730 (Reitz) from the National Institute on Aging, NIH; grant K12CA090625 (Shinall) from the National Cancer Institute, NIH; grant U01 TR002393 (Hall and Shireman) from the National Center for Advancing Translational Sciences and Office of the Director, NIH; grant 5T32HL0098036 (Reitz) from the National Heart, Lung, and Blood Institute, NIH. These funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit for publication.

REFERENCES

1. Hsiao WC, Braun P, Dunn D, et al.Resource-Based Relative Values: An Overview. JAMA. 1988;260:2347–2353. [PubMed: 3050169]
2. AMA/Specialty Society. RVS Update Process. Am Med Assoc.
3. American Medical Association. Understanding the RUC Survey Instrument: Physician ServicesAvailable from: https://www.youtube.com/watch?v=nu5unDX8VIs&feature=youtu.be.2015. Accessed May 5, 2020.
4. Childers CP, Dworsky JQ, Russell MM, et al.Supplement - Association of Work Measures and Specialty with Assigned Work Relative Value Units among Surgeons. JAMA Surg. 2019;154:915–921. [PubMed: 31314063]
5. Childers C, Maggard-Gibbons M. Assessment of the Contribution of the Work Relative Value Unit Scale to Differences in Physician Compensation Across Medical and Surgical Specialities. JAMA Surg. Epub ahead of print2020. DOI: 10.1001/jamasurg.2020.0422.
6. Laugesen MJ, Wada R, Chen EM. In setting doctors' medicare fees, CMS almost always accepts the relative value update panel's advice on work values. Health Aff. 2012;31:965–972.
7. Shinall MC, Arya S, Youk A, et al.Association of Preoperative Patient Frailty and Operative Stress with Postoperative Mortality. JAMA Surg. 2019;15213:1–9.
8. Shinall MC, Youk A, Massarweh NN, et al.Association of Preoperative Frailty and Operative Stress With Mortality After Elective vs Emergency Surgery. JAMA Netw Open. 2020;3:10–13.
9. Hall DE, Arya S, Schmid KK, et al.Development and initial validation of the Risk Analysis Index for measuring frailty in surgical populations. JAMA Surg. 2017;152:175–182. [PubMed: 27893030]
10. McIsaac DI, Bryson GL, Van Walraven C. Association of frailty and 1-year postoperative mortality following major elective noncardiac surgery: A population-based cohort study. JAMA Surg. 2016;151:538–545. [PubMed: 26791334]
11. Shah R, Attwood K, Arya S, et al.Association of frailty with failure to rescue after low-risk and high-risk inpatient surgery. JAMA Surg;153. Epub ahead of print2018. DOI: 10.1001/jamasurg.2018.0214.
12. Wahl TS, Graham LA, Hawn MT, et al.Association of the modified frailty index with 30-day surgical readmission. JAMA Surg. 2017;152:749–757. [PubMed: 28467535]

