Performance of a Novel Risk Model for Deep Sternal Wound Infection after Coronary Artery Bypass Grafting

Bianca Maria Maglia Orlandi
Harvard Medical School

Omar Asdrubal Vilca Mejia (omar.mejia@incor.usp.br)
Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina do Estado de São Paulo (INCOR)

Jennifer Loría Sorio
University of Costa Rica

Pedro Barros e Silva
Hospital Samaritano Paulista

Marco Antonio Praça Oliveira
Beneficência Portuguesa de São Paulo

Marcelo Arruda Nakazone
Hospital de Base de São José do Rio Preto

Marcos Gradim Tiveron
Irmandade da Santa Casa de Misericórdia de Marília

Valquíria Pelliser Campagnucci
Irmandade da Santa Casa de Misericórdia de São Paulo

Luiz Augusto Ferreira Lisboa
Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina do Estado de São Paulo (INCOR)

Jorge Zubelli
Instituto Nacional de Matemática Pura e Aplicada

Sharon-Lise Normand
Harvard Medical School

Fabio Biscegli Jatene
Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina do Estado de São Paulo (INCOR)

Research Article

Keywords: Prediction models, Validation, CABG, Mediastinitis, Surgical wound infection, Quality improvement

Posted Date: November 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1041348/v1

Performance of a novel risk model for deep sternal wound infection after Coronary Artery Bypass Grafting

Bianca Maria Maglia Orlandi[1,2]; Omar Asdrubal Vilca Mejia[1,3,*]; Jennifer Loría Sorio[4,5]; Pedro de Barros e Silva[3,+]; Marco Antonio Praça de Oliveira[6,+]; Marcelo Arruda Nakazone[7,+]; Marcos Gradim Tiveron[8,+]; Valquíria Pelliser Campagnucci[9,+]; Luiz Augusto Ferreira Lisboa[1]; Jorge Zubelli[5]; Sharon-Lise Normand[2,10]; Fabio Biscegli Jatene[1].

[1]Department of Cardiovascular Surgery, Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina do Estado de São Paulo (INCOR), São Paulo, São Paulo, Brazil.

[2]Department of Health Care Policy, Harvard Medical School, Boston, United States.

[3]Department of Cardiovascular Surgery, Hospital Samaritano Paulista, São Paulo, São Paulo, Brazil.

[4]Universidad de Costa Rica, Costa Rica, America Central.

[5]Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, Brazil.

[6]Department of Cardiovascular Surgery, Beneficência Portuguesa de São Paulo, São Paulo, São Paulo, Brazil.

[7]Department of Cardiovascular Surgery, Hospital de Base de São José do Rio Preto, São José de Rio Preto, São Paulo, Brazil.

[8]Department of Cardiovascular Surgery, Irmandade da Santa Casa de Misericórdia de Marília, Marília, São Paulo, Brazil.

[9]Department of Cardiovascular Surgery, Irmandade da Santa Casa de Misericórdia de São Paulo, São Paulo, São Paulo, Brazil.

[10]Department of Biostatistics, Harvard TH Chan School of Public Health, Boston, United States.

+these authors contributed equally to this work.

*Corresponding author: email: omar.mejia@incor.usp.br.

Abstract

Clinical prediction models for deep sternal wound infections (DSWI) after coronary artery bypass graft (CABG) surgery exist, although they have a poor impact in external validation studies. We developed and validated a new predictive model for 30-day DSWI after CABG (REPINF) and compared it with the Society of Thoracic Surgeons model (STS). The REPINF model was created through a multicenter cohort of adults undergoing CABG surgery (REPLICCAR II Study) database, using least absolute shrinkage and selection operator (LASSO) logistic regression, internally and externally validated comparing discrimination, calibration in-the-large (CL), net reclassification improvement (NRI) and integrated discrimination improvement (IDI), trained between the new model and the STS PredDeep, a validated model for DSWI after cardiac surgery. In the validation data, c-index = 0.83 (95% CI 0.72–0.95). Compared to the STS PredDeep, predictions improved by 6.5% (IDI). However, both STS and REPINF had limited calibration. Different populations require independent scoring systems to achieve the best predictive effect. As the STS, the REPINF external validation across multiple centers it's important to guide healthcare professionals as a quality improvement tool in the prevention of DSWI after CABG surgery.

