Published in final edited form as:

ASAIO J. 2015 ; 61(3): 313–323. doi:10.1097/MAT.0000000000000209.

A New Bayesian Network-Based Risk Stratification Model for Prediction of Short-term and Long-term LVAD Mortality

Natasha A. Loghmanpour, MS^{*}, Manreet K. Kanwar, MD[†], Marek J. Druzdzal, PhD^{#,‡},
Raymond L. Benza, MD[†], Srinivas Murali, MD[†], and James F. Antaki, PhD^{*}

^{*}Department of Biomedical Engineering, Carnegie Mellon Engineering, Pittsburgh, PA [#]Faculty of Computer Science, Bialystok University of Technology, Bialystok, Poland [†]Cardiovascular Institute, Allegheny General Hospital, Allegheny Health Network, Pittsburgh, PA [‡]Decision Systems Laboratory, School of Information Sciences, University of Pittsburgh, Pittsburgh, PA

Abstract

Existing risk assessment tools for patient selection for left ventricular assist devices (LVADs) such as the Destination Therapy Risk Score (DTRS) and HeartMate II Risk Score (HMRS) have limited predictive ability. This study aims to overcome the limitations of traditional statistical methods by performing the first application of Bayesian analysis to the comprehensive INTERMACS dataset and comparing it to HMRS. We retrospectively analyzed 8,050 continuous flow (CF) LVAD patients and 226 pre-implant variables. We then derived Bayesian models for mortality at each of five time endpoints post-implant (30 day, 90 day, 6 month, 1 year, and 2 year), achieving accuracies of 95, 90, 90, 83, and 78%, Kappa values of 0.43, 0.37, 0.37, 0.45, and 0.43, and area under the ROC of 91, 82, 82, 80 and 81% respectively. This was in comparison to the HMRS with an ROC of 57 and 60% at 90-days and 1-year, respectively. Pre-implant interventions such as dialysis, ECMO, and ventilators were major contributing risk markers. Bayesian models have the ability to reliably represent the complex causal relationships of multiple variables on clinical outcomes. Their potential to develop a reliable risk stratification tool for use in clinical decision making on LVAD patients encourages further investigation.

Keywords

LVAD; Mortality; Bayesian; Risk Assessment; Decision Support System

Corresponding Author: James F. Antaki, Department of Biomedical Engineering, 700 Technology Drive, Carnegie Mellon University, Pittsburgh, PA, 15219, Tel: 412-268-9857, Fax: 412-268-9807, antaki@cmu.edu.

Supplemental Information

Requests for access to the beta version of this interactive decision support tool can be made to the corresponding author.

Disclosures

None of the authors have any financial relationship with any commercial entities that have an interest in the subject of the presented manuscript or other conflicts of interest to disclose.

Introduction

As the prevalence of advanced heart failure (HF) steadily increases,¹ the number of patients requiring consideration for mechanical circulatory support (MCS) is also on the rise. Left ventricular assist devices (LVADs) have been used for nearly 25 years to support patients with advanced HF and improve mortality.^{2,3} In fact, LVADs are now considered standard of care for end stage HF patients, providing alternatives for patients awaiting transplant (Bridge to Transplant – BTT Listed), for those ineligible for transplant (Destination Therapy – DT), or for those patients whose eventual transplant candidacy is unknown (BTT – Likely).

Given the complexity of medical co-morbidities ailing end stage HF patients, careful selection is critical in predicting outcomes. The main goal of this selection is to derive the maximum benefit from this expensive and evolving technology, and keeping the cost per patient life year optimal. Therefore, we need a simple, reliable risk tool that can predict optimal clinical outcomes, quality of life and functional benefit with LVAD therapy in individual HF patients.

Several risk scores and risk factors for both post-op morbidity^{4–6} and mortality^{2,7} have been derived from LVAD population and are currently being used in the clinical setting. Unfortunately, all published risk assessment scores to date are limited by having a narrow scope and inability to generalize across a rather heterogeneous HF population. For example, the commonly cited Destination Therapy Risk Score (DTRS) derived from first generation pulsatile flow pumps demonstrated less than satisfactory performance⁸ in estimating mortality when applied to continuous flow LVAD (CF LVAD) populations.² The more recent risk score derived exclusively for patients receiving the most widely used CF LVAD, HeartMate II, demonstrated only marginal improvement over the DTRS. This HeartMate II Risk Score (HMRS) was able to predict 90-day survival based on five variables with an area under the curve (AUC) of 70%, but only when applied to the same data from which it was derived. Furthermore, the predictors of long term (1 year) survival were limited to only two variables: age and implant center experience.⁷ A recent study to test the validity of the HMRS to a single site demonstrated minimal statistical differences between the survival of the low-, medium- and high-risk cohorts.⁹

We hypothesize that the prevailing conventional scores are limited in scope due to their assumption of linear relationships between significant clinical variables and that these limitations can be overcome by utilizing modern machine learning algorithms. The purpose of this study was to evaluate the efficacy of one such algorithm, the Bayesian network classifier, for predicting mortality at multiple post-implant time points (30 day, 90 day, 6 months, 1 year and 2 years). While the final web-based application of the risk model is a work in progress, this study presents the performance of the predictive Bayesian models, introduced here as the Cardiac Outcomes Risk Assessment (CORA) models, along with its comparison to the HMRS.