13. Rothenberg KA, Stern JR, George EL, et al.Association of Frailty and Postoperative Complications With Unplanned Readmissions After Elective Outpatient Surgery. JAMA Netw open. 2019;2:e194330. [PubMed: 31125103]
14. Davenport DL, Henderson WG, Khuri SF, et al.Preoperative risk factors and surgical complexity are more predictive of costs than postoperative complications: A case study using the National Surgical Quality Improvement Program (NSQIP) database. Ann Surg. 2005;242:463–471. [PubMed: 16192806]
15. Fagenson AM, Powers BD, Zorbas KA, et al.Frailty Predicts Morbidity and Mortality After Laparoscopic Cholecystectomy for Acute Cholecystitis: An ACS-NSQIP Cohort Analysis. J Gastrointest Surg.. Epub ahead of print2020. DOI: 10.1007/s11605-020-04570-1.
16. Varley PR, Borrebach J, Arya S, et al.Clinical Utility of the Risk Analysis Index as a Prospective Frailty Screening Tool within a Multi-practice, Multi-hospital Integrated Healthcare System. Ann Surg.
17. Massarweh NN, Kaji AH, Itani KMF. Practical guide to surgical data sets: Veterans Affairs Surgical Quality Improvement Program (VASQIP). JAMA Surg. 2018;153:768–769. [PubMed: 29617543]
18. Shiloach M, Frencher SK, Steeger JE, et al.Toward Robust Information: Data Quality and Inter-Rater Reliability in the American College of Surgeons National Surgical Quality Improvement Program. J Am Coll Surg2010;210:6–16. [PubMed: 20123325]
19. Medicare Physician Fee Schedule. National Physician Fee Schedule Relative Value FileAvailable from: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeeSched/PFS-Relative-Value-Files.2018.
20. Waskom M, Botvinnik O, Gelbart M, et al.mwaskom/seaborn: v0.11.0. ZenodoAvailable from: 10.5281/zenodo.592845.2020. Accessed August 12, 2020.
21. Reback J, McKinney W, Van den Bossche J, et al.Pandas 1.2.0rc0. ZenodoAvailable from: 10.5281/zenodo.3509134.Accessed August 12, 2020.
22. von Elm E, Altman DG, Egger M, et al.The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. Int J Surg. 2014;12:1495–1499. [PubMed: 25046131]
23. Arya S, Varley P, Youk A, et al.Recalibration and External Validation of the Risk Analysis Index. Ann Surg.. Epub ahead of print2019. DOI: 10.1097/sla.0000000000003276.
24. George EL, Hall DE, Youk A, et al.Association between Patient Frailty and Postoperative Mortality across Multiple Noncardiac Surgical Specialties. JAMA Surg. 2020;94305:1–9.
25. Mukaka MM. A guide to appropriate use of Correlation coefficient in medical research. Malawi Med J. 2012;24:69–71. [PubMed: 23638278]
26. Haga Y, Ikei S, Ogawa M. Estimation of physiologic ability and surgical stress (E-PASS) as a new prediction scoring system for postoperative morbidity and mortality following elective gastrointestinal surgery. Surg Today. 1999;29:219–225. [PubMed: 10192731]
27. Smith PK, Mayer JE, Kanter KR, et al.Physician Payment for 2007: A Description of the Process by Which Major Changes in Valuation of Cardiothoracic Surgical Procedures Occurred. Ann Thorac Surg. 2007;83:12–20. [PubMed: 17184623]
28. Balentine CJ, Naik AD, Berger DH, et al.Postacute care after major abdominal surgery in elderly patients intersection of age, functional status, and postoperative complications. JAMA Surg. 2016;151:759–766. [PubMed: 27144881]
29. Handoll HHG, Cameron ID, Mak JCS, et al.Multidisciplinary rehabilitation for older people with hip fractures. Cochrane Database Syst Rev.. Epub ahead of print2009. DOI: 10.1002/14651858.CD007125.pub2.
30. Carli F, Bousquet-Dion G, Awasthi R, et al.Effect of Multimodal Prehabilitation vs Postoperative Rehabilitation on 30-Day Postoperative Complications for Frail Patients Undergoing Resection of Colorectal Cancer: A Randomized Clinical Trial. JAMA Surg. 2020;155:233–242. [PubMed: 31968063]
31. Moran J, Guinan E, McCormick P, et al.The ability of prehabilitation to influence postoperative outcome after intra-abdominal operation: A systematic review and meta-analysis. Surg (United States). 2016;160:1189–1201.