Keywords: Prediction models, Validation, CABG, Mediastinitis, Surgical wound infection, Quality improvement

Introduction

Although outcomes of cardiovascular surgery have improved over time, the incidence of deep sternal wound infection (DSWI) has varied considerably. Morbidity related to DSWI includes prolonged hospital stay, increased antibiotics use, and, consequently, increased costs. Patients who evolved with DSWI increased the cost of hospitalization by up to 3 times relative to those without DSWI [1]. The incidence of DSWI ranged between 1.3% and 2.4% in 2014 with mortality of 5.4% in the first 30 days following CABG surgery [2].

Prediction models for DSWI exist but may not be generalizable in different geographic settings. Incidence of DSWI in developed countries may be lower, considering their resources, application of best practices to avoid these complications and population characteristics. Moreover, models that include more than one cardiac surgery type may have heterogeneous in case-mix, thereby limiting the discriminant ability [3-7]. Healthcare systems, patients' characteristics, and quality protocol adherence varies widely between institutions.

The Magendanz score, a specific prediction tool for Mediastinitis, was based on data from a single center for which some risk factors were missing, data quality assessment was incomplete, and no external validation in independent samples was conducted, thus impacting generalizability [4,5]. Methodologically, addressing the observational nature of the data due to lack of randomization of patients to institutions and missing information for key patient-level variables are two critical challenges requiring attention [6].

Valid statistical approaches to missing confounder information include multiply imputing the missing information, inverse-probability weighting, or comprehensive sensitivity analyses. The inclusion of more confounders and multiple imputation of missing information should enhance the predictive performance of a model [3,7,8]. We hypothesized that specific clinical characteristics, including pre and intraoperative factors, would be associated with better accuracy to predict DSWI on a multicentric registry. This study aimed to develop and validate a prediction model, the REPINF, using data from the Cardiovascular Surgery Registry of the state of São Paulo, Brazil (REPLICCAR II), and compare with the validated model STS PredDeep.

Methods

The REPLICCAR II study is an observational, multicenter, prospective cohort study (9 hospitals in the state of São Paulo) conducted between August 2017 and June 2019. The Ethical Committee Board of the Heart Institute of the Hospital das Clínicas, Medicine School, University of São Paulo, Brazil approved this study as a sub-analysis of the REPLICCAR project (CAPPesq: 2.507.078). Thus, informed consent was waived due to the analysis of pre-established data logs.

All methods were carried out in accordance with relevant guidelines and regulations. All consecutive patients over 18 years undergoing isolated CABG surgery (first cardiac surgery) constituted the sample. The indications for CABG surgery were according to guidelines [9]. Patients received antibiotic prophylaxis at least one hour before skin incision according to institutional policies.

The variables included in REPLICCAR II were defined using the STS ACSD (Adult Cardiac Surgery Database) collection tool (version 2.9, 2017). Approximately 760 variables were collected preoperatively, intraoperatively, and postoperatively, and included risk factors, clinical and laboratory characteristics, and complications of surgery. The data were collected using a secure web application for building and managing online surveys and databases, the REDCap platform (Research Electronic Data Capture, https://www.project-redcap.org/).

The participating hospitals, their researchers, and data managers participated in meetings and data training before and during the data collection period. Data were audited twice by the REPLICCAR II team to evaluate the accuracy and validity of the information collected by the trained data managers [10].

A trained surgical clinical nurse reviewed the infection criteria and definitions following the infection control surveillance system (Standard CDC National Healthcare Safety Network definitions following the National Healthcare Safety Network - NHSN) [11]. All infections involving the subcutaneous tissue to the mediastinum within 30 days following CABG surgery were considered DSWI. This involves fascia and muscle layers as well as organs, spaces and/or deep soft tissues. “Mediastinitis” refers to an infection of the mediastinum, which can

be caused by different etiologies, including DSWI following sternotomy [11]. Patients who have a fascia or muscle affected by an infection during hospitalization often receive surgical wound debridement, antibiotics and negative-pressure wound therapy (vacuum-assisted closure) to prevent mediastinitis.