Methods

Patient Cohort

The retrospective data for this study was provided by the Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS) database, which in turn was collected under IRB approval by over 150 participating hospitals. Patients or designated surrogates provided written informed consent for their information to be recorded in the registry before the device implant. We excluded data from patients not receiving a CF LVAD as the primary implant, pediatric patients (defined by INTERMACS as age < 19), or those records for which no pre-implant data (within 60 days prior to implant) was available. Patients who also received a Right Ventricular Assist Device (RVAD) were included as long as the primary or initial implant was an LVAD. Data were obtained for 8,050 patients from July 2006 to June 2013. Data from patients whose LVAD was electively removed (e.g., due to transplantation or recovery) were included for post-op outcomes but eventually censored at the time of explant (See Table 1). Medical co-morbidities, such as events or other surgeries during hospitalization and interventions prior to 48 hours of implant, as well as complications were defined using the INTERMACS definitions.

Pre-processing

The initial dataset consisted of 226 pre-implant clinical variables. These were initially screened on the basis of sparseness, wherein variables with greater than 20%, 50%, and 80% missing entries were excluded, leaving for analysis 96, 111 and 184 variables, respectively. Missing values were replaced with either the mean or the mode (for continuous and discrete variables, respectively). Supplemental Table 1 provides the list of variables with their respective formats and input values.

Bayesian Network Models

Bayesian networks (BNs)¹⁰ are classification algorithms in which *nodes* represent variables and *directed edges* (depicted as arrows between nodes) represent influences between those variables. Absence of an arrow between a pair of nodes implies independence between those variables. This allows for significant savings in the number of parameters necessary to represent the complete probability distribution of predictive factors in this complex patient population, making BNs highly practical. In addition to the graph structure, a BN is equipped with conditional probability tables (CPTs) associated with each node, which not only describe the direction of influence amongst variables but allows for representation of the degree of influence.

Consider a simple BN model in Supplemental Figure 1, containing risk factors related to LVAD survival. The percentages in Supplemental Figure 1a correspond to the prevalence of each factor in this example population. The network represents the joint probability distribution of the four clinical factors on survival: age, center experience, albumin, and creatinine contributing to a predicted 2-year mortality of 27%. Now considering a specific patient (Supplemental Figure 1b) over the age of 70 at an inexperienced center for which albumin and creatinine values are unavailable, the model predicts a 58% chance of survival. If albumin and creatinine values were made available for this patient, supplemental Figure

1c demonstrates a reduction in the chance of survival to 41%. By contrast, Supplemental Figure 1d illustrates a much more favorable prognosis (94% chance of survival) using the same BN for a different patient, age 51–60 at an experienced center and normal albumin and creatinine values. This example illustrates the ability of BN models to accommodate incomplete data sets.^{11,12}

The methods used for the present study evolved from our prior experience with machine learning for decision support of optimal VAD weaning,¹³ the need for right ventricular support due to right ventricular failure in LVAD recipients^{14–16} a two-center study to predict 90-day survival for continuous flow LVADs^{17–19} and previous mortality studies using INTERMACS^{20,21}. For this study, we investigated three BN classification algorithms: the Naïve Bayes, Tree-Augmented Naïve Bayes (TAN), and Hill Climber Bayes Net for their unique features, each based on a subset of clinical variables. Naïve Bayes assume that all clinical variables affect the outcome (mortality), but are independent of each other. TAN allows representation of correlations/dependence between the variables, as well as their impact on outcome, represented as multiple arrows. For example, Naïve Bayes could link pre-op INR and albumin to mortality and TAN would take this initial Naïve Bayes structure and then add an arrow between INR and albumin. Hill Chamber Bayes Net²² adds, deletes, and reverses edges (arrows) as it searches through the feature space and terminates when an optimal model structure is achieved.

The subsets of clinical variables were derived using a process called *feature selection*, which reduces the total number of variables to avoid over-fitting the model to the data. In the present study, this was performed by three different evaluators including *Chi-Square* correlation analysis and *information gain*.²³ These processes allow for variables to be ranked based on their predictive power. All BN classifiers were evaluated on an independent dataset comprised of a training set of approximately 90% of the data records and a testing set from the remaining approximate 10%, also known as ten-fold cross validation. Multiple time points were chosen for mortality post-op to demonstrate the ability of BN risk scores to be consistently reproducible and clinically relevant. The final step was to re-rank variables at each mortality end point by sequentially removing those that were least predictive, in order to obtain a subset that was most accurate and relevant for that particular mortality endpoint.

Comparison to HeartMate II Risk Score

The HMRS, which is the only predictive score derived and validated on a cohort of CF LVADs, was derived initially for 90-day survival, but also applied and validated for the 1-year mortality endpoint. In this study, we calculated the 90-day and 1-year mortality stratified by the HMRS for comparison with the Bayesian (CORA) models. The HMRS variables are: center experience (0 if < 15 VADs/year, 1 if ≥ 15 VADs/year), age (per 10 years), albumin, creatinine, and INR. The HMRS equation = $(0.0274 \times [\text{age in years}]) - (0.723 \times [\text{albumin}]) + (0.74 \times [\text{creatinine}]) + (1.136 \times [\text{INR}]) + (0.807 \times [\text{center LVAD volume} < 15])$.⁷

Since the INTERMACS data for age were recorded by decade (e.g., 40–49 or 50–59 years), it was coded by the median value (e.g., 55 for 50–59). The site of LVAD implant for each

patient was also not provided, so we assumed that all centers were considered experienced, which ultimately neutralized the variable. Missing values were replaced by the mean of the associated feature: albumin (844 missing, mean= 3.39), INR (434 missing, mean= 1.34), and creatinine (35 missing, mean= 1.43). There were no missing entries for the age variable. The performance of the HMRS was assessed by the AUC of the receiver operator characteristic (ROC) curve, the Kaplan Meier curves, and the general distribution of predicted low-, medium-, and high risk patients.