32. Hall DE, Arya S, Schmid KK, et al.Association of a frailty screening initiative with postoperative survival at 30, 180, and 365 days. JAMA Surg. 2017;152:233–240. [PubMed: 27902826]

33. Min L, Hall K, Finlayson E, et al.Estimating risk of postsurgical general and geriatric complications using the VESPA preoperative tool. JAMA Surg. 2017;152:1126–1133. [PubMed: 28768325]

34. Johnston KJ, Wen H, Joynt Maddox KE. Relationship of a claims-based frailty index to annualized Medicare costs: A cohort study. Ann Intern Med. 2020;172:533–540. [PubMed: 32252070]

35. Johnston KJ, Bynum JPW, Joynt Maddox KE. The Need to Incorporate Additional Patient Information into Risk Adjustment for Medicare Beneficiaries. JAMA. 2020;323:925–926. [PubMed: 31999298]

36. Kupfer J, Witmer RS, Do V. Caring for those who serve: Potential implications of the veterans affairs maintaining internal systems and strengthening integrated outside networks act of 2018. Ann Intern Med. 2018;169:487–489. [PubMed: 30039157]

37. Kullgren JT, Fagerlin A, Kerr EA. Completing the MISSION: a Blueprint for Helping Veterans Make the Most of New Choices. J Gen Intern Med. 2020;35:1567–1570. [PubMed: 31646457]

38. Wilson NJ, Kizer KW. The VA Health Care System: An Unrecognized National Safety Net. Health Aff1997;16:200–204.

39. Lorence DRegional Variation in Medical Classification Agreement: Benchmarking the Coding Gap. J Med Syst2003;27:435–443. [PubMed: 14584620]

