Published in final edited form as:

Med Care. 2016 June ; 54(6): 623–631. doi:10.1097/MLR.0000000000000536.

Practice Variation among Hospitals in Revascularization Therapy and Its Association with Procedure-Related Mortality

Jarrod E. Dalton, PhD,

Departments of Quantitative Health Sciences and Outcomes Research, Cleveland Clinic, Cleveland, Ohio

David A. Zidar, MD, PhD,

Harrington Heart and Vascular Institute, Case Western Reserve University/University Hospitals Case Medical Center, Cleveland, Ohio

Belinda L. Udeh, PhD,

Department of Outcomes Research, Cleveland Clinic, Cleveland, Ohio

Manesh R. Patel, MD,

Duke Clinical Research Institute, Duke University, Durham, North Carolina

Jesse D. Schold, PhD, and

Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, Ohio

Neal V. Dawson, MD

Center for Healthcare Research and Policy, Case Western Reserve University/MetroHealth Medical Center, Cleveland, Ohio

Abstract

Background—While substantial practice variation in coronary revascularization has been described and deviation from clinical practice guidelines has been associated with worse outcomes, the degree to which this is driven by flawed decision-making and/or appropriate deviation associated with comorbid conditions is unknown. We evaluated heterogeneity in procedure use, and the extent to which hospital-level practice variation is related to surgical mortality.

Methods—We analyzed data on 554,563 inpatients undergoing either percutaneous coronary intervention (PCI) or coronary artery bypass grafting (CABG) at 391 centers in six states. Procedure-specific risk models were developed based on demographics and comorbidities, allowing for differential effects of comorbidities for each sex. For each patient, the revascularization procedure which minimized predicted probability of in-hospital mortality was designated as the model-preferred procedure.

Hospital-level discordance rates (HDRs) – the proportion of cases in each hospital for which the opposite from the model-preferred procedure was performed – were calculated. Hierarchical linear

Address Correspondence To: Jarrod E. Dalton, PhD, 9500 Euclid Avenue (P77), Cleveland, OH, 44195, Telephone: (216) 210-5197, Fax: (216) 444-6135, daltonj@ccf.org.

Conflicts of Interest: None.

models were used to analyze the relationship between HDRs and hospital-level risk-standardized mortality ratios (RSMRs).

Results—Comorbidities and demographics alone explained between 68% and 86% of overall variation in in-hospital mortality (corresponding C-statistics of 0.84–0.93). The mean (SD) HDR was 26.3% (9.6%). There was a positive independent association between HDRs and in-hospital mortality, with a 10% increase in HDR associated with an 11% increase in RSMR ($P<0.001$).

Conclusions—Variance in procedure use according to model preference was strongly associated with worse outcomes. A systematic approach to incorporating comorbidity as part of the decision-making process for coronary revascularization is needed.

Keywords

revascularization; mortality; risk model; coronary artery bypass graft; percutaneous coronary intervention

Approximately 1.35 million coronary revascularization procedures are performed annually in the United States, of which 71% are percutaneous coronary interventions (PCI) and 29% are coronary artery bypass grafting surgeries (CABG)¹. While meta-analysis of ten major cardiology trials² found comparable overall long-term mortality between the two procedures (with exception to elderly or diabetic patients, in whom CABG was associated with reduced mortality rates), the recent SYNTAX trial³ found that anatomic complexity of coronary artery disease (CAD) was associated with varying treatment effects; particularly, patients in the highest tertile of anatomic complexity (as represented by the SYNTAX score) had reduced mortality under CABG, although the extent to which disease anatomy was *independently* responsible for causing the observed differences in outcomes – particularly regarding comorbid conditions that co-vary with anatomical complexity – is unclear.

Practice guidelines for revascularization⁴ reflect these findings, suggesting that both procedures are appropriate for multi-vessel CAD, although CABG is more appropriate among patients with certain types of complex 3-vessel CAD or CAD with left main stenosis. Yet, wide practice variation in revascularization therapy persists^{5, 6}, and to date little is known as to why this is the case. One potential source of practice variation relates to decisions to proceed with major surgical revascularization among patients presenting with complex sets of comorbidities.

Preoperative evaluation is critical for informing surgical mortality risk under CABG. In practice, two risk models are commonly used: the EuroSCORE II⁷ and the Society of Thoracic Surgeons (STS) score⁸. However, the consistency among providers and hospitals with which revascularization decisions are informed by these risk scores is unknown. Furthermore, both risk scores incorporate only a small amount of information on comorbid conditions, despite evidence that comorbidities are strongly associated with post-operative in-hospital mortality. Validation studies^{9–11} have produced C-statistics ranging from 0.71 to 0.84 for the EuroSCORE II and STS models, depending on the populations studied. These estimates suggest that, in the best case, the models account for 68% of overall rank order discrimination for in-hospital mortality, and in the worst case, only 42%. The remaining

uncertainty is determined by either factors external to these models or by more complex relationships among factors internal to the models.

We have recently proposed a modeling methodology that yields measures of differential predicted risk among two (or more) treatments for a specific covariate profile¹². In the current study, we applied this methodology to evaluate observed hospital-level variation in the use of PCI and CABG and its relationship with in-hospital mortality after revascularization. Specifically, we compared risk-standardized mortality rates among hospitals in relation to a model that mapped distinct sets of comorbidities to predicted risk of mortality under both procedures. We hypothesized that hospitals with higher rates of *discordance* from the model-preferred procedure (defined as the procedure that minimized predicted mortality risk) would have higher risk-adjusted mortality.