Performance Metrics

The Bayesian network (CORA) models were assessed by their accuracy, Kappa statistic, area under the ROC curve (AUC %) and Kaplan Meier curve. Accuracy was defined as the sum of true positive and true negative instances divided by the total number of instances. Standard ROC curves were constructed to capture the relationship between overall sensitivity (true positive rate) and specificity (true negative rate). The area under the ROC curve (AUC %) provided a measure of overall performance. The Kappa statistic measured the agreement of the prediction with the true class (actual outcome), where 1.0 signifies complete agreement and 0 is agreement based on chance.

Comparison of Feature Selection Methods

Univariate statistical analysis was performed to corroborate the statistical significance of the variables that were identified by the BN feature selection. Cox proportional hazards regression model was used to identify statistically significant variables (defined as $p < 0.05$) using MedCalc Statistical Software (version 13.0.6, Ostend, Belgium, 2014) and SAS (version 9.4, Cary, North Carolina).

Survival Analysis

Patient survival estimated by each of the predictive models (BN and HMRS) was plotted as Kaplan-Meier curves for 90-days and 1-year. The latter plots were stratified according to HMRS classification: low-, medium-, high risk. BN plots were stratified according to the prediction of survival at the associated endpoint. Differences between stratified groups were compared using log-rank statistics.²⁴ The log-rank test was used to compare the survival curves, with statistical significance defined as $p < 0.05$. Survival was calculated from the day of implant.

Results

The most predictive variables (according to information gain) for each CORA Bayesian models are provided in Table 2. Variables that were common among all five models across the different time points were: (1) intervention within the last 48 hours, (2) creatinine, (3) events experienced during the hospitalization closest to LVAD implantation, (4) previous cardiac operations, (5) IV inotrope therapy agent, (6) primary diagnosis, (7) hemoglobin, (8) LVAD device strategy, and (9) INTERMACS profile. The specific interventions and events found to be most predictive of mortality were: cardiac arrest, intubation, dialysis, ECMO, feeding tube and IABP. However it is noteworthy that the relative importance of predictive variables differed amongst survival endpoints given their varying influence on outcomes

with passing time. The significant univariate correlations with 90-day and 1-year mortality are reported in Supplemental Table 2 and Supplemental Table 3, respectively. In general, the predictive variables in the CORA Bayesian models were also found to be significant ($p < 0.05$) by the univariate Cox proportional hazard model.

Resulting from model optimization, the best performance resulted from using the *information gain* evaluator, the ranker method for ordering the predicting variables, the TAN model structure (maximum of two arrows directed at each node), and varying variable subset sizes depending on the endpoint (30-day: $n=60$; 90-day: $n=68$; 6-month: $n=80$; 1-year: $n=89$; 2-year: $n=65$). The models were derived using the cutoff point for inclusion of at least 50% completion for each variable, as opposed to the 20% or 80% thresholds. Although the model performances of all three cutoffs for data completeness were comparable, we chose 50% for the final model derivation to preserve the maximal number of clinically relevant variables. A summary of the performance of each model is provided in Table 3, which reports accuracies as large as 96%, AUC of the ROC as large as 89%, and Kappa values as large as 0.47. Figure 1 shows the CORA Bayesian models for (a) 90-days and (b) 1-year. These networks indicate the predicted mortality for an individual exemplary patient, which are 78% at 90-days and 76% at 1-year. The shade of red indicates the sensitivity of a given variable for that particular patient; the thickness of the arrows indicates the strength of influence between two variables. These results predict that the majority of the risk is within the initial three months post-implant and the risk remains relatively constant for the remainder of the year. The exact clinical variable entries are listed in the figure caption.

Supplemental Table 4 summarizes the HMRS risk profile distributions for 90-day and 1-year mortality, with the majority of patients predicted as low risk (93% for both). This overwhelming skew is also depicted in the Figure 2 frequency charts.

The ROC in Figure 3 illustrates a consistent superiority of the CORA Bayesian model predictions for five mortality time end points over the two HMRS time intervals, which were close to the line of unity (ROC curve if solely based on chance). The 90-day and 1-year HMRS stratifications had AUC of 60% and 57%, respectively, whereas the Bayesian 90-day and 1-year predictions exhibited AUC of 81% and 79%, respectively.

Figures 4 and 5 show the Kaplan Meier survival curves stratified by the (a) HMRS and (b) CORA Bayesian models at the 90-day and 1-year endpoints, respectively. Whereas the differences between HMRS risk groups are barely discernable, the CORA Bayesian network models provide clear distinctions in the survival curves.

Discussion

Appropriate patient selection is vital to optimal post-surgical outcomes and cost-effectiveness of LVAD therapy. There is a critical need for accurate, flexible, and improved predictive model to account for the heterogeneity of end stage HF patients and BN analyses can provide the necessary tools to achieve that, as demonstrated in the aforementioned analysis. In contradistinction to the traditional statistical methods, which are comprised of weighted combinations of independent variables, BNs provide the advantages of a rigorous

probabilistic framework in which to perform inference of multiple variables and a visual representation that is easy to interpret by clinicians. These qualities provide a more accurate depiction of human decision-making process and improved performance than the “black box” type risk scores, which can only take into account restricted number of variables that fail to represent the complexity of an end stage HF patient.