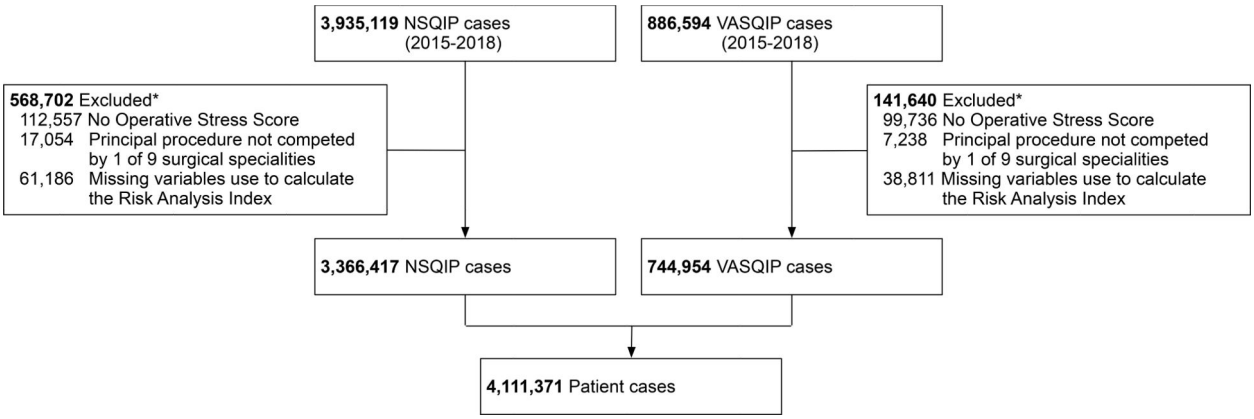

FIGURE 1.
Case Number and Exclusion Criteria

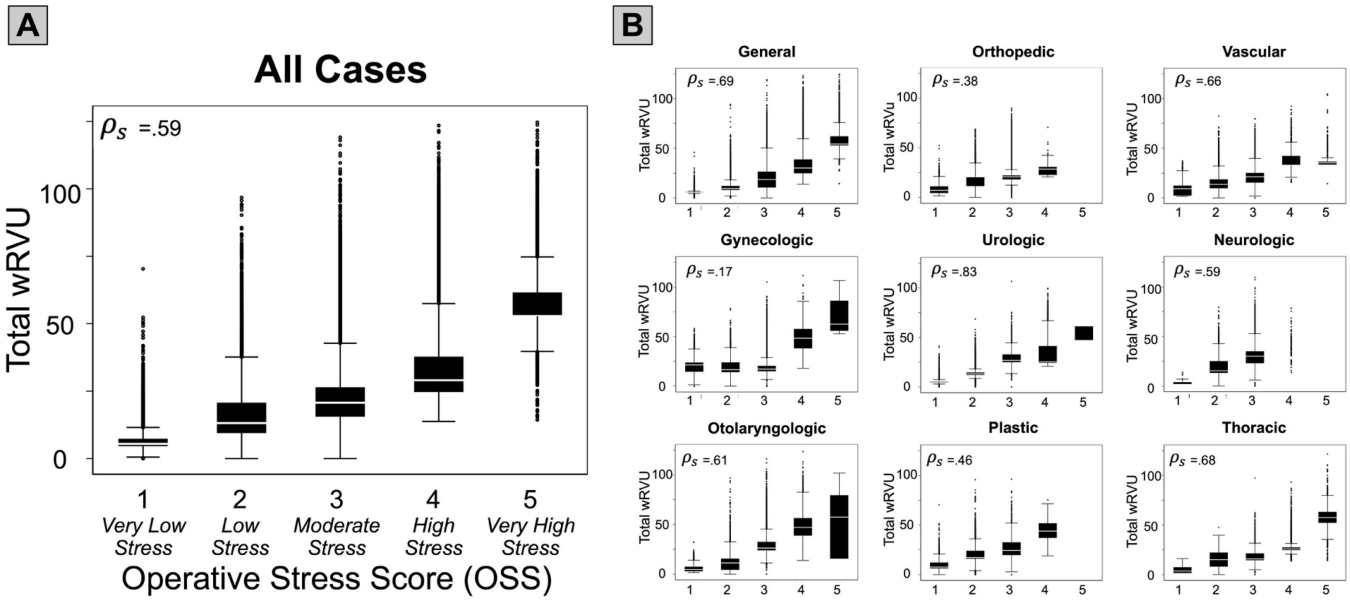

FIGURE 2.

Correlation between operative stress and total work Relative Value Unit for all procedures (Panel A) and by surgical specialty (Panel B)