Methods

Data Sources and Inclusion/Exclusion Criteria

We obtained administrative data on all hospital inpatient discharges occurring in six states (Arizona, California, Florida, Maryland, Michigan, and New Jersey) between 2009 and 2011 from the US Agency for Healthcare Research and Quality. We included discharges having either CABG (International Classification of Diseases, Ninth Revision, Clinical Modification, ICD-9-CM, procedure code of 36.1x) or PCI (ICD-9-CM procedure code of 36.09 or 00.66) as the principal procedure. CABG patients with stenting were classified into the CABG group. Hospitals with <100 total revascularizations over the three-year time period were excluded.

Risk Modeling

Our risk modeling approach was centered on the creation of a decision rule that mapped patients' covariate values to a model-preferred revascularization treatment – either CABG or PCI – based on which procedure had the lower estimated probability of in-hospital mortality. These estimated probabilities were based on procedure-specific logistic regression models (developed in a manner that ignored hospital assignment). We felt that the use of separate models for CABG and PCI more directly allowed the possibility of modeling fundamental differences between the two procedures in the nature of the relationship between covariates and mortality. Details of this treatment effectiveness modeling approach are presented in Dalton et al. (2015)¹².

The following variables were used as covariates in the logistic regression models: patient age and sex, state and year of discharge, type of admission (e.g., emergency, transfer, etc.), patient locale (National Center for Health Statistics' Urban-Rural Code), and present-on-admission diagnoses (ICD-9-CM codes). Since many diagnosis codes are sparsely-represented, we aggregated certain diagnoses together by truncating trailing digits. Diagnoses were truncated (up to a minimum of 3 characters) if they were represented by fewer than 500 discharges. As a result of this aggregation routine, 746 distinct diagnosis-related covariates were created and considered for the models; 50.1% of diagnoses were not aggregated at all, 43.0% were truncated to 4 digits, 5.5% were truncated to 3 digits, and

1.4% did not meet minimal sample size requirements after aggregation and were thus mapped to an all-purpose “other diagnoses” covariate.

Age, sex, state and year of discharge, type of admission, transfer status, and patient locale were forced into all risk models, while diagnosis-related covariates were selected using “elastic net” penalization¹³. Since revascularization may have different risk profiles between men and women^{14–19}, we further considered sex-specific diagnosis effects by evaluating two-way interactions. Issues involving computational complexity necessitated the use of a simplified routine to identify these interactions – specifically, we included an interaction between sex and a given diagnosis if, among patients with the diagnosis in question, there were at least 10 deaths for each sex and a univariate chi-squared test for association in the sex-by-mortality contingency table was significant at the 0.05 level. Once all predictors and two-way interactions with sex were selected, we refit the final models and estimated the probability of in-hospital mortality for each discharges’ specific covariate values, under each of the two treatments (regardless of which one was actually administered).

Discordance from the Model—Discharges were subsequently characterized as either *concordant* or *discordant*. Concordance was defined as occurring when the treatment actually received was the same as the model-preferred treatment (i.e., the one with the lower estimated probability of in-hospital mortality). Discordance was defined as occurring when patients received the opposite of the model-preferred treatment. *Hospital discordance rates* (HDRs) were then defined as the proportion of discordant discharges among all those treated at a given hospital. We clarify here that our use of the term ‘discordance’ simply meant that the model-preferred treatment and actual treatment administered were not the same. It does not imply that physicians’ actively chose a treatment against information provided by the model.

HDRs were assessed for heterogeneity using both a histogram and a Bayesian beta-binomial model²⁰. The beta-binomial distribution is a *mixture* distribution for binary outcomes characterized by multiple component binomial distributions, each with their own unique probability parameters; we thus modeled hospital-specific binomial distributions, each with their own HDR parameter. We used a non-informative Uniform(0,1) prior probability distribution for the HDRs. Posterior estimates of mean and standard deviation of the HDRs were then obtained from this model.

Relationship between Hospital Discordance Rates and Risk Standardized Mortality Ratios—We implemented our primary analysis investigating the relationship between HDRs and risk-adjusted in-hospital mortality using risk-standardized mortality ratios (RSMRs). Although the objective of this analysis was to estimate associations between hospital-level characteristics and mortality – and not to report hospital-specific quality measures – we followed the methodology for estimating RSMRs commissioned by the Committee of Presidents of Statistical Societies and funded by the Centers for Medicare and Medicaid Services²¹, with the exception that we used a frequentist hierarchical linear model instead of a Bayesian hierarchical linear model.

RSMRs were defined for each hospital as the number of observed deaths (during the three-year study period) divided by the number of deaths predicted by our risk-adjustment model. Specifically, the expected number of deaths was defined (for each respective hospital) as the total number of predicted deaths (sum of individual patients' predicted probabilities) under the model-preferred treatment.