The utility of the Bayesian approach was only recognized within the past 25 years,²⁵ with the more recent application of BN based decision support being published in a wide variety of medical disciplines.^{26–31} In 2010, the FDA released a guidance for the use of Bayesian statistics in medical device clinical trials.³² In 2013, United Network for Organ Sharing proposed the adoption of a new Bayesian methodology to better identify those transplant programs that may be underperforming in the area of patient and graft survival.³³

To our knowledge, this is the first application of Bayesian analysis to any LVAD cohort and the first study to report the findings of predictive models derived from both Bayesian and traditional statistical methods (HMRS). The final BN models included both non-modifiable/historical variables (such as implant year) and modifiable variables (such as nutritional assessment, renal function etc.). The former are useful in extrapolating future predictions based on the current trajectory and the latter provide a meaningful way of applying this model prospectively. In the present analysis, there were several variables found to have significant impact on the predicted mortality. These included *clinical* and *non-clinical* variables, both of which play a vital role in decision-making that occurs on a day-to-day basis with these often critically ill patients. An example of a non-clinical variable would be an inability to perform a quality of life questionnaire, due to *patient related reasons* as opposed to *administrative reasons* (as defined by INTERMACS) suggesting poor prognosis and correlating with increased mortality. Not only were certain pre-operative variables associated with higher mortality, having several of these risk factors compounded their impact in an incremental fashion. For example, a patient on dialysis had a predicted 28% risk of death at 90 days post-implant (baseline risk of 10%), which increased to 44% if they were also intubated, and further increased to 66% if they also experienced cardiac arrest. In another example, a patient on ECMO within 48 hours of implant had a 20% risk of death at 90-days, which increased to 31% if they were also on a ventilator and increased even further to 61% if they were additionally on IABP and feeding tube. BN analyses can not only show how clinical variables (e.g. low cardiac output and renal dysfunction) impact the predicted class value (mortality) independently but also analyze how they impact each other (e.g. low cardiac output contributes to renal dysfunction, and thereby mortality). This allows a user to input these various scenarios and calculate the changes in predicted chance of mortality. Traditional risk scores are only able to show how each clinical variable relates to the outcome and not to each other.

To provide a real-life comparison of the BN and HMRS performances, consider the following two real-life, representative case studies. Patient A is a 60 year old Caucasian male, who is INTERMACS level 1 with NYHA class IV symptoms, on ventilator and IABP, with a creatinine level of 2.0 mg/dL. Patient B is a 70 year old Caucasian female who is INTERMACS level 3, with NYHA class IV symptoms, chronic renal disease, with creatinine level of 3.3 mg/dL. Applying the HMRS, patient A is predicted to be at low risk

(8% risk of 90-day mortality), while CORA predicted this patient to have a 44% chance of survival at 90 days. In case B, HMRS this patient to have a medium risk of mortality (11% at 90-days), while CORA predicted the same patient to have a 96% chance of survival at 90 days. In reality, patient A died during their initial hospitalization of multi-organ failure while patient B continued to thrive on pump support after 2 years. These examples demonstrate the improved utility of CORA over HMRS and highlight the already noted shortcomings of HMRS. Also noted in Figure 2, the HMRS exhibited an inherent bias towards identifying patients as low risk. This may stem from the conservative stratification utilizing five variables used to calculate the HMRS. This is in comparison to the 60 – 89 variables included in the CORA models, which comprehensively include: demographics, co-morbidities, hemodynamics, laboratory values, medications and quality of life metrics.

We acknowledge that this study has several important limitations, including extensive missing data pertaining to the independent variables. Although the INTERMACS database is large and representative, it suffers from sparsity of many of the data elements. This prompted us to exclude a large quantity of variables, and compelled us to impute missing values. We elected to use the arithmetic mean, which we understand is more likely to introduce bias than multiple imputation. However, it is more appropriate for data which is not missing at random.³⁴ Our ongoing objective is to continually update the database with prospectively collected data specifically for the prognostic model.

Additional limitations include: uneven distribution of classes, which may impact the accuracy of learning; uneven distribution of many continuous variables and skew of categorical variables; inherent retrospective bias (all patients were already chosen to receive a VAD); and finally, only FDA approved VAD devices were included in registry. Furthermore the preponderance of the INTERMACS data set is derived from HeartMate-II, and therefore might not accurately reflect outcomes of competitive devices. However, despite these limitations, our study does not suffer from other, more common limitations (e.g., single centered) as we utilized the most comprehensive and robust registry currently available for LVAD recipients.

In the current model, we chose five discrete endpoints, although it is possible to include more specificity with respect to time (e.g. 30 day, 45 day, 60 day) the resulting accuracy would degrade. It is ostensibly possible to provide a prediction of mortality as a continuous variable but at the expense of accuracy.

The CORA Bayesian models demonstrated a remarkable improvement over the HMRS with respect to accuracy, sensitivity, specificity, and more realistic Kaplan Meier survival distinctions comparisons. The CORA models consistently outperformed HMRS due to their ability to (1) learn from prior probability, (2) account for relationships between variables, and (3) tolerate missing data elements and allow for blank data entries without loss of continuity. In addition, Bayesian networks are able to more closely reflect the natural clinical decision-making process as compared to traditional risk scores and therefore provide greater confidence as a tool for those making medical decisions. These results encourage continued validation and expansion of the models with a prospective, multicenter study.