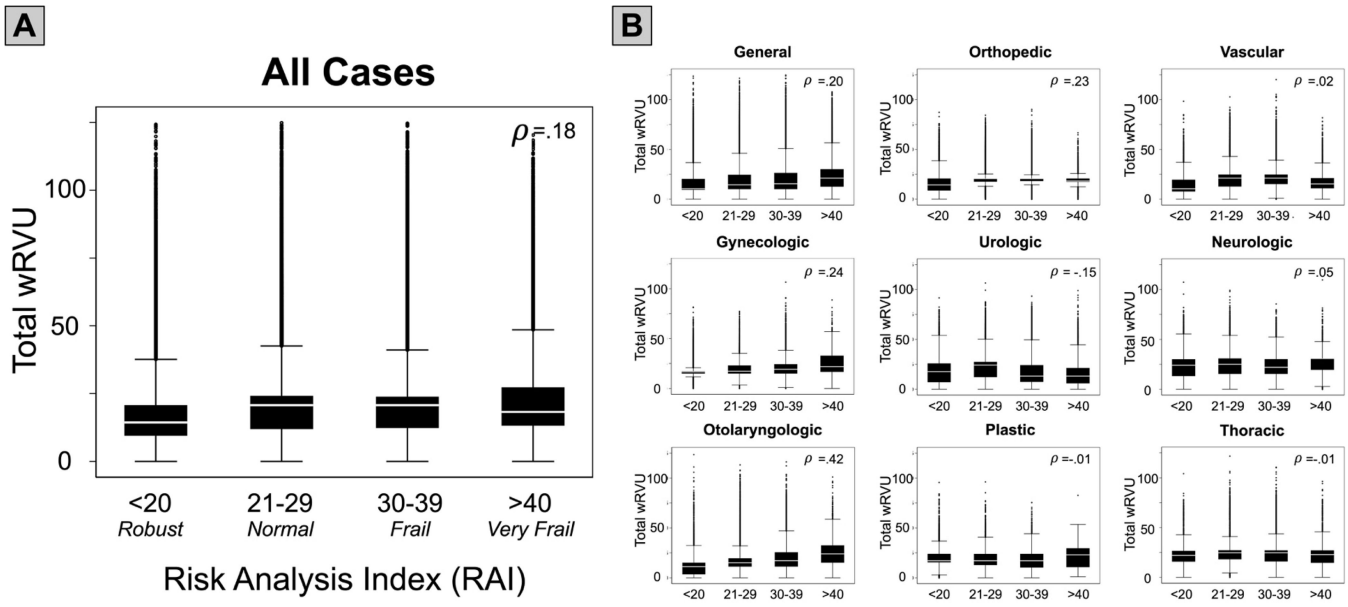

FIGURE 3.

Correlation between patient frailty and total work Relative Value Unit for all procedures (Panel A) and by surgical specialty (Panel B)

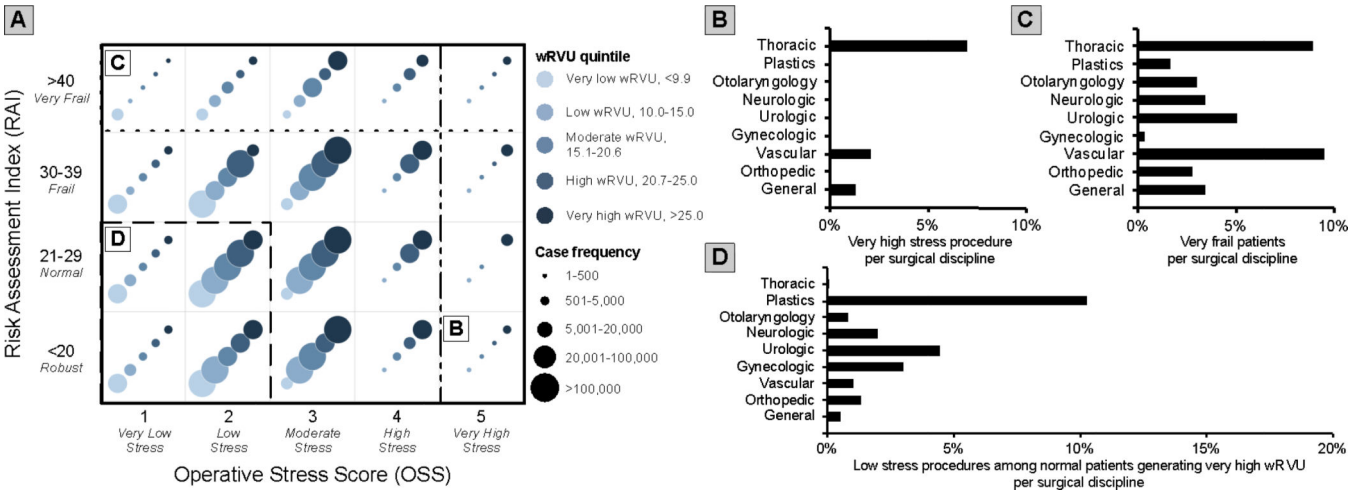

FIGURE 4.

Relationship between case frequency, operative stress, patient frailty, and total work Relative Value Unit

TABLE 1.