Hierarchical logistic regression models²² were used to analyze the relationship between HDRs and hospital-level RSMRs. These models contained random intercepts to account for correlation among patients treated at the same hospital. An offset term, equal to each patient's predicted log-odds of mortality under the preferred treatment (or actual treatment, in the case of the negative control analysis; see below), was included in order to directly model RSMRs. The following variables were included as fixed effects: HDR, source of admission, an indicator variable for whether or not the hospital performs CABG, hospital revascularization volume, and a propensity score defined as patients' estimated probability of receiving CABG given their age, sex, and aggregated (as described above under "Risk Modeling") diagnoses. The slope for the HDR effect, and corresponding 95% confidence interval (CI) was then reported as a measure of the primary relationship of interest between hospital-level discordance from the risk model and hospital-level risk-standardized mortality. To evaluate potential nonlinearities, a separate model that incorporated restricted cubic splines for the HDR effect was visualized. Propensity scores were estimated using logistic regression.

Sensitivity Analyses

Changing the definition of expected mortality for a given hospital from the sum of patients' estimated probabilities under the model-preferred treatment to the sum of estimated probabilities *under the actual treatment administered* would impact the hospital's RSMR as the number of expected deaths would be greater. The reason for the relative increase in expected number of deaths is that, while estimated probabilities would not change for patients assigned the treatment that was in agreement with the model recommendation, estimated probabilities would be greater under the actual treatment given among patients assigned the opposite of the model-preferred treatment. (For example, if predicted mortality risk for a given group of patients was 2% under PCI and 4% under CABG, and the patients received PCI, then the expected number of deaths under the model-preferred treatment would be 0.02 under both the model-preferred treatment and the actual treatment, which in this case would both be PCI. Had the patients received CABG, however, then the expected number of deaths under the actual treatment would have been 0.04.)

Assuming our models accurately captured risk, using the number of expected deaths under the actual treatment administered would therefore effectively nullify the association between hospital-level discordance and RSMRs as the extra risk would be accounted for within the denominator. We evaluated this assumption in a negative control analysis wherein the primary association of interest (percent change in hospital-level RSMR associated with a relative HDR increment of 10%) was estimated using patients' estimated risk under actual treatment to define the expected number of deaths for each hospital. If, under this model, there was no association observed (in the presence of a positive association in the primary

analysis), we would then conclude that observed associations were not artifacts due to erroneous risk estimation – at least with respect to the available covariates.

In a more detailed Monte Carlo sensitivity analysis (see Supplemental Digital Content, **under “Simulated Effects of Major Unobserved Confounding Variable”**), we assessed the degree to which preferred treatment – and ultimately our results – might have been affected by an unavailable independent major confounding variable (e.g., coronary lesion complexity).

Finally, in a third sensitivity analysis, we estimated the primary association of interest within certain restricted samples, in order to evaluate the extent to which our observed results may have been affected by inclusion of specific subpopulations. We excluded: i) patients with ST-elevated myocardial infarctions (STEMIs, ICD-9-CM diagnosis code of 401.X, excluding 401.7), who tend to systematically receive PCI and have greater risk of mortality; ii) transfer patients, whose risk profile may be systematically different from that of other sources of admission; iii) CABG patients also undergoing secondary PCI, whose risk might not appropriately be captured by our models; and iv) patients whose predicted risk was nearly equivocal between CABG and PCI (defined as predicted risk under CABG within ± 0.01 of predicted risk under PCI), effectively restricting the analysis to the subset of patients among whom discordance might most clearly be associated with higher risk of mortality.

Results

Of the 563,758 discharges meeting entry criteria, 1.3% of the records were removed due to missing data (Figure 1); 116/507 hospitals (constituting 0.4% of the remaining sample) treated <100 patients in the three-year time period and were excluded from analysis. We thus analyzed data on 554,563 patients (98.4% of all discharges) treated at 391 hospitals. The crude incidence of in-hospital mortality among these hospitals ranged from 0% to 6.6%, with a median [first quartile, third quartile] of 1.3% [0.9%, 1.7%].

Our model for in-hospital mortality after CABG included main effects for 266 diagnosis-related covariates (in addition to the demographic variables forced into the model), and 42 two-way interactions with sex (among those 266 covariates). Similarly, the model for mortality after PCI included 374 diagnosis-related covariates and 68 two-way interactions with sex. Discrimination and calibration of these models were good for each combination of sex and observed treatment (Figure 2); C-statistics ranged between 0.84 (females) and 0.86 (males) for the CABG models and between 0.91 (females) and 0.93 (males) for the PCI models.

Overall, the risk model-based decision rule preferred CABG for 10% of discharges and PCI for the remaining 90%. The frequency of discordance with the model-preferred treatment was 27.8%; this frequency was 77.2% among discharges with CABG as the model-preferred treatment and was 22.3% among discharges with PCI as the model-preferred treatment. A summary of model-preferred treatments, discordance rates and in-hospital mortality among selected subpopulations is presented in Table 1.

HDRs ranged from 3.6% to 59.0% among the 391 hospitals analyzed. Based on our Bayesian beta-binomial model, the HDRs had a mean (95% credibility interval) of 26.3% (25.3% – 27.2%) and a standard deviation of 9.6% (9.0% – 10.4%). Hospitals with discordance rates >30% generally had greater volume, more elective transfers, fewer emergency cases, more patients from non-urban locales, were more likely to perform CABG, and had higher median propensity scores (estimated probability of receiving CABG given patients' covariates) than hospitals with discordance rates ≤20% (Table 2).