Translation to Clinical Practice

Now that the validity of Bayesian analysis in predicting clinical outcomes has been established, the next step is the translation of CORA to clinical practice. This will entail its incorporation into an easily interpretable online application that is accessible at key decision points along the continuum of a patient's clinical course. For this reason, we are currently in the process of programming a demonstration model available to clinicians to solicit feedback to improve both the aesthetics and usability of the interface. The application is intended to integrate with the electronic health record to compute the prognosis (risk of death or adverse events) for an LVAD patient, based on the most current clinical data available. An additional feature will allow LVAD centers to customize the decision support tool according to the unique protocols in their individual programs.

We acknowledge that heart failure is a continuum of diseases, and that a computer decision support system is not necessary to identify the extreme cases for whom LVAD therapy is obviously not appropriate. Hence, the potential utility of the CORA models is to assist the clinical team in decision making with patients for whom the merits or contraindications to an LVAD are not immediately apparent. Accordingly we hope that CORA may promote more judicious use of LVAD therapy: by sparing the questionable patients who would do poorly, and by providing supporting evidence for those patients who are likely to benefit, but might otherwise be denied LVAD therapy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

We would like to thank the Data Access, Analysis, and Publications Committee of INTERMACS for allowing us to use their registry for the study. Models described in this paper have been constructed using GeNIe, a modeling environment for graphical probabilistic models available at <http://genie.sis.pitt.edu/>.

Funding Source

Funding was provided by the 1R41 HL120428-01, 1R01HL086918, and U01HL101066-01 NIH grants.

References

1. Bleumink GS, Knetsch AM, Sturkenboom MCJM, et al. Quantifying the heart failure epidemic: Prevalence, incidence rate, lifetime risk and prognosis of heart failure: The rotterdam study. *European heart journal*. 2004; 25(18):1614–1619. [PubMed: 15351160]
2. Lietz K, Long JW, Kfoury AG, et al. Outcomes of left ventricular assist device implantation as destination therapy in the post-rematch era - implications for patient selection. *Circulation*. 2007; 116(5):497–505. [PubMed: 17638928]
3. Lietz K, Miller LW. Destination therapy: Current results and future promise. *Seminars in thoracic and cardiovascular surgery*. 2008; 20(3):225–233. [PubMed: 19038733]
4. Matthews JC, Koelling TM, Pagani FD, Aaronson KD. The right ventricular failure risk score a pre-operative tool for assessing the risk of right ventricular failure in left ventricular assist device candidates. *J Am Coll Cardiol*. 2008; 51(22):2163–2172. [PubMed: 18510965]
5. Drakos SG, Janicki L, Horne BD, et al. Risk factors predictive of right ventricular failure after left ventricular assist device implantation. *Am J Cardiol*. 2010; 105(7):1030–1035. [PubMed: 20346326]

6. Nakajima I, Kato TS, Komamura K, et al. Pre- and post-operative risk factors associated with cerebrovascular accidents in patients supported by left ventricular assist device. -single center's experience in japan. *Circulation journal : official journal of the Japanese Circulation Society*. 2011; 75(5):1138–1146. [PubMed: 21372403]
7. Cowger J, Sundareswaran K, Rogers JG, et al. Predicting survival in patients receiving continuous flow left ventricular assist device the heartmate ii risk score. *J Am Coll Cardiol*. 2013; 61(3):313–321. [PubMed: 23265328]
8. Teuteberg JJ, Ewald GA, Adamson RM, et al. Risk assessment for continuous flow left ventricular assist devices: Does the destination therapy risk score work? *J Am Coll Cardiol*. 2012; 60(1):44–51. [PubMed: 22541834]
9. Thomas SS, Nahumi N, Han J, et al. Pre-operative mortality risk assessment in patients with continuous flow left ventricular assist devices: Application of the heartmate ii risk score. *The Journal of Heart and Lung Transplantation*. 2014
10. Pearl, J. Probabilistic reasoning in intelligent systems: Networks of plausible inference. Morgan Kaufmann; 1988.
11. Oni ko A, Druzdel MJ. Impact of precision of bayesian network parameters on accuracy of medical diagnostic systems. *Artificial intelligence in medicine*. 2013; 57(3):197–206. [PubMed: 23466438]
12. Onisko, A.; Druzdel, MJ. Impact of bayesian network model structure on the accuracy of medical diagnostic systems. Paper presented at: 13th International Conference on Artificial Intelligence and Soft Computing (ICAISC-2014; Zakopane, Poland. 2014.
13. Santelices LC, Wang Y, Severyn D, et al. Development of a hybrid decision support model for optimal ventricular assist device weaning. *The Annals of thoracic surgery*. 2010; 90(3):713–720. [PubMed: 20732482]
14. Wang Y, Simon MA, Bonde P, et al. Decision tree for adjuvant right ventricular support in patients receiving a left ventricular assist device. *J Heart Lung Transplant*. 2012; 31(2):140–149. [PubMed: 22168963]
15. Wang Y, Simon M, Bonde P, et al. Prognosis of right ventricular failure in patients with left ventricular assist device based on decision tree with smote. *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*. 2012; 16(3):383–390.
16. Wang, Y.; Rose, CP.; Ferreira, A., et al. A classification approach for risk prognosis of patients on mechanical ventricular assistance. *Proceedings of the ... International Conference on Machine Learning and Applications*. International Conference on Machine Learning and Applications; 2010. p. 293-298.
17. Loghmanpour N, Kanwar M, Wang Y, et al. A novel, risk stratification model using bayesian networks for continuous flow left ventricular assist device 90-day survival. *The Journal of Heart and Lung Transplantation*. 2013; 32(4):S232.
18. Loghmanpour NA, Kanwar MK, Wang Y, et al. A novel, risk stratification model using bayesian networks for continuous-flow left ventricular assist device 90-day survival. *J Am Coll Cardiol*. 2013; 61(10_S)
19. Loghmanpour NA, Kanwar MK, Benza RL, Murali S, Antaki JF. Development of predictive models for continuous flow left ventricular assist device patients using bayesian networks. *International Journal of Statistics in Medical Research*. 2014; 3(4):423–434.
20. Loghmanpour NA, Druzdel MJ, Antaki JF. Cardiac health risk stratification system (chris): A bayesian-based decision support system for left ventricular assist device (lvad) therapy. *PloS one*. 2014; 9(11):e111264. [PubMed: 25397576]
21. Loghmanpour NA, Kanwar MK, Bailey SH, et al. Development of a novel predictive model for mortality post continuous flow lvad implant using bayesian networks (bn). *J Heart Lung Transpl*. 2014; 33(4):S74–S74.
22. Cooper GF, Herskovits E. A bayesian method for the induction of probabilistic networks from data. *Machine Learning*. 1992; 9(4):309–347.
23. Kullback S, Leibler RA. On information and sufficiency. *The Annals of Mathematical Statistics*. 1951:79–86.

