DEVELOPMENT AND UTILIZATION OF BAYESIAN PROGNOSTIC MODELS IN A LEFT VENTRICULAR ASSIST DEVICE DECISION SUPPORT TOOL

by

Lisa Carey Lohmueller

BSE, Johns Hopkins University, 2008

MS, University of Pittsburgh, 2015

Submitted to the Graduate Faculty of

Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Lisa Carey Lohmueller

It was defended on

December 5, 2017

and approved by

James F Antaki, PhD, Department of Biomedical Engineering, Carnegie Mellon University

Rema Padman, PhD, Professor, Management Science & Healthcare Informatics, Carnegie Mellon University

Alan Hirschman, PhD, Professor, Department of Bioengineering, University of Pittsburgh

Dissertation Director: Sanjeev Shroff, PhD, Professor, Department of Bioengineering, University of Pittsburgh

DEVELOPMENT AND UTILIZATION OF BAYESIAN PROGNOSTIC MODELS IN A LEFT VENTRICULAR ASSIST DEVICE DECISION SUPPORT TOOL

Lisa Carey Lohmueller, PhD

University of Pittsburgh, 2018

Heart failure is a chronic, progressive condition that affects over 6 million Americans. The gold standard treatment for advanced heart failure is heart transplant. However, when a donor heart is not available, or the patient is not eligible, patients may receive a mechanical circulatory support device such as a left ventricular assist device (LVAD).

LVADs can improve patient survival and increase patient quality of life but they also require significant changes in lifestyle and carry with them risks of adverse events, such as re-hospitalization, gastrointestinal bleeding (GI), stroke, or right heart failure. LVAD decision making for physicians and patients requires extensive discussion of the trade-off between benefits, risks, and associated lifestyle changes. Decision support tools for patients and their caregivers are in development but are not personalized and are limited to general educational information.

Using Bayesian modeling, a machine learning method of data analysis, I developed novel predictive models for three sets of LVAD outcomes: all-cause mortality, recurrent

gastrointestinal (GI) bleeding, and pump-dependent ischemic stroke. The mortality models performed better than current risk scores with receiver operating characteristic area under the curve (ROC AUC) of 70-71% in a multi-center validation cohort and 76-79% in a contemporary single-center study. The recurrent GI bleeding models performed with ROC AUCs of 68% and 60%, revealed the importance of hemoglobin/hematocrit levels and inflammation in driving risk, and are the first models for this outcome. The ischemic stroke models out-performed the current ischemic risk score with ROC AUCs of 64-66%.

In addition to model development, I explored how to present prognostic information to decision making stakeholders: physicians, patients, and caregivers. I accomplished this with three studies: pilot testing the usability of an online application for physicians, surveying potential LVAD patients' interest in healthcare engagement, and comparing the interpretation of prognostic information in different visual formats between patients and the general population. The results of these studies indicated that survival predictions are the most important outcome in decision making; patient numeracy is a key determinant of decision making engagement; and use of line graphs to present prognostic information is well-suited to all stakeholders.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This work is dedicated to my husband, Jason, and my wonderful family of scientists and academics for their example of passion and intellectual curiosity, my principal investigator, Jim Antaki, for taking a gamble on a student with no data science background, my academic advisor, Sanjeev Shroff, for the support to entirely change my area of study, the patients and their families at the University of Pittsburgh Medical Center, for teaching me the most important lessons about medical decision making, and my baby girl, Andromeda Eve, who motivated me to finish this work.

1.0 INTRODUCTION

1.1 HEART FAILURE ETIOLOGY AND PROGRESSION

Heart failure is a chronic, progressive condition that affects over 6 million Americans. It is characterized by a decline in function of the heart to pump enough blood to perfuse the body. To compensate for the loss of power, the heart may enlarge (cardiac dilation), the muscles of the heart may increase in mass, and/or the heart may pump faster. The vasculature may respond by narrowing blood vessels to increase overall pressure or diverting blood perfusion away from less important tissues. As the heart continues to under-perform, these compensation methods begin to fail. Symptoms resulting from heart failure are fatigue, shortness of breath, and difficulty moving. Patients with heart failure may not know they have the condition until the symptoms begin inferring with activities of daily living [1].