Approach

1. Confounders and predictors

First, we eliminated all variables missing in more than 30% of the patients because the imputed values would be driven by the imputation model. We next identified variables related to incidence of DSWI in the scientific literature and found 160 variables in our database. Of these, 55 variables with statistical association or clinical significance for DSWI were considered as predictors (supplementary table 1).

2. Treating missing data with multivariate imputation

We used chained equations (MICE) to impute missing data and created 10 imputed datasets. Sample distribution was captured with histograms and descriptive statistics.

3. Statistical analysis

The training sample was created using the REPLICCAR II database that included 4,085 patients. Information from an additional 498 patients from a different set of hospitals was assembled to create an external validation set (from 2015 to 2016). The model development for variables selection and regularization was performed with the least absolute shrinkage and selection operator (LASSO) logistic regression 10-fold cross-validation, to enhance the prediction accuracy and interpretability of the statistical model produced.

We calculated the area under the receiver operating characteristic curve (*c-index*) to evaluate the discriminatory performance of the model and calibration in-large (CL) containing the observed and predicted values (ratio of observed/predicted). The discriminative ability was also evaluated by net reclassification improvement (NRI) and integrated discrimination improvement (IDI) [12]. The results were plotted to compare the new model (REPINF) with the STS PredDeep in both the training and validation databases.

We follow the guidelines recommended in the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis statement checklist (TRIPOD) [13].

Results

Nine hospitals started data collection, but 7 hospitals actively participated during the 2 years of the project. After exclusion of the patients from 2 hospitals (n = 53), our final sample size was 4,085 patients undergoing isolated CABG surgery as the first cardiac surgery. The mean age was 63.3 years (95% CI 62.9–63.5) and 74% were male. The mean body mass index (BMI) was 27 kg/m^2 (95% CI 26.9–27.2) and common comorbidities included diabetes (49%), hypertension (88%), dyslipidemia (62%) and previous myocardial infarction (52%). The baseline characteristics and missing percentages are described in supplementary table 2. After excluding variables with more than 30% missing, 5% of all patients had at least one missing variable in the REPLICCAR II database (n = 4,085 and 160 variables).

The incidence of DSWI during 30 days from surgery was 2.47% (n = 101). We observed 104 deaths, a competing risk for DSWI, within 30 days (3.1%); of these, 3 patients died with DSWI in the period (2.9%). Characteristics between infected and non-infected patients are described in Table 1.

Figure 1. Receiver operating characteristic curve (ROC; c-index) in the external validation sample of the REPINF and STS PredDeep in patients undergoing isolated CABG. Sao Paulo, Brazil, 2017-19.

