JOURNAL OF HEALTH CARE FINANCE

Fall 2017

Predicting Costs Associated with Open Heart Surgery Based on Clinical, Administrative and Cost Data

Tze Chiam PhD, Associate Director Research Informatics
Value Institute, Christiana Care Health System, Newark, Delaware

Paul Kolm PhD, Director Biostatistics
Value Institute, Christiana Care Health System, Newark, Delaware

Sangita Godbole, Senior Business Analyst, Section of Cardiology
Christiana Care Health System, Newark, Delaware

Matthew Esham, Director Business Intelligence & Analytics
LUMEDX Corporation, Bellevue, Washington

Richard Caplan PhD, Senior Biostatistician
Value Institute, Christiana Care Health System, Newark, Delaware

Mitchell Fawcett MBA, Senior Systems Engineer
Value Institute, Christiana Care Health System, Newark, Delaware

William Weintraub MD, Director Outcomes Research
MedStar Heart & Vascular Institute, MedStar Washington Hospital Center, Washington, DC

Acknowledgements/Source of Funding
This work is supported by Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number U54-GM104941 (PI: Binder-Macleod).

Abstract

Cardiovascular disease is a leading cause of mortality and morbidity in the United States. Coronary artery bypass graft (CABG) surgery, a leading revascularization procedure to treat coronary artery disease, is a high cost procedure that results in large economic impact. Ability to predict CABG surgery cost could enable clinicians and administrators to better manage hospital resources. To plan for CABG surgery cost based on individual patient characteristics, this study develops predictive models using clinical, administrative and cost data. We applied semiparametric regression to develop (i) a cost model that consists of pre-operative variables, and (ii) a cost model that consists of pre-, peri- and post-operative variables. Adding perioperative and postoperative variables increased model accuracy by 25%. Statistically significant variables can inform clinicians and administrators to focus on areas for quality and process improvement initiatives. Potential limitation for model adoption is that costing methodology and accounting methods might vary across hospitals.

Keywords:

Predictive model, coronary artery bypass graft, cost, semiparametric, risk factors

1. Background and Significance

Cardiovascular disease is a leading cause of mortality and morbidity in the United States (Sidney et al. 2016). It also places a large economic burden of illness on the society (Eisenberg 2005). Of all the procedures related to cardiovascular disease, coronary artery bypass graft surgery (CABG) is a leading revascularization procedure in the treatment of coronary artery disease. With an estimated direct cost in the US averaging $57,577 (median: $61,445) with a range from $17,731 to $124,221 (Nicholson et al. 2016). CABG results in a large economic impact to society as well as to individual hospitals. In order to plan for appropriate level of hospital care and facilitate optimal resource allocation, there have been multiple studies on cost prediction of major cardiovascular conditions and interventions, including CABG based on clinical and non-clinical characteristics.

Several studies suggested that models predicting post-CABG mortality can also be used to predict cost (T. S. Kurki et al. 2001),(T S Kurki, Kataja, and Reich 2002),(Pinna Pintor et al. 2003),(Nilsson et al. 2004). Some other studies developed models specifically to predict costs associated with CABG. Sokolovic et al suggested that pre-operative and intra-operative variables are predictors of costs for patients with open heart surgery (Sokolovic et al. 2002).

Despite studies that had success with cost prediction using either cost-specific or mortality prediction models, other studies concluded that in general, risk stratification models fail to predict hospital cost of cardiac surgery patients. Badreldin et al utilized six preoperative scoring models (EuroSCORE, Parsonnet, Ontario, French, Pons and CABDEAL) and concluded that these models are not reliable at predicting costs of cardiac surgical patients due to low correlation between preoperative risk scores with hospital cost and reimbursement. As a result, the authors do not recommend the use of these models for such purpose (Badreldin et al. 2013). The same study also concluded that cost-prediction models that utilized ICU length-of-stay as a variable for cost prediction *do not* help the resource allocation upon hospital admission as ICU length-of-stay can only be calculated after discharge. The study suggested development of more accurate morbidity scores with appropriate weights may be an option to achieve a more reliable financial risk model.