Propensity score distributions (estimated probability of receiving coronary artery bypass grafting, given information on patient characteristics and present-on-admission diagnoses, see Supplemental Digital Content, Figure S-1) indicate considerable overlap between patients who received PCI and patients who received CABG, despite general tendency towards lower propensity scores for PCI patients than CABG patients. Propensity scores for both groups ranged from <0.1% to >99%.

Our primary hierarchical linear model indicated a positive association between HDR and in-hospital mortality after adjusting for source of admission, hospital-level use of CABG, volume, and propensity score, with a 10% increase in HDR associated with a 11% increase (95% CI: 5% – 17% increase) in risk-standardized mortality ($P<0.001$; Figure 3). Comparing representative HDRs from the two tails of the observed distribution, hospitals with discordance rates of 40% had 37% (16% – 60%) higher risk-standardized mortality than hospitals with discordance rates of 10%. Based on the dashed lines in Figure 3 – which represent pointwise 95% confidence interval estimates for the displayed relationships under (respective) models that allowed for nonlinearities via restricted cubic splines – we concluded that the above linear characterization was justifiable.

In our negative control analysis, where we switched the definition of expected mortality from the number expected under the model-preferred treatment to the number expected under the actual treatment administered, this association was no longer significant [estimated slope for a 10% increase in HDR of -1% ($-7\% - +5\%$); $P=0.69$], indicating that our models accurately captured risk with respect to the available covariates.

Effects of an unavailable independent covariate on the primary observed association were minimal despite large hypothetical prevalences and large hypothetical differential effects on mortality between PCI and CABG (see Supplemental Digital Content, Figure S-2).

In the third sensitivity analysis, exclusion of neither discharges with STEMI ($n = 108,191$; 10% increase in HDR associated with a 21% [13% – 30%] increase in hospital-level RSMR after removing these patients), transfer patients ($n = 98,114$; 11% [5% – 18%]), CABG patients who had secondary PCI ($n = 2,351$; 11% [5% – 18%]), nor patients with equivocal predicted risk under PCI and CABG ($n = 365,125$ with predicted risk under CABG within ± 0.01 of predicted risk under PCI; 15% [9% – 21%]) resulted in meaningful qualitative changes to the primary relationship of interest.

Discussion

Model-based treatment recommendations can be desirable when their incorporation into shared decision-making practice leads to improved patient-centered outcomes. Here, we sought to use novel risk-based models as a platform to assess hospital-level variation in revascularization activities, independent of guidelines or appropriateness criteria. We found that discordance in the mode of revascularization in relation to our risk-based model's preferred treatment was strongly associated with increased risk-adjusted in-hospital mortality. This suggests that, in aggregate, variation in preoperative evaluation practice in relation to patients' comorbidity profiles may be associated with worse outcomes and highlights the need to better understand its sources.

Many possibilities exist. First, undesirable variation may be introduced as a result of socio-demographic characteristics. Hannon et al. found that African-American and Hispanic patients were 35–40% less likely to undergo coronary artery bypass grafting (CABG) than similarly-indicated and similarly-insured white non-Hispanic patients²³. Female Medicare beneficiaries were observed to have a significantly lower risk-standardized chance of receiving CABG (27%) than their male counterparts (33%) and a significantly higher chance of receiving PCI (42% vs. 38%)²⁴. Finally, socioeconomic status has been associated with both access to²⁵ and outcomes after²⁶ CABG. In this study, females and African Americans received CABG less often despite comparable or even increased proportions that were recommended for CABG by the model. While we did not directly evaluate socioeconomic status, we note that the proportion of patients with private insurance who had model recommendations for CABG was lower than that among patients with other payers, despite comparable proportions actually receiving CABG.

Practice variation is also influenced by physician and practice-level factors, like regret aversion, perceived low risk, avoidance of patient anxiety, and litigation risk²⁷. Hospital-level variation may also be a function of logistical considerations including the preference of the referring physician, the on-site availability of cardiothoracic surgery, and economic incentives. Group-based decision-making processes like those of heart teams may reduce some forms of practice variation, but do not eliminate it altogether. For example, Bennett-Guerrero et al. identified enormous hospital-level variation in the use of erythrocyte transfusion among CABG patients²⁸. Furthermore, heart-team based processes may encourage an overly consistent approach to evaluating comorbidity risk within teams when variation in comorbidity risk may best be characterized by a complex predictive model. While clinical practice guidelines and appropriateness criteria can be used to inform decision making tendencies, substantial knowledge gaps limit the applicability in many real world settings, and limit their utility in the quantification of practice variation. Moreover, studies have suggested that physicians, like all humans, vary in their ability to accurately assess probabilities of events^{29–31}.

Our risk modeling framework used modern statistical algorithms to distinguish in-hospital mortality risk separately for both CABG and PCI using baseline covariates derived from a large, diverse, population-based registry. While this model did not account for anatomic factors such as atherosclerotic burden and lesion complexity – which are likely to heavily

influence the decision for PCI versus CABG in practice – our models nonetheless explained between 68% and 86% of overall rank order discrimination for in-hospital mortality among all four combinations of sex and procedure, with C-statistics ranging from 0.84 to 0.93. (We arrived at these figures via conversion of C-statistics to percent of rank-order discrimination explained, given by $\% \text{ explained} = 2(C - 0.5)$.) This suggests that only modest amounts of residual discrimination (14–32%) are left to be independently explained by unavailable factors such as complex anatomy, socioeconomic status, education, and preferences. Furthermore, our sensitivity analyses suggest that our observed relationship was robust to large and independent *differential* effects of a major unobserved covariate on in-hospital mortality between PCI and CABG, and that removing the most appropriate subclass of patients indicated for PCI (namely, those with STEMI) – thereby restricting analysis to those for whom there isn't clear consensus – actually strengthened the observed relationship.