To better describe and then treat heart failure, a system of classification is used called the New York Heart Association (NYHA) Functional Classification. There are four classes of patients, characterized by the functional capability of the patient from Class I – No limitation of physical activity, to Class IV – Unable to carry on any physical activity without discomfort, with symptoms present even when at rest [2].

In early stages (Class I & II), heart failure can be managed with medication, reduced sodium diet, and exercise. However, these solutions do not solve the underlying issue of the

weakened heart and gradually become ineffective over time. When medication and lifestyle changes no longer are effective at managing symptoms, a patient is considered to have advanced heart failure (Class III & IV) [3].

To further delineate the condition of advanced heart failure patients, the International Mechanical Circulatory Support registry (INTERMACS) classifies patients with end-stage HF into seven (7) profiles, with decreasing severity of illness [4]. The correlation between classification systems and their relationship to treatment modalities is presented in Figure 1 [5].

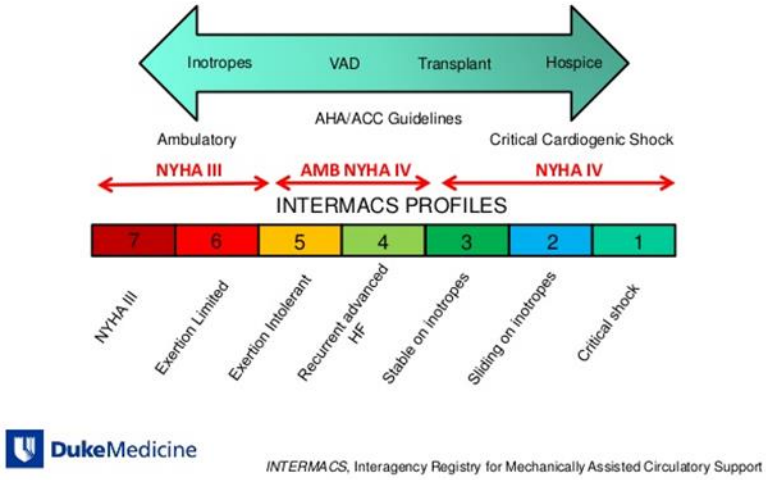


Figure 1. Schematic of NYHA class and INTERMACs profile describing the heart failure treatment continuum [5].

Advanced heart failure treatments range from intravenous drug delivery (inotropes) to surgical interventions (cardiac resynchronization therapy, coronary artery bypass, percutaneous coronary intervention, valve repair or replacement, heart transplant) and device implantation (implantable cardioverter defibrillator, left and/or right ventricular assist device). At the end of the disease progression, palliative care and hospice are often considered. The tumultuous clinical course of disease progression and treatment intensity is depicted in Figure 2 [6].