Another study that examined seven risk stratification scoring systems (EuroSCORE, Cleveland, Parsonnet, Ontario, French, Pons, and CABDEAL) also concluded that none of the common risk stratification models accurately predicted hospital costs when applied to cardiac surgical patients at the study site (Hekmat et al. 2005). This study suggested that a "bottom-up" approach where a predictive model is built based on cost associated with individual patient should be used in financial risk studies instead of a "top-down" approach where total budget was divided into aliquots (Edbrooke and Nightingale 1998) for model development.

A study performed by Celi et al also concluded that scoring systems developed using heterogeneous patient population from various centers universally lack clinically acceptable accuracy at an individual patient level (Celi et al. 2012),(Strand and Flaatten 2008). Rather than using models with good external validity, hospitals should consider the alternative approach to build models for specific patient subsets using one's own local database. In addition, a study performed by Pasquali et al found large variations in cardiac surgeries costs among institutions studied. The authors also suggested institutions that wish to use predictive cost model for cardiac surgeries should develop their own model using local clinical and administrative databases (Pasquali et al. 2011).

2. Objectives

This study aims to utilize lessons learnt from previous studies to develop cost prediction models for patients undergoing CABG surgeries. Since patients' preoperative conditions and postoperative complications can contribute to hospital cost, we develop two models to address both cases. As the goal of this paper is to address cost prediction prior to patient discharge to inform hospital resources allocation, outcomes prediction is addressed in a separate paper. As outlined above, length-of-stay is not used as a variable for this prediction as it is only known after patient is discharged from the hospital, thus it is not relevant for this study.

3. Methods

3.1 Study population and data sources

Adult patients >18 years of age who underwent at least one elective CABG surgery at (de-identified) Health System, between 1/7/2013 to 2/10/2016 are included in the study. Visit data and patient demographics data were extracted from the hospital main data warehouse. Clinical data was extracted from data mart developed by the Heart and Vascular Center. Cost data were obtained by extracting direct cost associated with each visit through the hospital's financial accounting system.

3.2 Costing Method

As we aim to derive a predictive model for costs associated with care provided to the study population, only direct costs are used in the analysis. Direct costs, which include fixed and variable costs, are costs incurred by hospital departments that provide direct patient care, such as actual labor of individual employees within the department or cost of supplies used while providing care. Indirect costs (including overhead costs) such as costs related to administration, security, housekeeping are not included in direct cost.

3.3 Statistical Methods

Because cost data were right skewed, semiparametric regression was used to develop cost models. In this analysis, coefficients are estimated by ordinary least squares regression, but standard errors are estimated by bootstrap (500 replications). Preoperative independent variables used for modeling cost included age centered at the mean, sex, race, body mass index (BMI), body surface area (BSA), diabetes, diabetes and on insulin, family history of CAD, cerebrovascular disease, dyslipidemia, unstable angina, number of diseased vessels, ejection fraction (categorized as ≤ 40%, > 40% or not done), previous cardiovascular (CV) intervention, prior myocardial infarction (MI), preoperative medications that included: ACE/ARB, ADP inhibitor, anticoagulant, GPIIBIII inhibitor, inotropes, nitrates and steroids, dialysis, renal failure, peripheral vascular disease (PVD) and hypertension.

We included all variables in the initial model and applied penalized regression using the least absolute shrinkage and selection operator (LASSO) to select a reduced number of risk factors (Tibshirani 2007). We then included peri- and postoperative variables, operative bleeding and readmission to ICU and assessed model accuracy by the change in R-squared. We assessed model accuracy by examining residuals (observed cost minus predicted cost) within one and two standard deviations of the mean of the residuals (zero). In addition, we calculated the percentage of patients with predicted cost deviations of ± \$500, ± \$1,000, ± \$5,000, and ± \$10,000 from the observed cost. Positively- signed residuals indicate that the predicted cost is underestimating the observed cost; negatively-signed residuals indicate that the predicted cost is overestimating the observed cost.

