JACC: HEART FAILURE
VOL. 4, NO. 9, 2016

ISSN 2213-1779/$36.00
PUBLISHED BY ELSEVIER
http://dx.doi.org/10.1016/j.jchf.2016.04.004

A Bayesian Model to Predict Right Ventricular Failure Following Left Ventricular Assist Device Therapy

Natasha A. Loghmanpour, PhD,[a] Robert L. Kormos, MD,[b] Manreet K. Kanwar, MD,[c] Jeffrey J. Teuteberg, MD,[b] Srinivas Murali, MD,[c] James F. Antaki, PhD[a]

ABSTRACT

OBJECTIVES This study investigates the use of a Bayesian statistical model to address the limited predictive capacity of existing risk scores derived from multivariate analyses. This is based on the hypothesis that it is necessary to consider the interrelationships and conditional probabilities among independent variables to achieve sufficient statistical accuracy.

BACKGROUND Right ventricular failure (RVF) continues to be a major adverse event following left ventricular assist device (LVAD) implantation.

METHODS Data used for this study were derived from 10,909 adult patients from the Inter-Agency Registry for Mechanically Assisted Circulatory Support (INTERMACS) who had a primary LVAD implanted between December 2006 and March 2014. An initial set of 176 pre-implantation variables were considered. RVF post-implant was categorized as acute (<48 h), early (48 h to 14 daysays), and late (>14 days) in onset. For each of these endpoints, a separate tree-augmented naïve Bayes model was constructed using the most predictive variables employing an open source Bayesian inference engine.

RESULTS The acute RVF model consisted of 33 variables including systolic pulmonary artery pressure (PAP), white blood cell count, left ventricular ejection fraction, cardiac index, sodium levels, and lymphocyte percentage. The early RVF model consisted of 34 variables, including systolic PAP, pre-albumin, lactate dehydrogenase level, INTERMACS profile, right ventricular ejection fraction, pro-B-type natriuretic peptide, age, heart rate, tricuspid regurgitation, and body mass index. The late RVF model included 33 variables and was predicted mostly by peripheral vascular resistance, model for end-stage liver disease score, albumin level, lymphocyte percentage, and mean and diastolic PAP. The accuracy of all Bayesian models was between 91% and 97%, with an area under the receiver operator characteristics curve between 0.83 and 0.90, sensitivity of 90%, and specificity between 98% and 99%, significantly outperforming previously published risk scores.

CONCLUSIONS A Bayesian prognostic model of RVF based on the large, multicenter INTERMACS registry provided highly accurate predictions of acute, early, and late RVF on the basis of pre-operative variables. These models may facilitate clinical decision making while screening candidates for LVAD therapy. (J Am Coll Cardiol HF 2016;4:711–21)

From the [a]Department of Biomedical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania; [b]Heart and Vascular Institute, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania; and the [c]Cardiovascular Institute, Allegheny Health Network, Pittsburgh, Pennsylvania. Funding for this work was provided by National Institutes of Health, Division of National Heart, Lung, and Blood Institute grants R41 HL120428 STTR Phase I Cardiac Health Risk Stratification System, R01 HL122639, CORA: a Personalized Cardiac Counselor for Optimal Therapy, and R01 HL086918, Identification and Optimization of Ventricular Recovery of Patients on Ventricular Assistance. Data for this study was provided by the International Registry for Mechanical Circulatory Support (INTERMACS), funded from the National Heart, Lung, and Blood Institute, National Institutes of Health, under Contract No. HHSN268201100025C. Dr. Kanwar has received a research grant through her institution from Thoratec. Dr. Teuteberg has industry and financial relationships with HeartWare, Abiomed, CareDx, Thoratec, and Sunshine Heart. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

Manuscript received August 10, 2015; revised manuscript received April 11, 2016, accepted April 14, 2016.