24. Kaplan EL, Meier P. Nonparametric-estimation from incomplete observations. *J Am Stat Assoc*. 1958; 53(282):457–481.
25. Ashby D. Bayesian statistics in medicine: A 25 year review. *Statistics in medicine*. 2006; 25(21): 3589–3631. [PubMed: 16947924]
26. Hoot, N.; Aronsky, D. Using bayesian networks to predict survival of liver transplant patients. Paper presented at: AMIA annual symposium proceedings; 2005.
27. Lucas, P. Bayesian networks in medicine: A model-based approach to medical decision making. Paper presented at: Proceedings of the EUNITE workshop on Intelligent Systems in patient Care; 2001.
28. Eom J-H, Kim S-C, Zhang B-T. Aptacds-e: A classifier ensemble-based clinical decision support system for cardiovascular disease level prediction. *Expert Systems with Applications*. 2008; 34(4): 2465–2479.
29. Sadeghi S, Barzi A, Sadeghi N, King B. A bayesian model for triage decision support. *International Journal of Medical Informatics*. 2006; 75(5):403–411. [PubMed: 16140572]
30. Gevaert O, De Smet F, Timmerman D, Moreau Y, De Moor B. Predicting the prognosis of breast cancer by integrating clinical and microarray data with bayesian networks. *Bioinformatics*. 2006; 22(14):e184–e190. [PubMed: 16873470]
31. Charitos T, van der Gaag LC, Visscher S, Schurink KA, Lucas PJ. A dynamic bayesian network for diagnosing ventilator-associated pneumonia in icu patients. *Expert Systems with Applications*. 2009; 36(2):1249–1258.
32. Administration FaD. Guidance for the use of bayesian statistics in medical device clinical trials: FDA. Washington, D. C: 2010.
33. Membership and Professional Standards Committee (MPSC). [Accessed April 8, 2014] Proposal to revise the current method for flagging for transplant program post-transplant performance reviews. 2013. http://optn.transplant.hrsa.gov/PublicComment/pubcommentPropSub_327.pdf
34. Lin J-H, Haug PJ. Exploiting missing clinical data in bayesian network modeling for predicting medical problems. *Journal of biomedical informatics*. 2008; 41(1):1–14. [PubMed: 17625974]

Figure 1.

(a) Top: 90-day and (b) Bottom: 1-year Bayesian models case study with a pre-VAD candidate. Relative intensity of red for each node shows the sensitivity of the outcome to the variable, darker correlates to increased sensitivity and lighter correlates to decreased sensitivity (with the order dark red, lighter red, dark grey and finally, white). Arrow thicknesses illustrate the strength-of-influence between two nodes. Patient baseline: age: 72, limtx advanced age: yes, limtx chronic renal disease: yes, limtx frailty: no, limtx limited social support: no, limtx other comorbidities: yes, limtx pulmonary disease: yes, device

strategy: DT, Time since first cardiac diagnosis: >2 years, primary diagnosis: cancer, hemoglobin: 11.9, potassium: 4.5, platelet: 165, WBC: 2.88, pro BNP: 4725, pre-albumin: 17.2, albumin: 3.3, INR: 1.1, creatinine: 1.19, PCWP: 30, mean right atrial pressure: 9, RVEF: 29, LVEDD: 66, LVEF 20–25, aortic regurgitation: none, tricuspid regurgitation: mild, mitral regurgitation: moderate to severe, ECG rhythm: atrial fibrillation, Amiodarone: yes, arrhythmia: no, temporary MCS: no, INTERMACS level 7, NYHA III. All other nodes were unknown or missing and left blank.

Figure 2. Frequency charts stratified by the HMRS risk groups: (a) Top: for 90-day mortality and (b) Bottom: for 1-year mortality.

Figure 3. Receiver operator characteristic curve of 30-day, 90-day, 6-month, 1-year and 2-year Bayesian models and HMRS stratification at 90-day and 1-year compared to the line of unity.