Cases by surgical specialty with single (n = 2,817,440) or multiple procedures (n = 1,293,931) performed

Surgical specialty	All cases				Cases with a single wRVU contributing procedure*		Cases with multiple wRVU contributing procedures†	
	n Case	Total wRVU, mean (SD)	RAI, mean (SD)	OSS, median (Range)	n Case (% per surgical specialty)	Total wRVU, mean (SD)	n Case (% per surgical specialty)	Total wRVU, mean (SD)
All	4,111,371	18.4 (10.0)	24.9 (8.3)	3 (2–3)	2,817,440 (69%)	15.7 (7.4)	1,293,931 (31%)	25.8 (12.7)
General	1,803,368	17.4 (11.4)	23.0 (8.6)	3 (2–3)	1,356,638 (75%)	14.2 (8.2)	446,730 (25%)	26.9 (14.2)
Orthopedic	1,067,411	18.5 (6.4)	25.9 (7.2)	2 (2–3)	904,946 (85%)	17.8 (5.5)	162,465 (15%)	22.2 (9.3)
Vascular	279,115	19.8 (9.6)	30.6 (7.1)	3 (2–3)	206,705 (74%)	18.1 (7.6)	72,410 (26%)	24.9 (12.6)
Gynecologic	200,620	18.3 (5.7)	18.3 (5.6)	3 (3–3)	119,333 (59%)	15.6 (2.1)	81,287 (41%)	22.3 (7.0)
Urologic	289,841	18.6 (11.7)	29.1 (6.4)	2 (2–3)	211,569 (73%)	15.6 (9.0)	78,272 (27%)	26.4 (14.4)
Neurologic	229,632	23.8 (9.3)	25.3 (7.4)	3 (2–3)	172,099 (75%)	20.7 (7.3)	57,533 (25%)	33.2 (8.5)
Otolaryngologic	95,936	15.1 (11.3)	21.4 (9.4)	2 (2–2)	72,106 (75%)	11.9 (6.8)	23,830 (25%)	24.8 (15.8)
Plastic	86,338	18.3 (8.2)	19.5 (7.7)	2 (2–2)	40,805 (47%)	13.6 (5.3)	45,533 (53%)	22.6 (8.0)
Thoracic	59,110	25.3 (11.5)	29.6 (7.4)	4 (3–4)	34,432 (58%)	22.2 (9.1)	24,678 (42%)	29.5 (13.1)

*Cases with a single procedure contributing to work Relative Value Unit summation for the primary surgical team.

†Includes only multiple procedures which contribute to the work Relative Value Unit, as defined by the *Current Procedural Terminology* code(s) multiple procedure modifier (Supplemental Table 1).

SD indicates standard deviation; RAI, risk analysis index; OSS< operative stress score; wRVU, work Relative Value Unit.

TABLE 2.