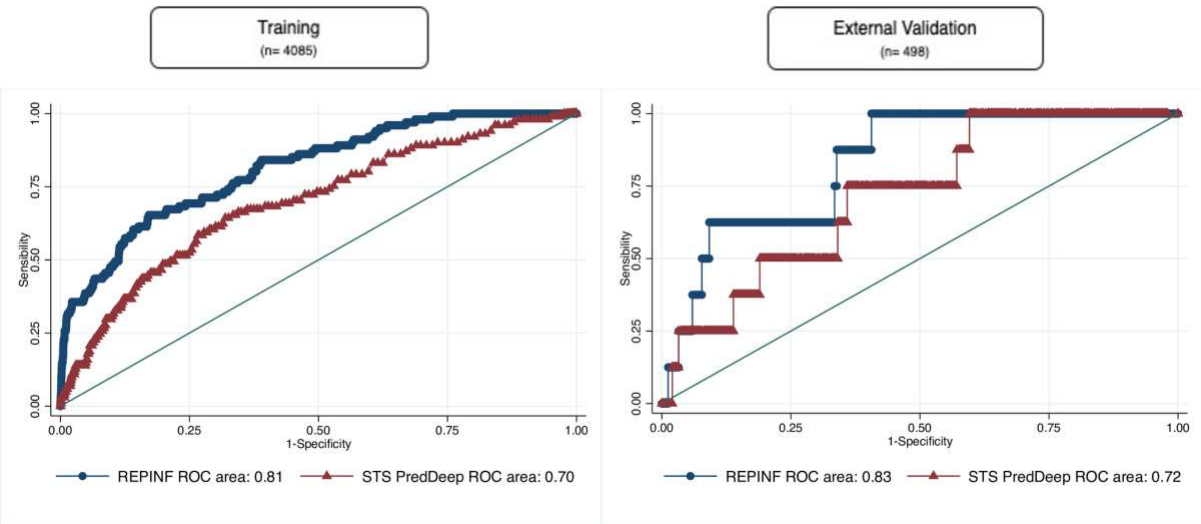


Table 1. Baseline characteristics of patients undergoing isolated CABG surgery with and without DSWI (n = 4,085). REPLICCAR II, São Paulo, Brazil, 2017-19.

	DSWI			
	Yes (n = 101)		No (n = 3,984)	
	n	%	n	%
Age (years)*	63.6 ± 9.5		63.2 ± 9.2	
Gender Male	50	49.5	2,984	74.9
BMI (kg/m^2)*	29 ± 5.5		27.4 ± 4.3	
Diabetes	72	71.3	1,938	48.6
Hemoglobin (mg/dL)*	12.6 ± 1.86		13.5 ± 1.79	
Hematocrit (%)*	38 ± 5.1		40 ± 4.9	
NYHA ≥ III	3	3.0	140	3.5
Three-vessel disease	15	14.9	1,060	26.6
Surgery status				
Elective	54	53.5	2,600	65.3
Urgency	43	42.6	1,371	34.4
Emergency	4	4.0	13	0.3
Lowest intraoperative temperature (ºC)*	33.2 ± 1.9		33.8 ± 1.9	
Surgery duration (hours)*	5.9 ± 1.7		4.7 ± 1.6	
CPB time (minutes)*	87.1 ± 31.6		75.5 ± 29.2	
Anoxia time (minutes)*	69.8 ± 30.8		58.4 ± 24.8	
BITA	19	18.8	450	11.3
Intraoperative high glucose (mg/dL)*	205.1 ± 61.4		179.6 ± 59.9	
Intraoperative blood transfusion	30	29.7	700	17.6

* Mean ± SD; BMI: body mass index; MI: myocardial infarction; NYHA: New York Heart Association; CPB: Cardiopulmonary bypass; BITA: bilateral internal thoracic artery.

Model development

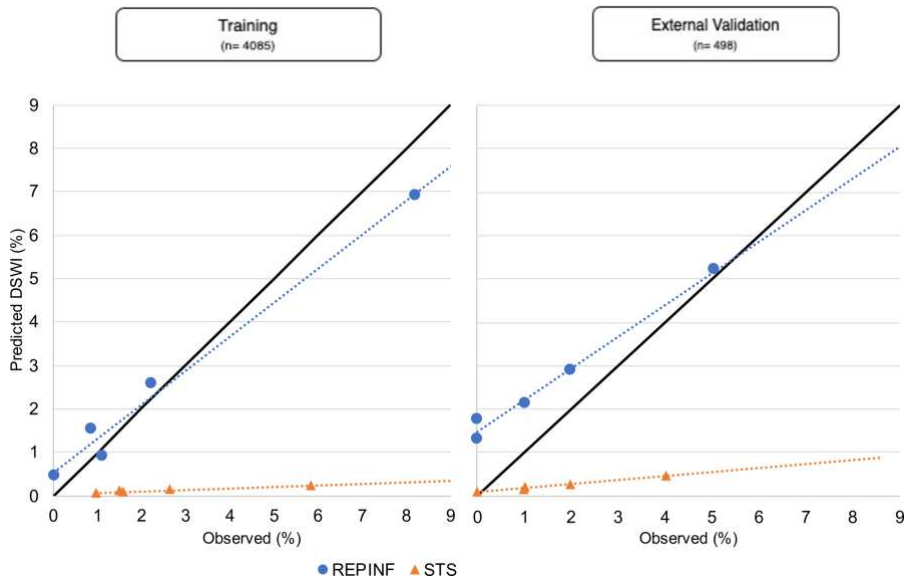
Out of a total 55 of variables related to pre- and intraoperative factors, 7 were included in the Lasso modeling after 10-fold cross-validation (Table 2). In the training sample (n = 4,085), the

REPINF had a *c-index* of 0.81 (95% CI 0.77–0.86) compared to STS PredDeep *c-index* of 0.70 (95% CI 0.64–0.75) (Figure 1). The predicted mean for DSWI was 0.12% (SD = 0.08) using the STS PredDeep. The calibration in-the-large plot (Figure 2) demonstrated that STS predictions tended to underestimate the DSWI risk in our sample.