Figure 4. Kaplan Meier survival estimates for the 90-day mortality endpoint stratified by: (a) Top: HMRS risk profile and (b) Bottom: Bayesian model predictions.

Figure 5. Kaplan Meier survival estimates for the 1-year mortality endpoint stratified by: (a) Top: HMRS risk profile and (b) Bottom: Bayesian model predictions.

Table 1
Mortality Statistics Censored for Explant and Transplant

Summary of the patient cohorts used to derive the Bayesian models at each endpoint.

Endpoint	Survival No. (%)	Death No. (%)	Total
30 day	7620 (95.2)	387 (4.8)	8007
90 day	7024 (90.5)	737 (9.5)	7761
6 month	6245 (86.1)	1007 (13.9)	7252
1 year	5241 (79.7)	1334 (20.3)	6575
2 year	4432 (72.7)	1667 (27.3)	6099

Table 2
List of most predictive variables (in order) included in each Bayesian model

INR= international normalized ratio, ICD= implantable cardioverter defibrillator, BNP= B-type natriuretic peptide, WBC= white blood cell, ALT= alanine transaminase, AST= aspartate aminotransferase, LVEF= left ventricle ejection fraction, LVEDD= left ventricle end diastolic diameter, RVEF= right ventricle ejection fraction, PCWP= pulmonary capillary wedge pressure, NYHA= New York Heart Association functional class, QoL= quality of life, BP= blood pressure.

#	30 Day (n=60)	90 Day (n=68)	6 Month (n=80)	1 Year (n=89)	2 Year (n=65)
1	Intervention within the last 48 hours	Intervention within the last 48 hours	Intervention within the last 48 hours	Implant year	Implant year
2	Events this hospitalization	Events this hospitalization	Events this hospitalization	Intervention within the last 48 hours	Cardiac index
3	Previous cardiac operation	Previous cardiac operation	Previous cardiac operation	Events this hospitalization	Events this hospitalization
4	INTERMACS profile	Creatinine	Creatinine	Previous cardiac operation	LVEDD
5	IV inotrope therapy agent	INTERMACS profile	Hemoglobin	Hemoglobin	Intervention within the last 48 hours
6	Creatinine	IV inotrope therapy agent	Device strategy	Creatinine	QoL main occupation
7	AST	Device strategy	INTERMACS profile	LVEDD	ECG rhythm
8	Primary diagnosis	Hemoglobin	IV inotrope therapy agent	Primary diagnosis	Previous cardiac operation
9	Device strategy	Primary diagnosis	Primary diagnosis	INTERMACS profile	Creatinine
10	VO2 max	AST	LVEDD	IV inotrope therapy agent	Mean right atrial pressure
11	6 minute walk	Albumin	Peripheral edema recorded	ECG rhythm	IV inotrope therapy agent
12	Amiodarone	Peripheral edema recorded	ECG rhythm	QoL main occupation	Primary diagnosis
13	Hemoglobin	Platelet	Work status	Device strategy	6 minute walk
14	Albumin	Amiodarone	Albumin	Work status	Hemoglobin
15	Temporary circulatory support	VO2 max	Implant year	Peripheral edema recorded	Pulmonary artery diastolic pressure
16	Platelet	6 minute walk	Mean right atrial pressure	VO2 max	Pulmonary artery systolic pressure
17	Work status	Work status	VO2 max	Mean right atrial pressure	INTERMACS profile
18	ECG rhythm	LVEDD	LVEF	6 minute walk	VO2 max
19	Peripheral edema recorded	ECG rhythm	AST	Cholesterol	Cholesterol
20	ALT	Mean right atrial pressure	6 minute walk	Platelet	Work status

#	30 Day (n=60)	90 Day (n=68)	6 Month (n=80)	1 Year (n=89)	2 Year (n=65)
21	QoL mobility	QoL main occupation	Platelet	Lim for tx listing-chronic renal disease	RVEF
22	PCWP	Lim for tx listing-chronic renal disease	QoL main occupation	Albumin	Aortic regurgitation
23	Tricuspid regurgitation	LVEF	Amiodarone	LVEF	LVEF
24	Mean right atrial pressure	Tricuspid regurgitation	Aldosterone	AST	AST
25	Peak R	Aldosterone	Cholesterol	Aldosterone	Aldosterone
26	Pulmonary artery diastolic pressure	ALT	Lim for tx listing-chronic renal disease	C-Reactive Protein	Lim for tx listing-chronic renal disease
27	Lim for tx listing-chronic renal disease	Diastolic BP	Ascites	Amiodarone	Pro BNP
28	Aldosterone	Temporary circulatory support	C-Reactive Protein	Lim for tx listing-frailty	Albumin
29	INR	INR	Lim for tx listing-peripheral vascular disease	RVEF	PCWP
30	Pre albumin	QoL mobility	Diastolic BP	Diastolic BP	C-Reactive Protein
31	QoL main occupation	Pre albumin	Lim for tx listing-advanced age	NYHA	Lim for tx listing-other comorbidities
32	LVEDD	Ascites	Lim for tx listing-frailty	Aortic regurgitation	Admission reason
33	Trail making status	Cholesterol	Temporary circulatory support	Pulmonary artery systolic pressure	Lim for tx listing-frailty
34	Time since first cardiac diagnosis	NYHA	NYHA	Ascites	Ascites
35	Lim for tx listing-unfavorable mediastinal anatomy	Admission reason	QoL mobility	Admission reason	Amiodarone
36	Sodium	Education level	INR	PCWP	Trail making status
37	Mitral regurgitation	Mitral regurgitation	Lim for tx listing-thoracic aortic disease	QoL mobility	Platelet
38	Diastolic BP	C-Reactive Protein	Arrhythmia	Ace inhibitors	Mitral regurgitation
39	Implant year	Implant year	Tricuspid regurgitation	Trail making status	NYHA
40	QoL self-care	Arrhythmia	Blood type	Education level	Cardiac output
41	Cholesterol	Lim for tx listing-advanced age	Potassium	Pro BNP	Device strategy
42	Cardiac biopsy	Trail making status	Pre albumin	Blood type	Diastolic BP
43	Admission reason	Lim for tx listing-unfavorable mediastinal anatomy	Lim for tx listing-unfavorable mediastinal anatomy	BNP	Pre albumin
44	Aortic regurgitation	Lim for tx listing-other comorbidities	Ace inhibitors	Cardiac output	Tricuspid regurgitation