Patient and operative characteristics

Variables	Very low total wRVU: <9.9	Low total wRVU: 10.0–15.0	Moderate total wRVU: 15.1–20.6	High total wRVU: 20.7–25.0	Very high total wRVU: >25.0
	n = 827,393	n = 838,315	n = 709,740	n = 931,588	n = 804,335
Demographics					
Age, years, mean (SD)	52 (18)	55 (16)	58 (17)	64 (13)	61 (14)
Male sex, n (%)	538,639 (65%)	383,740 (46%)	298,015 (42%)	507,690 (54%)	450,131 (56%)
Race, n (%)					
White	576,721 (83.2%)	593,715 (83.6%)	488,731 (81.0%)	679,636 (85.0%)	583,677 (83.7%)
Black	88,357 (12.7%)	89,141 (12.6%)	93,906 (15.6%)	98,895 (12.4%)	92,188 (13.2%)
Asian or Pacific Islander	23,099 (3.3%)	21,540 (3.0%)	16,591 (2.7%)	15,185 (1.9%)	17,397 (2.5%)
American Indian or Alaska Native	5,377 (0.8%)	5,654 (0.8%)	4,392 (0.7%)	5,975 (0.7%)	4,249 (0.6%)
Ethnicity, n (%)					
Not Hispanic	623,536 (89.5%)	639,512 (89.3%)	556,902 (91.6%)	756,365 (94.4%)	660,724 (93.2%)
Hispanic	73,006 (10.5%)	76,310 (10.7%)	50,842 (8.4%)	44,996 (5.6%)	48,043 (6.8%)
Current smoking, n (%)	186,291 (22.5%)	171,272 (20.4%)	138,052 (19.5%)	186,646 (20.0%)	177,674 (22.1%)
Preoperative functional status and comorbidities					
Body mass index, mean (SD)	29.3 (10.5)	30.5 (7.2)	30.7 (10.0)	30.9 (7.9)	30.2 (20.2)
Risk Analysis Index, n (%)					
Robust (≤20)	355,359 (42.9%)	334,345 (39.9%)	245,933 (34.7%)	132,890 (14.3%)	166,823 (20.7%)
Normal (21–29)	292,070 (35.3%)	321,198 (38.3%)	262,437 (37.0%)	493,412 (53.0%)	381,763 (47.5%)
Frail (30–39)	157,153 (19.0%)	156,937 (18.7%)	164,761 (23.2%)	280,603 (30.1%)	213,152 (26.5%)
Very frail (≥40)	22,811 (2.8%)	25,835 (3.1%)	36,609 (5.2%)	24,683 (2.6%)	42,597 (5.3%)
Baseline Risk Factors, n (%)					
COPD	43,425 (5.2%)	47,120 (5.6%)	47,767 (6.7%)	67,688 (7.3%)	62,627 (7.8%)
Congestive heart failure	27,641 (3.3%)	30,338 (3.6%)	30,172 (4.3%)	42,607 (4.6%)	28,881 (3.6%)
Diabetes – oral therapy	65,745 (7.9%)	83,115 (9.9%)	78,131 (11.0%)	121,543 (13.0%)	98,101 (12.2%)
Diabetes – insulin	53,546 (6.5%)	59,304 (7.1%)	59,076 (8.3%)	58,706 (6.3%)	60,484 (7.5%)
Hypertension	320,016 (38.7%)	367,616 (43.9%)	357,703 (50.4%)	572,664 (61.5%)	444,582 (55.3%)

Variables	Very low total wRVU: <9.9	Low total wRVU: 10.0–15.0	Moderate total wRVU: 15.1–20.6	High total wRVU: 20.7–25.0	Very high total wRVU: >25.0
	n = 827,393	**n = 838,315**	**n = 709,740**	**n = 931,588**	**n = 804,335**
Renal failure	2,916 (0.4%)	2,387 (0.3%)	2,400 (0.4%)	1,555 (0.2%)	2,681 (0.4%)
Dialysis	138,17 (1.7%)	9,709 (1.2%)	11,810 (1.7%)	5,538 (0.6%)	8,282 (1.0%)
Disseminated cancer	6,927 (0.8%)	8,885 (1.1%)	12,603 (1.8%)	11,095 (1.2%)	38,932 (4.8%)
Operative information					
VASQIP, n (%)	129,047 (15.6%)	134,500 (16.0%)	125,014 (17.6%)	222,136 (23.8%)	134,257 (16.7%)
Anesthesia, n (%)					
General	725,272 (87.7%)	781,201 (93.2%)	656,973 (92.6%)	670,323 (72.0%)	780,362 (97.0%)
Monitored	75,339 (9.1%)	32,367 (3.9%)	18,548 (2.6%)	59,812 (6.4%)	8,031 (1.0%)
Spinal/Epidural	14,097 (1.7%)	18,125 (2.2%)	28,311 (4.0%)	187,325 (20.1%)	13,553 (1.7%)
Regional/Local	11,797 (1.4%)	6,013 (0.7%)	5,316 (0.7%)	13,343 (1.4%)	1,633 (0.2%)
Other	710 (0.1%)	435 (0.1%)	403 (0.1%)	680 (0.1%)	676 (0.1%)
Emergency case, n (%)	134,355 (16.2%)	57,692 (6.9%)	55,805 (7.9%)	32,081 (3.4%)	57,074 (7.1%)
American Society of Anesthesiologists Classification, n (%)					
1	132,828 (16.1%)	72,428 (8.6%)	38,307 (5.4%)	24,093 (2.6%)	14,452 (1.8%)
2	373,375 (45.1%)	393,170 (46.9%)	272,908 (38.5%)	344,769 (37.0%)	254,131 (31.6%)
3	278,289 (33.6%)	327,620 (39.1%)	340,035 (47.9%)	504,220 (54.1%)	458,167 (57.0%)
4	39,238 (4.7%)	42,128 (5.0%)	56,787 (8.0%)	56,343 (6.0%)	73,827 (9.2%)
5	405 (<1%)	1,010 (0.1%)	761 (0.1%)	1,238 (0.1%)	2,926 (0.4%)
Operative Stress Score, n (%)					
1	116,236 (14.0%)	13,872 (1.7%)	3,186 (0.4%)	500 (0.1%)	267 (<1%)
2	668,689 (80.8%)	473,261 (56.5%)	293,989 (41.4%)	412,734 (44.3%)	72,852 (9.1%)
3	42,468 (5.1%)	350,761 (41.8%)	407,527 (57.4%)	452,600 (48.6%)	516,452 (64.2%)
4	0 (0.0%)	278 (<1%)	4,986 (0.7%)	65,735 (7.1%)	180,931 (22.5%)
5	0 (0.0%)	143 (<1%)	52 (<1%)	19 (<1%)	33,833 (4.2%)
Postoperative Information					
Inpatient admission, n (%)	217,222 (26.3%)	305,720 (36.5%)	436724 (61.5%)	809,606 (86.9%)	729,009 (90.6%)
30-day mortality, n (%)	2,851 (0.3%)	4,374 (0.5%)	7,897 (1.1%)	5,866 (0.6%)	13,198 (1.6%)