Table 2. LASSO logistic regression 10-fold cross-validation coefficients. REPINF, REPLICCAR II, São Paulo, Brazil, 2017-19.

Covariates	Coefficients	Logistic Regression Standard error
Female gender	0.246	0.267
Body mass index	0.041	0.025
Diabetes	0.134	0.279
Hemoglobin	-0.182	0.236
Surgery emergency status	0.132	0.793
Surgery duration	0.433	0.091
Bilateral internal thoracic artery used	0.020	0.308
Constant	-3,851	-

Figure 2. Calibration in-the-large plot on training and external validation. Sao Paulo, Brazil, 2017-19.



External model validation

The validation database included 498 patients undergoing isolated CABG during 2015–2016. The mean age was 61.7 (SD = 9.5), 78.5% were male, 54.6% had diabetes and 88% had hypertension. The incidence of DSWI was 1.61%, the predicted STS risk was 0.24% (SD= 0.15) for DSWI and 2.65% (SD= 1.58) for REPINF. The *c-index* was 0.83 (95% CI 0.72–0.95) and 0.72 (95% CI 0.56–0.88) for REPINF and STS in the external validation sample, respectively. Relative to the STS PredDeep, REPINF demonstrated improved classification with a net improvement (NRI) of 29% (Table 3) and the IDI was 0.065.

Table 3. Reclassification data table with quintiles for net reclassification improvement (NRI). REPINF, REPLICCAR II, São Paulo, Brazil, 2017-19.

Event		REPINF						
		Quintile	1	2	3	4	5	Total
No	STS	1	379	217	138	67	36	837
		2	223	221	175	137	63	819
		3	128	181	172	184	123	788
		4	71	132	203	187	186	779
		5	16	57	122	224	342	761
			817	808	810	799	750	3984
Yes	STS	1	0	2	3	1	2	8
		2	0	3	0	4	6	13
		3	0	2	2	2	6	12
		4	0	1	0	4	16	21
		5	0	1	2	7	37	47
	Total		0	9	7	18	67	101

Discussion

The development of prognostic models that combine patient characteristics, risk profiles, and surgical practice to produce predictions about future outcomes allow informed clinical decision-making [8]. Risk models can be used for quality measurement, clinical practice improvement, voluntary public reporting, and research.

In the STS database the incidence of DSWI across all cardiac procedures is low (less than 0.5%). The model systematically underestimates the risk of infection in CABG patients, possibly due to the low rate of this complication for this type of procedure, yielding a c-index of 0.68 for CABG surgery patients [3].

Differences in health management and performance across countries may also affect model discrimination. The MagedanzSCORE[4] (2010) was created using information from a single Brazilian institution among adults (n = 2,809) undergoing isolated CABG and valve surgery. The score was developed and validated in the same population and thus provides overly optimistic model performance metrics [5].

In this study, the incidence of DSWI was 2.47% (30-day follow-up) and the prognostic model was restricted to patients undergoing primary isolated CABG. Of 55 variables candidates in REPINF, 7 emerged from the LASSO logistic regression: female gender [14-16], BMI [14,17-19], diabetes [1,17,18,19], hemoglobin [18], emergency surgery status [18,20,21], surgery duration [18,21] and bilateral internal thoracic artery (BITA) used [22,23,24]. All variables included in the LASSO regression in the new model have already been described as risk factors for DSWI. In fact, the REPINF model included intraoperative variables already described for DSWI endpoint, as surgical timing [18,21] and use of BITA[22,23,24].