#	30 Day (n=60)	90 Day (n=68)	6 Month (n=80)	1 Year (n=89)	2 Year (n=65)
45	WBC	VAD indication	RVEF	Temporary circulatory support	Quarter of Implant
46	Lim for tx listing-other comorbidities	QoL self-care	Peak R	VAD indication	QoL mobility
47	C-Reactive Protein	Peak R	Mitral regurgitation	INR	Temporary circulatory support
48	Ascites	WBC	Admission reason	Pulmonary artery diastolic pressure	Administrative reason KCCQ not completed
49	Systolic BP	Lim for tx listing-frailty	ALT	Arrhythmia	Patient complete EuroQoL
50	LVEF	Lim for tx listing-pulmonary disease	Education level	Lim for tx listing-advanced age	Lim for tx listing-peripheral vascular disease
51	VAD indication	Lim for tx listing-chronic infectious concerns	Lim for tx listing-pulmonary disease	Pre albumin	Arrhythmia
52	Potassium	Antiplatelet	VAD indication	Mitral regurgitation	BNP
53	Quarter of Implant	Systolic BP	Trail making status	Lim for tx listing-peripheral vascular disease	ALT
54	QoL pain	Blood type	Pro BNP	Tricuspid regurgitation	Blood type
55	Education level	Time since first cardiac diagnosis	Lim for tx listing-malnutrition cachexia	Lim for tx listing-malnutrition cachexia	Lim for tx listing-unfavorable mediastinal anatomy
56	NYHA	Lim for tx listing-peripheral vascular disease	Pulmonary artery systolic pressure	Peak R	INR
57	Antiplatelet	Aortic regurgitation	QoL self-care	Antiplatelet	Education level
58	Arrhythmia	Pro BNP	BNP	Lim for tx listing-thoracic aortic disease	Age interval
59	Lim for tx listing-immunosuppression	PCWP	PCWP	Lim for tx listing-pulmonary disease	Sodium
60	Heart rate	Visual Analog Status health state	Time since first cardiac diagnosis	Lim for tx listing-limited social support	WBC
61		Potassium	Antiplatelet	Lim for tx listing-unfavorable mediastinal anatomy	QoL self-care
62		RVEF	Lim for tx listing-other comorbidities	Lim for tx listing-other comorbidities	Ace inhibitors
63		Beta blockers	Cardiac output	ALT	Potassium
64		Lim for tx listing-limited social support	Aortic regurgitation	QoL self-care	VAD indication
65		Patient complete EuroQoL	WBC	Quarter of Implant	Time since first cardiac diagnosis
66		BNP	Heart rate	Patient complete EuroQoL	

#	30 Day (n=60)	90 Day (n=68)	6 Month (n=80)	1 Year (n=89)	2 Year (n=65)
67		Cardiac biopsy	Sodium	Time since first cardiac diagnosis	
68		Ace inhibitors	Quarter of Implant	Lim for tx listing-allosensitization	
69			Lim for tx listing-allosensitization	Heart rate	
70			Lim for tx listing-chronic infectious concerns	Beta blockers	
71			Patient complete EuroQoL	Administrative reason KCCQ not completed	
72			QoL activities	Lim for tx listing-recent pulmonary embolus	
73			Lim for tx listing-recent pulmonary embolus	Lim for tx listing-chronic infectious concerns	
74			Lim for tx listing-limited social support	WBC	
75			Age interval	Current ICD	
76			Lim for tx listing-severe diabetes	Lim for tx listing-major stroke	
77			BMI	Potassium	
78			Beta blockers	Currently smoke cigarettes	
79			Systolic BP	Sodium	
80			Visual Analog Status health state	Lim for tx listing-severe diabetes	
81				Lim for tx listing-history of illicit drug use	
82				BMI	
83				Visual Analog Status health state	
84				QoL activities	
85				Age interval	
86				Lim for tx listing-repeated non-compliance	
87				Neseritide	
88				Angiotensin	
89				Lim for tx listing-pulmonary hypertension	

Table 3**Model Performance Summary**

ROC= receiver operator characteristic, ROC %= area under the ROC curve

Endpoint	Accuracy (%)	ROC (%)	Kappa
30 day	96.3	89.4	0.44
90 day	91.4	81.2	0.38
6 month	89.3	80.9	0.47
1 year	84.0	79.4	0.46
2 year	78.4	80.0	0.43