Ideally, models similar to the one we employ in this study might in the future provide meaningful guidance in complex clinical situations. However, this was not the purpose of the present study. Instead of defining optimal decision-making strategies *per se*, our efforts served as a means for uniform evaluation of hospital-level practice variation under observed decisions. Whether or not the current approach could serve as a “tool” for clinical decision support would need to be carefully and prospectively evaluated, in particular with respect to patient preferences, discount rates (especially the tradeoff between avoiding in-hospital mortality and maximizing chances of long-term survival), numeracy and health literacy. Other key limitations inherent to this study also deserve to be highlighted. For instance, our analysis relies on coding of diagnoses, which can be inaccurate^{32, 33}. Our use of administrative data limited the ability to identify patients who were ineligible for either procedure based on clinical factors. Finally, our analysis did not account for referral practice or readmissions, which represent important directions for future work; in particular, the reasons why hospitals without CABG capabilities had lower rates of discordance and their impact to the interpretation of the hospital-level performance measures need to be further studied.

In summary, the current study strongly suggests that a systematic approach to incorporating comorbidity as part of the decision-making process for coronary revascularization is needed. Our models, based largely on comorbid conditions, identified hospital-level practice variation that explained differences in risk-adjusted in-hospital mortality after revascularization. These findings demand explanation, and should be linked to ongoing quality improvement efforts and refinement of appropriate use criteria^{4, 34, 35}. Additional study is also needed to determine the extent to which patient-centered, provider-related, and hospital/system level factors drive the observed association between variation in revascularization strategy and adverse outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial support for this study was provided in part by the Clinical and Translational Science Collaborative of Cleveland, KL2TR000440 from the National Center for Advancing Translational Sciences component of the National Institutes of Health (NIH) and the NIH roadmap for Medical Research. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing and publishing the report.

References

- Centers for Disease Control and Prevention. Number of all-listed procedures for discharges from short-stay hospitals, by procedure category and age: United States. 2010. URL: http://www.cdc.gov/nchs/data/nhds/4procedures/2010pro4_numberprocedureage.pdf
- Hlatky MA, Boothroyd DB, Bravata DM, Boersma E, Booth J, Brooks MM, Carrie D, Clayton TC, Danchin N, Flather M, Hamm CW, Hueb WA, Kahler J, Kelsey SF, King SB, Kosinski AS, Lopes N, McDonald KM, Rodriguez A, Serruys P, Sigwart U, Stables RH, Owens DK, Pocock SJ. Coronary artery bypass surgery compared with percutaneous coronary interventions for multivessel disease: a collaborative analysis of individual patient data from ten randomised trials. *Lancet*. 2009; 373:1190–7. [PubMed: 19303634]
- Mohr FW, Morice MC, Kappetein AP, Feldman TE, Stahle E, Colombo A, Mack MJ, Holmes DR Jr, Morel MA, Van Dyck N, Houle VM, Dawkins KD, Serruys PW. Coronary artery bypass graft surgery versus percutaneous coronary intervention in patients with three-vessel disease and left main coronary disease: 5-year follow-up of the randomised, clinical SYNTAX trial. *Lancet*. 2013; 381:629–38. [PubMed: 23439102]
- Patel MR, Dehmer GJ, Hirshfeld JW, Smith PK, Spertus JA. ACCF/SCAI/STS/AATS/AHA/ASNC/HFSA/SCCT 2012 Appropriate use criteria for coronary revascularization focused update: a report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, Society for Cardiovascular Angiography and Interventions, Society of Thoracic Surgeons, American Association for Thoracic Surgery, American Heart Association, American Society of Nuclear Cardiology, and the Society of Cardiovascular Computed Tomography. *Journal of the American College of Cardiology*. 2012; 59:857–881. [PubMed: 22296741]
- Jones WS, Patel MR, Holleran SA, Harrison JK, O'Connor CM, Phillips HR 3rd. Trends in the use of diagnostic coronary angiography, percutaneous coronary intervention, and coronary artery bypass graft surgery across North Carolina. *American Heart Journal*. 2011; 162:932–7. [PubMed: 22093211]
- The Dartmouth Institute for Health Policy and Clinical Practice. The Dartmouth Atlas of Health Care. 2015. URL: <http://www.dartmouthatlas.org>
- Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, Lockowandt U. EuroSCORE II. *Eur J Cardiothorac Surg*. 2012; 41:734–44. discussion 744–5. [PubMed: 22378855]
- Anderson RP. First publications from the Society of Thoracic Surgeons National Database. *The Annals of Thoracic Surgery*. 1994; 57:6–7. [PubMed: 8279920]
- Chalmers J, Pullan M, Fabri B, McShane J, Shaw M, Mediratta N, Poullis M. Validation of EuroSCORE II in a modern cohort of patients undergoing cardiac surgery. *European Journal of Cardio-Thoracic Surgery*. 2013; 43:688–694. [PubMed: 22833541]
- Nilsson J, Algotsson L, Høglund P, Luhrs C, Brandt J. Early mortality in coronary bypass surgery: the EuroSCORE versus The Society of Thoracic Surgeons risk algorithm. *The Annals of Thoracic Surgery*. 2004; 77:1235–9. discussion 1239–40. [PubMed: 15063242]
- Barili F, Pacini D, Capo A, Rasovic O, Grossi C, Alamanni F, Di Bartolomeo R, Parolari A. Does EuroSCORE II perform better than its original versions? A multicentre validation study. *European heart journal*. 2013; 34(1):22–29. [PubMed: 23028171]
- Dalton JE, Dawson NV, Sessler DI, Schold JD, Love TE, Kattan MW. Empirical Treatment Effectiveness Models for Binary Outcomes. *Med Decis Making*. 2016; 36:101–14. [PubMed: 25852080]

13. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J Roy Stat Soc: Series B.* 2005; 67:301–320.
14. de Boer SP, Roos-Hesselink JW, van Leeuwen MA, Lenzen MJ, van Geuns RJ, Regar E, van Mieghem NM, van Domburg R, Zijlstra F, Serruys PW, Boersma E. Excess mortality in women compared to men after PCI in STEMI: an analysis of 11,931 patients during 2000–2009. *International Journal of Cardiology.* 2014; 176:456–63. [PubMed: 25127966]
15. Otten AM, Maas AH, Ottervanger JP, Kloosterman A, van 't Hof AW, Dambrink JH, Gosselink AT, Hoorntje JC, Suryapranata H, de Boer MJ, Zwolle Myocardial Infarction study G. Is the difference in outcome between men and women treated by primary percutaneous coronary intervention age dependent? Gender difference in STEMI stratified on age. *European Heart Journal Acute Cardiovascular Care.* 2013; 2:334–41. [PubMed: 24338292]
16. Parolari A, Dainese L, Naliato M, Polvani G, Loardi C, Trezzi M, Fusari M, Beverini C, Tremoli E, Biglioli P, Alamanni F. Do women currently receive the same standard of care in coronary artery bypass graft procedures as men? A propensity analysis. *The Annals of Thoracic Surgery.* 2008; 85:885–90. [PubMed: 18291164]
17. Pendyala LK, Torguson R, Loh JP, Kitabata H, Minha S, Badr S, Dvir D, Barbash IM, Satler LF, Pichard AD, Waksman R. Comparison of adverse outcomes after contemporary percutaneous coronary intervention in women versus men with acute coronary syndrome. *The American Journal of Cardiology.* 2013; 111:1092–8. [PubMed: 23352262]
18. Rogers MA, Blumberg N, Saint SK, Kim C, Nallamotheu BK, Langa KM. Allogeneic blood transfusions explain increased mortality in women after coronary artery bypass graft surgery. *American Heart Journal.* 2006; 152:1028–34. [PubMed: 17161047]
19. Shrestha R, Xu J, Xie D, Liu Z, Xu T, Ye F, Din S, Qian X, Yang S, Liu Y, Li F, Zhang A, Chen S. Comparison of clinical outcomes of Chinese men and women after coronary stenting for coronary artery disease: a multi-center retrospective analysis of 4,334 patients. *Journal of Biomedical Research.* 2014; 28:368–75. [PubMed: 25332708]
20. Kruschke, J. Doing Bayesian data analysis: A tutorial introduction with R. Academic Press; 2010.
21. Ash, AS.; Fienberg, SE.; Louis, TA.; Normand, SLT.; Stukel, TA.; Utts, J. Statistical issues in assessing hospital performance. 2012. URL: <http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Downloads/Statistical-Issues-in-Assessing-Hospital-Performance.pdf>
22. McCulloch, CE.; Neuhaus, JM. Generalized Linear Mixed Models. Wiley Online Library; 2001.
23. Hannan EL, van Ryn M, Burke J, Stone D, Kumar D, Arani D, Pierce W, Rafii S, Sanborn TA, Sharma S, Slater J, DeBuono BA. Access to coronary artery bypass surgery by race/ethnicity and gender among patients who are appropriate for surgery. *Medical Care.* 1999; 37:68–77. [PubMed: 10413394]
24. Rathore SS, Foody JM, Radford MJ, Krumholz HM. Sex differences in use of coronary revascularization in elderly patients after acute myocardial infarction: a tale of two therapies. *Chest.* 2003; 124:2079–2086. [PubMed: 14665483]
25. Ancona C, Agabiti N, Forastiere F, Arca M, Fusco D, Ferro S, Perucci CA. Coronary artery bypass graft surgery: socioeconomic inequalities in access and in 30 day mortality. A population-based study in Rome, Italy. *Journal of Epidemiology and Community Health.* 2000; 54:930–5. [PubMed: 11076990]
26. Taylor F, Ascione R, Rees K, Narayan P, Angelini G. Socioeconomic deprivation is a predictor of poor postoperative cardiovascular outcomes in patients undergoing coronary artery bypass grafting. *Heart.* 2003; 89:1062–1066. [PubMed: 12923028]
27. Lin GA, Dudley RA, Redberg RF. Cardiologists' use of percutaneous coronary interventions for stable coronary artery disease. *Arch Intern Med.* 2007; 167:1604–9. [PubMed: 17698682]
28. Bennett-Guerrero E, Zhao Y, O'Brien SM, et al. Variation in use of blood transfusion in coronary artery bypass graft surgery. *JAMA: the Journal of the American Medical Association.* 2010; 304:1568–1575. [PubMed: 20940382]
29. Dolan JG, Bordley DR, Mushlin AI. An Evaluation of clinicians' subjective prior probability estimates. *Med Dec Making.* 1986; 6:216–223.