A recent prospective multicentric study with 16 centers of cardiac surgery in 6 European countries (England, Finland, France, Germany, Italy and Sweden) reported an incidence of DSWI of 2.5% and the following independent predictors: female gender, BMI ≥ 30kg/m^2, estimated glomerular filtration rate <45mL/min/1.73m^2, diabetes, chronic lung disease, preoperative atrial fibrillation, critical preoperative state and BITA grafting. The model achieved a better discrimination compared to the common scores used (Alfred hospital risk

index, Friedman score and the Brompton-Harefield infection score). The IDI improvement ranged from 1.2% to 2.1% compared to these scores, but no comparison to the validated model for cardiac surgery, the STS, was undertaken [24].

A single ‘‘calibrated model’’ to make predictions across patients undergoing many different surgery types is challenging [25,26]. Our model achieved better accuracy than the STS having *c-indices* of 0.83 (REPINF) and 0.72 for STS in the validation cohort. It is important to note that the external validation database corresponds to an active participant in STS reports since 2014, which is center that is likely not representative of all Brazilian hospitals. The REPINF demonstrated better discrimination, an improvement 29% NRI and IDI improvement of 6.5% compared to STS model. The STS PredDeep systematically underestimated DSWI risk in both the internal and external validation datasets (Calibration: Figure 2).

Our score overestimated risk in external validation cohort for those at the lower end of the risk scale, however, we see that this may have happened because for the elaboration of the REPINF the largest volume of patients came from the public health system. For validation, the REPINF was evaluated in a specific population of private network patients. This may have influenced REPINF to overestimate the risk of infection in the validation sample.

Calibration is an important aspect in models constructed for predictive purposes. It is necessary to keep data collection guided by rigorous quality registries and criteria to achieve and maintain the best accuracy in predictions, considering that this information may be often contaminated by noise [26]. To improve calibration, risk scores should be adjusted for the case-mix of hospitals, with recalibration or remodeling being recommended [27]. Models’ performance may be increased by adding more variables and optimizing estimates of improvement, but at the same time can cause overfitting. REPINF model was created considering these situations, where all variables included for LASSO regression were associated with DSWI creating a difficult task for the variable selection. Before LASSO, the stepwise approach is the most widespread method for choosing covariates. LASSO was originally formulated for linear regression, and it’s applied in statistics and machine learning for variable selection and regularization. Also, LASSO improves prediction error by shrinking

the sum of the squares of the regression coefficients to be less than a fixed value to reduce overfitting [26,28].

Accurate information is essential to access patient's prognosis, which simultaneously considers a number of factors and provides an estimate of the patient's absolute risk of an event and, for DSWI, it is a great challenge. Clinicians and surgeons need an accurate risk prediction for decision support, quality of care assessment, and patient education. Continuous evaluation of the model performance is important to ascertain that the classification performance does not degrade with time. Some models are redeveloped periodically to adjust for temporal trends [29-32].

We suggest the moment patient arrives to the ICU, a REPINF is estimated and the professional team establishes a clear plan of care for recovery organized according to patient risk, thus minimizing complications and reducing costs and hospital length of stay using specific protocols developed with the infection control team. However, should be performed more investigations to determine cut-offs on risk classification and determine when to apply preventive strategies.

Limitations related to data completeness and accuracy were carefully addressed during quality audit for all institutions. Still some important clinical aspects were not evaluated, as the pedicled or skeletonized harvesting conduits, glycated hemoglobin, albumin, bilirubin, local symptoms and variables related to DSWI treatment (fluid or tissue culture, antibiotics, wound intervention, bandages and others). However, it's possible to include more variables to the model without overfitting using this statistical approach.

Another important issue is related to the DSWI detection method (30-day follow-up), which may vary across institutions [24,25]. In our study, this limitation was controlled by having trained researchers to make contact 30 days after surgery with each patient, with only 5.97% of incomplete follow-up.

In summary, this study considered a structured, standardized approach to model development, and validation to identify factors to help multidisciplinary teams prevent DSWI after CABG. More studies should be performed to validate these findings, but we suggest that

REPINF, as well as the STS prediction models [33], provides the highest generalizability for future data. Thus, it's proven that different populations require independent scoring systems to achieve the best predictive effect.