COPD, Chronic Obstructive Pulmonary Disease; SD indicates standard deviation; VASQIP, Veterans Affairs Quality Improvement Project; wRVU, work Relative Value Unit.

Table 3.

Commonest Principal Procedures Among Very Low (OSS 1) and Low (OSS 2) Operative Stress Cases Generating a Very High Work Relative Value Unit (wRVU>25.0) in Non-Frail (RAI, <30) patients (n = 60,128)

Surgical specialty	Principal Procedure Description	Frequency n*	%
General	Mastectomy	7,238	12.0%
	Thyroidectomy	191	0.3%
	Parathyroidectomy	179	0.3%
	Incisional hernia repair	116	0.2%
Orthopedic	Knee arthroscopy or arthroplasty	6,905	11.5%
	Shoulder arthroscopy	4,697	7.8%
	Rotator cuff repair, open	160	0.3%
Vascular	Thromboendarterectomy (common, deep, or superficial artery)	1,007	1.7%
	Endovascular stent or angioplasty	868	1.4%
	Embolectomy	567	0.9%
Gynecologic	Hysterectomy	5,597	9.3%
	Stress incontinence sling	599	1.0%
Urologic	Laparoscopic partial nephrectomy	12,767	21.2%
Neurologic	Arthrodesis, below C2	4,862	8.1%
	Laminotomy or laminectomy for decompression	1,063	1.8%
	Discectomy	536	0.9%
Otolaryngologic	Thyroidectomy	538	0.9%
	Parotid resection	102	0.2%
Plastic	Breast reconstruction	5,739	9.5%
	Breast revision	1,214	2.0%
	Reduction mammaplasty	1,164	1.9%
	Panniculectomy	1,091	1.8%

*Includes only principal procedures for cases with an operative stress score <2, Risk Analysis Index<30, and a total work Relative Value Unit >25.0 for >100 cases. The total percentages are the frequency of low stress procedures, on non-frail patients, with high reimbursement cases among all cases in each specialty and therefore procedures listed do not sum to 100%.