4. Results

There were a total 1,165 patients who received CABG surgery during the study period with valid cost data. Preoperative characteristics are listed in Table I. The preoperative model was comprised of 7 variables: age, female gender, ejection fraction, preoperative anticoagulation, dialysis, PVD, and renal failure (Table II). Ejection fraction > 40% and “not done” decreased cost relative to EF ≤ 40%. The R-squared for the model was 0.278. Adding readmission to the ICU and reoperative bleeding increased the R-squared by 25%, to 0.369 ($p < 0.001$). Adding the peri- and postoperative variables had the most impact on the coefficients of PVD (16% decrease; from \$2318 to \$1,947), renal failure (23% decrease; from \$37,582 to \$28,858) and ejection fraction not done (26% increase in reduction of cost; from -\$2,849 to -\$3,600), as shown in Table III. For the other variables, the cost decrease impact was 10% or less.

Table I. Patient characteristics

Variable	**Summary statistic**
Age (years)	66 ± 10 (27 – 92)
Female	25.1 (292)
Race	
White	83.8 (976)
Black	11.7 (136)
Hispanic	1.6 (18)
Asian	2.2 (26)
Other	0.8 (9)
BMI	30.9 ± 6.4 (14.9 – 68.3)
BSA	2.0 ± 0.2 (1.3 – 3.0)
Family Hx of CAD	34.1 (397)
Diabetes	42.8 (498)
Insulin	30.3 (151/498)
Hypertension	88.8 (1,035)
Dialysis	3.1 (36)
Previous CV intervention	34.9 (407)
Peripheral vascular disease	16.1 (187)
Cerebrovascular disease	15.9 (185)
Unstable angina	15.6 (182)
Dyslipidemia	55.2 (643)
Number of diseased vessels	
0-1	5.1 (59)
2	18.3 (213)
3	76.6 (893)
Ejection fraction	
≤40%	31.7 (369)
>40%	62.9 (733)
Not done	5.4 (63)
Prior MI	31.0 (361)
Preop meds	
ACE/ARB	42.2 (492)
ADP inhibitor	3.8 (44)
Anticoagulant	40.0 (466)
GPIIBIII inhibitor	2.6 (30)
Inotropes	1.3 (15)
Nitrates	9.8 (114)
Steroids	2.8 (32)
Operative bleeding	2.8 (33)
ICU readmission	1.7 (20)
In-hospital death	3.0 (35)

Continuous variables are expressed as mean ± 1 standard deviation.
Categorical variables are expressed as % (n).

Table II. Preoperative cost model

Variable	Regression coefficient ($)	95% confidence interval ($)	p value
Age (centered)	79.55	1.44 : 140.70	0.026
Female	2,496.52	972.30 : 4,332.05	0.003
Ejection fraction			
≤ 40%	-----------	---------------------	------
>40%	-3,927.07	-5,510.99 : -2,388.84	< 0.001
Not done	-2,848.61	-6,346.62 : 323.11	0.075
Anticoagulant	5,158.05	4,015.84 : 6,686.27	< 0.001
Dialysis	11,995.82	5,557.25 : 20,232.33	0.001
PVD	2,317.69	18.18 : 4,383.11	0.039
Renal failure	37,582.13	23,771.21 : 50.669.61	< 0.001
Constant	25,002.89	23,653.72 : 26,527.88	< 0.001

R-squared = 0.278

Table III. Cost model adding postoperative variables

Variable	**Regression coefficient ($)**	**95% confidence interval ($)**	**p value**
Age (centered)	74.70	3.54 : 126.94	0.015
Female	2,230.11	743.28 : 3,794.90	0.003
Ejection fraction			
≤ 40%	-----------	---------------------	------
>40%	-4,183.56	-5,529.22 : -2,809.78	< 0.001
Not done	-3,599.93	-7,278.98 : -532.82	0.032
Anticoagulant	4,862.23	3,7295.04 : 6,106.21	< 0.001
Dialysis	11,273.57	5,797.99 : 20,047.99	0.001
PVD	1,947.49	-33.05 : 3,767.43	0.047
Renal failure	28,858.07	16,686.90 : 39.130.40	< 0.001
ICU readmission	24,792.30	13,727.08 : 40,112.84	< 0.001
Re-op bleeding	14,081.61	8,472.06 : 20,224.30	< 0.001
Constant	24,796.51	23,543.33 : 26,236.31	< 0.001