References

1. Hirahara N, et al. Procedure- and Hospital-Level variation of Deep Sternal Wound Infection from All-Japan Registry. *Ann Thorac Surg*. 2019;**109**(2):547–54.
2. Spartalis, E., *et al.* Results of the modified bi-pectoral muscle flap procedure for post-sternotomy deep wound infection. *Surg Today.* **46**, 460–465 (2016).
3. O'Brien SM, et al. The Society of Thoracic Surgeons 2018 Adult Cardiac Surgery Risk Models: Part 2—Statistical Methods and Results. *Ann Thorac Surg*. 2018;**105**(5):1419–28.
4. Magedanz EH, et al. Elaboração de escore de risco para mediastinite pós-cirurgia de revascularização do miocárdio. *Rev Bras Cir Cardiovasc*. 2010;**25**(2):154–9.
5. Sá MPBDO, et al. Validação do MagedanzSCORE como preditor de mediastinite após cirurgia de revascularização miocárdica. *Rev Bras Cir Cardiovasc*. 2011; **26**(3):386–92.
6. Normand SLT. Some old and some new statistical tools for outcomes research. *Circulation*. 2008;**118**(8):872–84.
7. Yulei H. Missing Data Analysis Using Multiple Imputation: Getting to the Heart of the Matter. *Circ Cardiovasc Qual Outcomes*. 2010;**3**(1):1–16.
8. Austin PC, et al. Effect of variable selection strategy on the performance of prognostic models when using multiple imputation. *Circ Cardiovasc Qual Outcomes*. 2019;**12**(11):1–14.
9. Hillis LD, et al. 2011 ACCF/AHA guideline for coronary artery bypass graft surgery a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2011;**124**(23):652–735.
10. Orlandi BMM, et al. REPLICCAR II study: Data quality audit in the Paulista cardiovascular surgery registry. *PLoS One*. 2020;**15**(7):1–13.

11. Centers for Disease and Control. CDC/NHSN Surveillance Definitions for Specific Types of Infections: Surveillance Definitions. *CDC*. 2021;(January):1–3. Available at: https://www.cdc.gov/nhsn/pdfs/pscmanual/17pscnosinfdef_current.pdf . Accessed May 7, 2021.
12. Pencina M, Agostino R, Agostino Jr. R, Vasan R. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;**27**:157–72.
13. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. *BMC Med*. 2015;**13**(1):1-10.
14. Fowler VG, et al. Clinical predictors of major infections after cardiac surgery. *Circulation*. 2005;**112**(9 suppl.):358–65.
15. Rogers MA, et al. Increased risk of infection and mortality in women after cardiac surgery related to allogeneic blood transfusion. *J Womens Health (Larchmt).* 2007 Dec;**16**(10):1412-20.
16. Ferreira GB, Donadello JCS, Mulinari LA. Healthcare-Associated Infections in a Cardiac Surgery Service in Brazil. *Braz. J. Cardiovasc. Surg*. 2020 Oct; **35**(5): 614-618.
17. Terada T, et al. Severe obesity is associated with increased risk of early complications and extended length of stay following coronary artery bypass grafting surgery. *J Am Heart Assoc*. 2016;**5**(6):9–11.
18. Tiveron MG, et al. Preoperative risk factors for mediastinitis after cardiac surgery: assessment of 2768 patients. *Brazilian J Cardiovasc Surg.* 2012;**27**(2):203–10.
19. Buja A, et al. An update review on risk factors and scales for prediction of deep sternal wound infections. *Int Wound J*. 2012;**9**(4):372–86.
20. Bustamante-Munguira J, et al. A New Surgical Site Infection Risk Score: Infection Risk Index in Cardiac Surgery. *J Clin Med*. 2019;**8**(480):1–12.