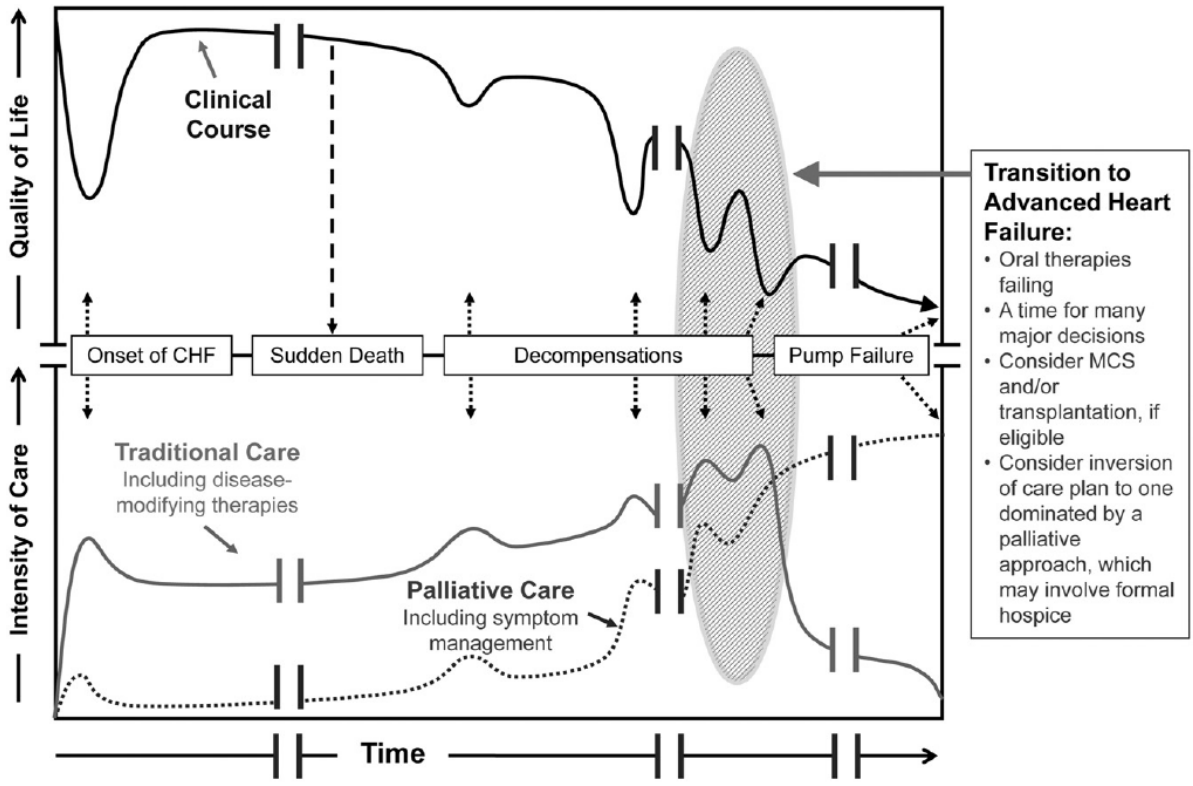


Figure 2. Heart failure disease progression and treatment intensity [6].

The gold standard treatment for end-stage heart failure is a heart transplant. However, the number of hearts available for transplant is far less than the number of patients who need a new heart. Additionally, not all advanced heart failure patients are eligible for heart transplant, due to their age, comorbid conditions, or lifestyle choices. When a heart is not available, or the patient is not eligible, patients may receive a left ventricular assist device (LVAD).

1.2 LEFT VENTRICULAR ASSIST DEVICES

Originally just used as a bridge to transplant, the current generation of LVADs can also be used as a destination therapy for patients who are ineligible for transplant. LVADs can add

years to a patient's life expectancy and increase quality of life [7, 8], but also require significant changes in daily life, investment of time and money, and present a heightened risk of severe adverse events [9].

At the time of this writing, only two FDA approved continuous flow LVAD devices were commercially available in the United States: Heartmate II (Abbott) and HVAD (Medtronic) [10] (see Figure 3.)

The Heartmate II was approved by the FDA in 2008 for bridge to transplant (BTT) and in 2011 for destination therapy (DT) [11]. The HVAD was approved as BTT in 2012 and as a destination therapy in 2017 [12]. (The DT approval occurred after the last data collection for this study, and therefore is only used as BTT in the analyses presented herein.)