R-squared = 0.369

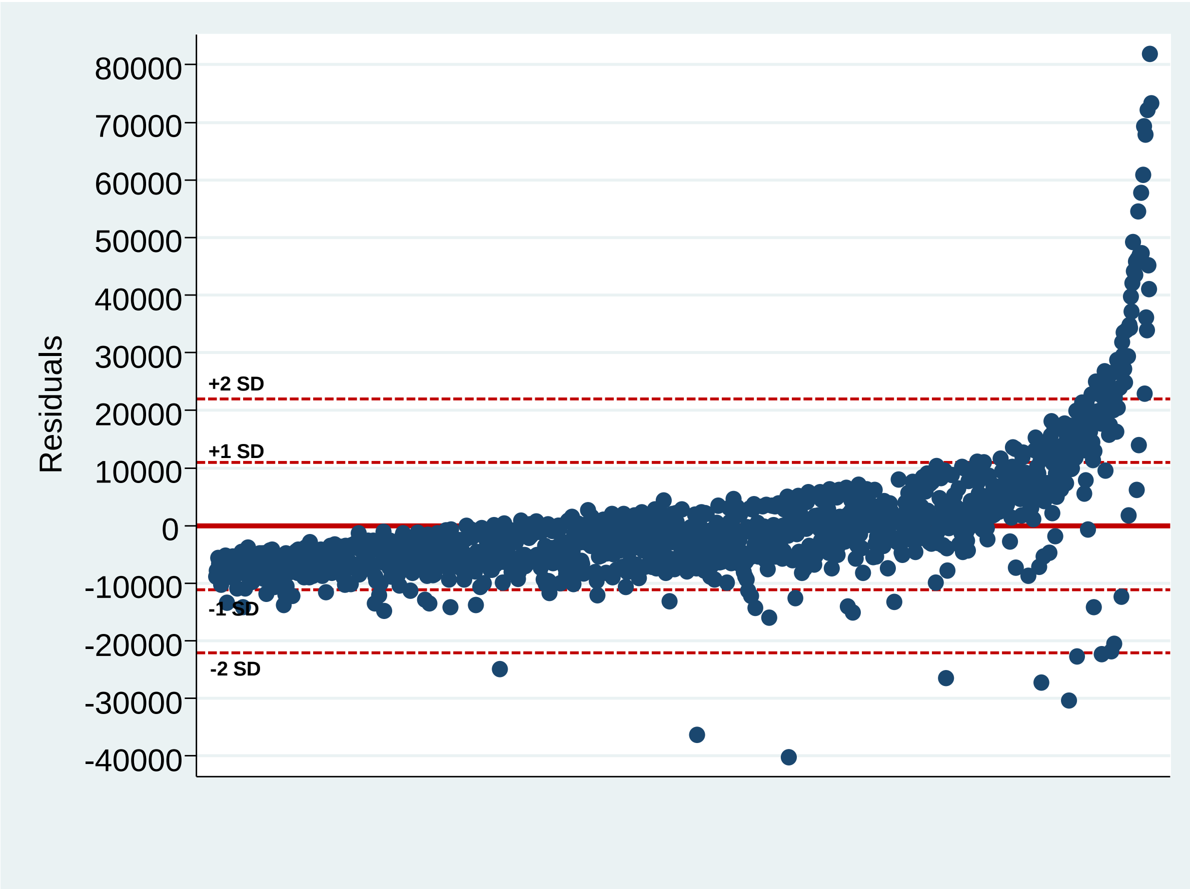

Figure 1. Cost residuals (observed cost minus predicted cost) ordered from lowest to highest cost.

Figure 1 shows residuals ordered from lowest to highest patient cost. The mean of the residuals was virtually zero (1.83×10^{-5}). Generally, lower costs were overestimated by the model whereas higher costs were underestimated by the model. About 5.5% of predicted costs were within $500 of the observed cost, 10% within $1,000, 50% within $5,000 and 85% within $10,000 (approximately 1 standard deviation of residuals).