ABBREVIATIONS AND ACRONYMS

BMI = body mass index

INTERMACS = Inter-Agency Registry for Mechanically Assisted Circulatory Support

LVAD = left ventricular assist device

PAP = pulmonary artery pressure

PVR = peripheral vascular resistance

rfMICE = random forest multiple imputation by chained equation

ROC = receiver operator characteristics

RVF = right ventricular failure

RVFRS = right ventricular failure risk score

SMILE = structural modeling, inference, and learning engine

VAS = visual analog scale

WBC = white blood cell

Left ventricular assist devices (LVADs) are increasingly used for management of patients with end-stage heart failure, both as a bridge to cardiac transplantation and as a destination therapy. Postoperative right ventricular failure (RVF) is known to contribute significantly to post-LVAD morbidity and mortality. The risk of developing RVF after LVAD implantation is multifactorial and dependent on hemodynamic variables such as RV preload and afterload as well as clinical variables such as hepatic and renal function, among others. There have been numerous publications examining risk factors associated with RVF over the past decade that have led to the development of several risk scores for RVF (1–13). These scores consist of weighted sums of 4 to 7 risk factors which do not contribute much sensitivity or specificity. Furthermore, accurate prediction of patients at risk to develop RVF after implantation of a continuous flow LVAD depends on the complex and dynamic interplay of multiple pre-operative variables which cannot be adequately captured by traditional multivariate modeling. In contradistinction, Bayesian network (BN) algorithms can account for nonlinear interactions between variables by identifying groups of risk factors and their conditional interdependency.

SEE PAGE 722

METHODS

We sought to develop a Bayesian-based prognostic model of RVF following implantation of a continuous flow LVAD, using the Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS).

PATIENT COHORT. This study was approved by the INTERMACS Data, Access, Analysis, and Publication Committee. The Data Coordinating Center at University of Alabama at Birmingham provided deidentified patient data for implantations undertaken between December 2006 and June 2014 (n = 10,909) (Online Table 1). The inclusion criteria for this study was the use of a continuous flow LVAD as the primary implant and age ≥18 years. Patients who received a right ventricular assist device (RVAD) were included as long as the initial implant was an LVAD. Total artificial heart recipients were excluded from this study.

RVF DEFINITION. The definition for RVF was based on the INTERMACS definition prior to 2014 (see INTERMACS Appendix A [14]). We studied 3 RVF endpoints: <48 h (acute onset), 48 h to 14 days (early onset), and >14 days (late onset). These endpoints were chosen to stress the fact that different clinical variables affect the risk of RVF at different time points. Therefore, we considered how the model might be used in clinical practice to highlight differences in the associated set of risk factors, which in turn might provide insights into the mitigation of risk. The first endpoint (acute) refers to the immediate intraoperative and postoperative period (<48 h) in which the surgical team might use the information to consider whether or not to implant a temporary or permanent RVAD. The next period (early, i.e., 48 h to 14 days) generally considers the duration from intensive care unit to initial discharge from hospital. The clinical response to this risk might involve pharmacological management of RV function or peripheral vascular resistance (PVR); or possibly the use of a temporary RVAD. Late endpoint (>14 days) would be used to alert the follow-up care provider to be vigilant about the risk of RVF and to be prepared to intervene if necessary. Also, chronic or late RVF (>14 days) is being increasingly recognized as an adverse event that may occur in patients who do not develop acute or early RVF.

MISSING DATA. Data entries in the INTERMACS registry exhibited varying degrees of "missing-ness" (Online Table 2). Data elements with excessive missing data (>90%) were excluded from analysis. Of the pre-implantation variables remaining, both numerical and categorical missing data elements were imputed (5 iterations) using the random forest multiple imputation by chained equation (rfMICE) method, using open source R statistical software (15). The average of all 5 iterations was used for the final imputed value.

DISCRETIZATION. Bayesian methodology requires discretization of continuous variables (Online Table 3). This was performed by assuming a normal, Gaussian distribution and defining breakpoints at 1 standard deviation (SD) above and below the mean, thereby creating 3 ranges, or classes (below average, average, and above average). Lower boundaries were truncated if the breakpoint was <0. In some cases, a fourth range was defined (well above average: if the distribution had a negative, left-sided skew). Examples of discretized intervals include body mass index BMI (≤22.0, 22.0 to 28.7, 28.7 to 35.3, and ≥35.3 kg/m^2), international normalized ratio (INR) (≤0.89, 0.90 to 1.34, 1.35 to 1.80, and ≥1.81), and central venous pressure (≤7.6, 7.7 to 11.4, 11.5