30. Kee F, Owen T, Leathem R. Offering a prognosis in lung cancer: when is a team of experts an expert team? *Journal of Epidemiology and Community Health*. 2007; 61:308–313. [PubMed: 17372290]
31. Specht MC, Kattan MW, Gonen M, Fey J, Van Zee KJ. Predicting nonsentinel node status after positive sentinel lymph biopsy for breast cancer: clinicians versus nomogram. *Ann Surg Oncol*. 2005; 12:654–659. [PubMed: 16021535]
32. Goldman LE, Chu PW, Osmond D, Bindman A. The accuracy of present-on-admission reporting in administrative data. *Health Serv Res*. 2011; 46:1946–62. [PubMed: 22092023]
33. Iezzoni LI. Finally present on admission but needs attention. *Medical Care*. 2007; 45:280–282. [PubMed: 17496709]
34. Doll JA, Patel MR. Self-regulation in the Era of Big Data: Appropriate Use of Appropriate Use Criteria. *Ann Intern Med*. 2015; 162:592–3. [PubMed: 25751687]
35. Mohareb MM, Qiu F, Cantor WJ, Kingsbury KJ, Ko DT, Wijeyesundera HC. Validation of the appropriate use criteria for coronary angiography: a cohort study. *Ann Intern Med*. 2015; 162:549–5. [PubMed: 25751586]

563,758 patients undergoing either coronary artery bypass grafting or percutaneous coronary intervention included in six states' AHRQ State Inpatient Datasets (AZ, CA, FL, MD, MI, NJ) between 2009 and 2011

556,356 remaining

554,563 patients included in final analysis (391 hospitals)

Figure 1.
Study flow diagram.

Figure 2.

Discrimination and calibration of predictive models for in-hospital mortality after revascularization, split by sex and actual procedure. Discrimination is characterized by the concordance index, which ranges from 0.5 (equivalent to random guessing) to 1.0 (perfect separation of outcomes). On the plots, perfect calibration is represented as a 45° line where all observed probabilities are equal to all predicted probabilities. Underlying histograms display the distribution of predicted risk.

Figure 3.

Relationship between hospital discordance rates and risk-adjusted (observed-to-expected) in-hospital mortality ratios. Lines and shaded regions indicate estimate and pointwise 95% confidence intervals, respectively; dashed lines indicate pointwise 95% confidence intervals from a model which incorporated restricted cubic splines to allow for potential nonlinearities. **Left panel:** primary model, where expected mortality is the sum of predicted probabilities of mortality for individual patients under the *model-preferred* procedure. **Right panel:** negative control sensitivity analysis, where expected mortality is the sum of predicted probabilities of mortality for individual patients under the *actual* procedure.

Table 1

Summary of model-preferred vs. actual procedure, discordance and in-hospital mortality, by model-recommended procedure and by select patient characteristics that might influence revascularization treatment decisions.

	Percent of Overall Sample (N=554,563)	Model-Recommended Treatment (%)		Actual Treatment (%)		Incidence of Discordance (%)			In-Hospital Mortality (per 1000 patients)			Crude Risk Ratio for Mortality [95% CI]
		CABG	PCI	CABG	PCI	Among All Patients	Among Patients Recommended for CABG	Among Patients Recommended for PCI	Among Patients Given Model-Recommended Treatment	Among Patients Given Discordant Treatment		
All Patients	100	10	90	22	78	28	77	22	7	25	3.7 [3.5, 3.9]	
Recommended	10	100	0	23	77	77	77	–	14	48	3.3 [2.9, 3.9]	
Treatment	90	0	100	22	78	22	–	22	7	17	2.5 [2.4, 2.7]	
Sex												
Male	68	10	90	24	76	29	74	24	5	22	4.2 [3.9, 4.5]	
Female	32	11	89	18	82	25	83	18	10	34	3.3 [3.1, 3.6]	
Age												
75 years	75	9	91	23	77	28	76	23	4	19	4.5 [4.2, 4.8]	
Yes	25	12	88	20	80	27	80	20	15	46	3.1 [2.9, 3.3]	
Type of												
All Emergent	44	14	86	14	86	24	85	14	10	39	4.0 [3.8, 4.2]	
Admission												
Elective Direct	39	5	95	31	69	32	59	30	3	14	4.3 [3.8, 4.8]	
Elective Transfer	18	11	89	24	76	28	71	23	7	26	3.6 [3.2, 4.1]	
Race												
White	70	10	90	22	78	28	77	22	7	25	3.7 [3.5, 3.9]	
Black	7	12	88	18	82	25	81	18	7	30	4.5 [3.8, 5.3]	
Hispanic	11	10	90	22	78	28	77	22	7	24	3.5 [3.0, 4.0]	
Asian	4	9	91	26	74	31	73	26	9	27	3.0 [2.4, 3.8]	
Other	3	10	90	24	76	28	74	24	5	24	4.5 [3.3, 6.1]	
Primary Payer												
Unknown	5	9	91	24	76	28	74	24	7	22	3.2 [2.5, 3.9]	
Medicare	54	11	89	23	77	28	77	23	10	32	3.2 [3.0, 3.4]	
Medicaid	7	12	88	23	77	30	77	23	7	27	4.0 [3.3, 4.8]	
Private	31	8	92	22	78	27	77	22	2	13	5.5 [4.8, 6.3]	
Self Pay/Other	8	11	89	20	80	26	81	20	4	22	5.8 [4.7, 7.1]	
Locale												
Central Metro	37	10	90	21	79	27	79	21	7	27	3.8 [3.5, 4.1]	
Fringe Metro	28	10	90	22	78	28	77	22	6	24	3.8 [3.5, 4.2]	
Medium Metro	19	9	91	25	75	29	75	25	7	26	3.6 [3.3, 4.1]	