21. Berríos-Torres SI, et al. Improved Risk Adjustment in Public Reporting: Coronary Artery Bypass Graft Surgical Site Infections. *Infect Control Hosp Epidemiol*. 2012;**33**(5):463–9.
22. Rubens FD, Chen L, Bourke M. Assessment of the Association of Bilateral Internal Thoracic Artery Skeletonization and Sternal Wound Infection after Coronary Artery Bypass Grafting. *Ann Thorac Surg*. 2016;**101**(5):1677–82.
23. Ohira S, et al. Deep sternal wound infection after bilateral internal thoracic artery grafting: Insights from a Japanese national database. *J Thorac Cardiovasc Surg*. 2019;**157**(1):166-173.e1.
24. Biancari F, et al. Preoperative risk stratification of deep sternal wound infection after coronary surgery. *Infect Control Hosp Epidemiol*. 2020;**41**(4):444–51.
25. Kirmani BH, et al. External validity of the Society of Thoracic Surgeons risk stratification tool for deep sternal wound infection after cardiac surgery in a UK population. *Interact Cardiovasc Thorac Surg*. 2013;**17**(3):479–84.
26. Cook NR. Quantifying the added value of new biomarkers: how and how not. *Diagnostic Progn Res*. 2018;**2**(1):1–7.
27. Ivanov J, Tu JV, Naylor CD. Ready-made, recalibrated, or remodeled? Issues in the use of risk indexes for assessing mortality after coronary artery bypass graft surgery. *Circulation*. 1999;**99**(16):2098-104.
28. Vasquez, MM, *et al.* Least absolute shrinkage and selection operator type methods for the identification of serum biomarkers of overweight and obesity: simulation and application. *BMC Med Res Methodol*. 2016; **16(**154).
29. Moore C, Doherty J. Role of the calibration process in reducing model predictive error. *Water Resour Res*. 2005; **41**(5):1–14.
30. Steyerberg EW, Harrell FE. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol*. 2016; **69**:245–7.
31. Alba AC, et al. Discrimination and calibration of clinical prediction models: Users' guides to the medical literature. *JAMA*. 2017;**318**(14):1377–84.

32. Matheny ME, Ohno-Machado L, Resnic FS. Discrimination and calibration of mortality risk prediction models in interventional cardiology. *J Biomed Inform*. 2005;**38**(5):367–75.
33. Shahian DM, et al. The Society of Thoracic Surgeons 2018 Adult Cardiac Surgery Risk Models: Part 1—Background, Design Considerations, and Model Development. *Ann Thorac Surg*. 2018;**105**(5):1411–8.

Acknowledgements: To the REPLICCAR Study Group: We thank the Harvard Medical School coordinating team (Haley Abing); InCor/HCFMUSP (Evelinda Trindade, MD and grant students: Daniella de L. Pes, Gabrielle Barbosa Borgomoni and Débora Maziero). To partner hospitals and their collaborators: Hospital Samaritano Paulista (Valter Furlan, MD; Nilza Lastra, RN and Mariana Okada); Hospital de Base de São José de Rio Preto (Mariana Pastor, MD); Beneficência Portuguesa (Flavia Cortez, MD, Gilmara Silveira da Silva, MD); Hospital Albert Einstein; Santa Casa de Marília; Santa Casa de São Paulo (Gabriel Mitsumoto, MD). To the Commission and Healthcare-Related Infection Control (Tania MV Strabelli, MD). To the Ministry of Health, FAPESP and PPSUS.

Author contributorship statement

OAVM, FBJ, LAFL and SLN conceived the study. BMMO, OAVM, PBS, MAPO, MAN, MGT, and VPC collected the data. BMMO, JLS and JPZ conducted the analysis. BMMO, OAVM and SLN wrote the paper. All authors participated in revisions and approved the final version of the current manuscript.

Additional information

Funding statement: This study was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP); Competing interests: None declared; Ethics approval: This study was submitted and approved by the Ethics Commission for Analysis of Research Projects (CAPPesq) under number 2016/15163-0. The free and informed consent was dismissed due

to the analysis dealing with pre-established data logs; Patient consent for publication: Not required

Figure legends

Figure 1. Receiver operating characteristic curve (ROC; c-index) in the external validation sample of the REPINF and STS PredDeep in patients undergoing isolated CABG. Sao Paulo, Brazil, 2017-19.

Figure 2. Calibration in-the-large plot on training and external validation. REPINF, Sao Paulo, Brazil, 2017-19.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterial.pdf