5. Discussion and Conclusion

Previous studies provided models to predict costs associated with CABG surgery. Some of these studies are multi-center studies that developed universal models. While such models can be used to provide predictions for benchmarking purposes, they often lack applicability for individual hospitals to use on their specific patient populations in order to inform appropriate resource and financial allocation to optimize hospital operations. Some of these studies also utilized length-of-stay as a predictor in the cost prediction model as hospital length-of-stay, particularly ICU length-of-stay is one of the main drivers of total cost associated with CABG surgeries. However, in order for the model to inform hospital resources allocation for improved hospital operations during patient stay and performance improvement efforts, length-of-stay will not be known before patient is discharged. As a result, we developed two cost prediction models to predict costs using only pre-operative variables, as well as using pre- and postoperative variables available prior to patient discharge.

Our analysis shows that the pre-operative variables provide less explanatory power than using both pre- and post-operative variables. The inclusion of postoperative variables increases the fit of the model by 25%. All the significant pre-operative variables remain significant after including postoperative variables. Confidence intervals of each predictor between the two models do not show much variation either. This shows that while the institution utilizing the model can benefit from using both pre- and postoperative variables to make cost predictions, it will still benefit from just using pre-operative variables if collecting data on ICU readmission and post op bleeding poses operational challenges.

We deliberately developed this model using only CABG population from (de-identified) Health System as previous studies have shown that universal models lack applicability to a specific population. Institutions that intend to perform cost predictions should validate and re-calibrate our model (or other models) against their patient populations.

Although this paper focuses on cost instead of clinical outcomes, the significant variables contributing to high costs can also inform clinicians and administrators to focus on specific areas for quality and process improvement initiatives, such as a pre-operative patient optimization clinic or resources needed for ICU readmission for this patient population.

There are several potential limitations to this study. Costing methodology and accounting methods might vary across hospitals. In addition, this study only utilizes total direct costs where some hospitals might find total costs that include direct and indirect costs to be more beneficial for their resource planning. As our model underestimates the costs for cases with higher costs, additional data is needed to further refine the model in order to provide a more conservative approach to resources allocation.

In conclusion, due to the high cost and high volume of CABG surgeries and the need for hospitals to contain cost and optimize resources allocation, we developed two models to predict costs associated with CABG surgeries. We have shown the variables that are significant to both models. We also showed that while the model incorporating pre- and postoperative variables provides higher explanatory power to the prediction, using only pre-operative variables can also provide insights into potential costs prior to patient discharge.

References

Badreldin, Akmal Ma, Fabian Doerr, Axel Kroener, Thorsten Wahlers, and Khosro Hekmat. 2013. “Preoperative Risk Stratification Models Fail to Predict Hospital Cost of Cardiac Surgery Patients.” *Journal of cardiothoracic surgery* 8(1): 126. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3718664&tool=pmcentrez&rendertype=abstract.

Celi, Leo Anthony, Sean Galvin, Guido Davidzon, Joon Lee, Daniel Scott, and Roger Mark. 2012. “A Database-Driven Decision Support System: Customized Mortality Prediction.” *Journal of personalized medicine* 2(4): 138–48.

Edbrooke, D., and P. Nightingale. 1998. “Relationship between TISS and Costs in Intensive Care.” *Intensive Care Medicine* 24(10): 995–96.

Eisenberg, M J. 2005. “Outcomes and Cost of Coronary Artery Bypass Graft Surgery in the United States and Canada.” *Arch. Intern. Med.* 165: 1506–13. http://dx.doi.org/10.1001/archinte.165.13.1506.

Hekmat, K., A. Raabe, A. Kroener, U. Fischer, M. Suedkamp, H. J. Geissler, R. H. Schwinger, S. Kampe, and U. Mehlhorn. 2005. “Risk Stratification Models Fail to Predict Hospital Costs of Cardiac Surgery Patients.” *Zeitschrift fur Kardiologie* 94(11): 748–53.

Kurki, T S, M J Kataja, and D L Reich. 2002. “Validation of a Preoperative Risk Index as a Predictor of Perioperative Morbidity and Hospital Costs in Coronary Artery Bypass Graft Surgery.” *Journal of cardiothoracic and vascular anesthesia* 16(4): 401–4. http://sfx.scholarsportal.info/ottawa?sid=OVID:embase&id=pmid:12154415&id=doi:&issn=1053-0770&isbn=&volume=16&issue=4&spage=401&pages=401-404&date=2002&title=Journal+of+cardiothoracic+and+vascular+anesthesia&atitle=Validation+of+a+preoperative+risk+index+.