	Percent of Overall Sample (N=554,563)	Model-Recommended Treatment (%)			Actual Treatment (%)			Incidence of Discordance (%)			In-Hospital Mortality (per 1000 patients)		Crude Risk Ratio for Mortality [95% CI]
		CABG		PCI	CABG		PCI	Among Patients Recommended for CABG		Among Patients Recommended for PCI	Among Patients Given Model-Recommended Treatment	Among Patients Given Discordant Treatment	
Diabetes	<i>Small/Metro</i>	12	12	88	24	24	76	30	76	24	6	23	3.6 [3.0, 4.4]
	<i>Metropolitan</i>	9	9	91	23	23	77	28	75	23	7	23	3.1 [2.5, 3.8]
	<i>Rural</i>	15	3	85	26	26	74	33	74	26	6	19	2.9 [2.2, 3.9]
	<i>No</i>	9	62	91	20	80	80	26	80	20	7	26	4.0 [3.7, 4.2]
PVD*	<i>Yes</i>	12	38	88	26	74	74	32	74	25	7	24	3.3 [3.0, 3.5]
	<i>No</i>	10	89	90	22	78	78	27	78	22	6	24	3.8 [3.6, 4.0]
	<i>Yes</i>	12	11	88	29	71	71	34	69	29	13	35	2.7 [2.4, 3.0]
	<i>No</i>	10	87	90	21	79	79	27	78	21	6	24	3.9 [3.7, 4.1]
COPD*	<i>Yes</i>	13	13	87	30	70	70	36	72	30	12	33	2.6 [2.4, 2.9]
	<i>No</i>	10	98	90	22	78	78	28	77	22	7	25	3.7 [3.5, 3.9]
	<i>Yes</i>	32	2	68	18	82	82	39	83	18	21	53	2.5 [2.0, 3.1]
	<i>No</i>	9	89	91	22	78	78	26	78	22	6	23	3.9 [3.7, 4.1]
Anemia	<i>Yes</i>	22	11	78	29	71	71	39	73	30	16	37	2.3 [2.1, 2.5]
	<i>No</i>	8	85	92	22	78	78	26	78	22	5	19	3.9 [3.7, 4.1]
	<i>Yes</i>	18	15	82	26	74	74	35	75	26	19	52	2.7 [2.5, 2.9]
	<i>No</i>												

* CABG = coronary artery bypass grafting; PCI = percutaneous coronary intervention; PVD = peripheral vascular disease; COPD = chronic obstructive pulmonary disease; RF/CKD = renal failure/chronic kidney disease.

Table 2

Hospital-level demographic characteristics vs. hospital discordance rate. Results presented as proportions, mean $\pm$ standard deviation, or median [first quartile, third quartile].

	Hospital Discordant Rate		
	20%	>20% and 30%	>30%
	(N = 66,525)	(N = 297,599)	(N = 190,439)
	95 Hospitals	166 Hospitals	130 Hospitals
# patients, 2009–11	437 [224, 854]	1346 [824, 2262]	1319 [826, 2014]
% female	33 $\pm$ 5	32 $\pm$ 4	31 $\pm$ 4
% elderly	24 $\pm$ 8	25 $\pm$ 6	24 $\pm$ 6
% non-white	21 [11, 39]	26 [14, 47]	25 [12, 42]
% Medicare	51 [42, 59]	53 [48, 58]	52 [46, 58]
% Medicaid	6 [4, 10]	6 [4, 10]	6 [4, 10]
% self-pay	9 [6, 13]	8 [5, 11]	8 [5, 12]
% elective transfer	1 [0, 4]	9 [3, 18]	11 [3, 19]
% emergent	75 [58, 83]	54 [40, 63]	52 [41, 61]
% small metro, micropolitan or rural	2 [1, 6]	1 [1, 6]	5 [1, 21]
% diabetic	34 $\pm$ 6	37 $\pm$ 7	37 $\pm$ 7
% with PVD	8 [6, 10]	9 [7, 12]	10 [8, 12]
% with COPD	12 [8, 16]	12 [9, 15]	13 [10, 16]
% with hypotension	2 [1, 3]	2 [1, 2]	2 [1, 2]
% anemic	10 [8, 13]	11 [8, 13]	10 [9, 13]
% with RF/CKD	34 $\pm$ 6	37 $\pm$ 7	37 $\pm$ 7
Median propensity score	12 [10, 15]	16 [13, 19]	18 [16, 21]
% not performing CABG	75	13	8
% discordant with model recommendation	14 [10, 17]	26 [23, 28]	36 [32, 39]