Kurki, T. S., U. Hakkinen, J. Lauharanta, J. Rämö, and M. Leijala. 2001. “Evaluation of the Relationship between Preoperative Risk Scores, Postoperative and Total Length of Stays and Hospital Costs in Coronary Bypass Surgery.” *European Journal of Cardio-thoracic Surgery* 20(6): 1183–87.

Nicholson, Gina, Shravanthi R. Gandra, Ronald J. Halbert, Akshara Richhariya, and Robert J. Nordyke. 2016. “Patient-Level Costs of Major Cardiovascular Conditions: A Review of the International Literature.” *ClinicoEconomics and Outcomes Research* 8: 495–506.

Nilsson, Johan, Lars Algotsson, Peter Höglund, Carsten Lührs, and Johan Brandt. 2004. “EuroSCORE Predicts Intensive Care Unit Stay and Costs of Open Heart Surgery.” *Annals of Thoracic Surgery* 78(5): 1528–34.

Pasquali, Sara K, Jie-lena Sun, Phil D’Almada, Robert D B Jaquiss, Andrew J Lodge, Neal Miller, Alex R Kemper, Carole M Lannon, and Jennifer S Li. 2011. “Center Variation in Hospital Costs for Patients Undergoing Congenital Heart Surgery.” *Congenital Heart Disease*.

Pinna Pintor, Plinio, Marco Bobbio, Salvatore Colangelo, Fabrizio Veglia, Roberta Marras, and Marco Diena. 2003. “Can EuroSCORE Predict Direct Costs of Cardiac Surgery?” *European Journal of Cardio-thoracic Surgery* 23(4): 595–98.

Sidney, Stephen, Charles P Quesenberry Jr, Marc G Jaffe, Michael Sorel, Mai N Nguyen-huynh, Lawrence H Kushi, Alan S Go, and Jamal S Rana. 2016. “Recent Trends in Cardiovascular Mortality in the United States and Public Health Goals.” *JAMA Cardiology* 94612(5): E1–6.

Sokolovic, E., D. Schmidlin, E. R. Schmid, M. Turina, C. Ruef, M. Schwenkglenks, and T. D. Szucs. 2002. “Determinants of Costs and Resource Utilization Associated with Open Heart Surgery.” *European Heart Journal* 23(7): 574–78.

Strand, K., and H. Flaatten. 2008. “Severity Scoring in the ICU: A Review.” *Acta Anaesthesiologica Scandinavica* 52(4): 467–78.

Tibshirani, R. 2007. “Regression Shrinkage and Selection via the Lasso Robert Tibshirani.” *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 58(1): 267–88.

Information About the Authors:

Tze Chiam PhD (corresponding author), Associate Director Research Informatics, Value Institute, Christiana Care Health System, Newark, DE, USA. tze.c.chiam@christianacare.org, 302-733-6104

Paul Kolm PhD, Director Biostatistics, Value Institute, Christiana Care Health System, Newark, DE, USA.

Sangita Godbole, Senior Business Analyst, Section of Cardiology, Christiana Care Health System, Newark, DE, USA.

Matthew Esham, Director Business Intelligence & Analytics, LUMEDX Corporation, Bellevue, Washington, USA.

Richard Caplan PhD, Senior Biostatistician, Value Institute, Christiana Care Health System, Newark, DE, USA.

Mitchell Fawcett MBA, Senior Systems Engineer, Value Institute, Christiana Care Health System, Newark, DE, USA.

William Weintraub MD, Director Outcomes Research, MedStar Heart & Vascular Institute, MedStar Washington Hospital Center, Washington, DC, USA.

Clinical Relevance Statement
Due to the high cost and high prevalence of CABG surgery, this paper discusses development of models that can be used to predict cost associated with CABG surgery. Such prediction will enable clinicians and hospital administrations to focus their interventions and performance improvement efforts to reduce cost.

Conflict of Interest Statement
The authors declare they have no conflicts of interests in this research.

Human Subjects Protections
Human and/or animal subjects were not included in this project.

Acknowledgements/Source of funding
This work is supported by Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number U54-GM104941 (PI: Binder-Macleod).