Both devices are classified as continuous flow pumps. They are both implanted inside the body and are connected through the skin to an external controller and power system. The HeartMate II has an axial flow impeller and operates at 6000-15000 rotations per minute (rpm) to provide up to 10 liters per minute of blood flow. It is cannulated from the apex of the left ventricle, with outflow to the aorta. The HVAD is a centrifugal flow device that is cannulated directly to the left ventricular apex. It operates at 2000-5000 rpm to provide up to 10 liters per minute [13]. The HVAD design originally had a smooth titanium inflow cannula, but was updated with a sintered inflow cannula in 2011 due to the high rate of pump thrombosis and stroke [14].

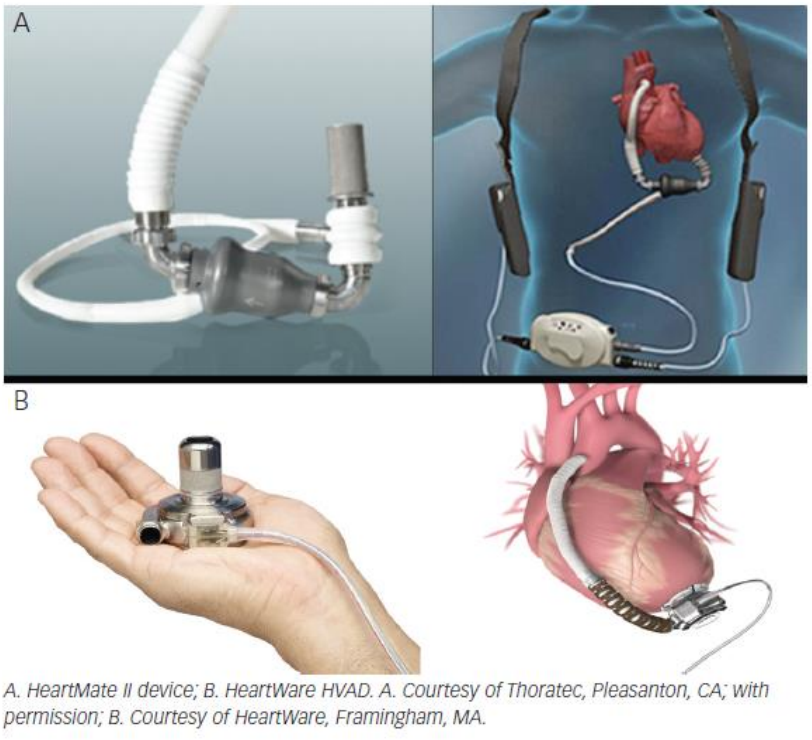


Figure 3. Continuous flow LVADs currently in use: Heartmate II (A) and HVAD (B) [13]

The driveline that exits the skin and the external controller require careful maintenance to avoid life-threatening infection or electrical failure. Both pumps are powered by AC power or batteries that plug into the controller. Therefore patients need to be near electrical outlets and/or carry extra charged batteries at all times [15]. In addition to maintaining the external hardware, patients are recommended to adhere to a low sodium diet. They also face increased risk of having to return to the hospital for pump-associated complications such as driveline infections, controller malfunction, pump thrombosis, or gastrointestinal bleeding [16]. Managing these life changes and risks not only demands the attention of the patient but the support of at least one dedicated caregiver [17].

Due to the trade-off of potentially improved survival and quality of life to change in lifestyle and risk of adverse events, the decision for a physician and patient to have an LVAD implanted requires careful review of educational information and discussion [6]. Accordingly, multidisciplinary heart failure teams must work together to educate patients and their caregivers on LVAD use, risks, and associated lifestyle changes as the patients are discerning their treatment options [18].

1.2.1 Left Ventricular Assist Device Decision Making

The decision to implant an LVAD is a daunting task for both patients and physicians; excess caution may deny a timely, life-saving intervention, but overzealous use may subject patients to significant morbidity, potentially diminishing their quality of life and/or hastening death. The complexity and challenge of clinical decision-making, therefore, lies in identifying the *right patient* who should receive an LVAD at the *right time*.

Most patients that are referred for an LVAD implant are INTERMACS profile I & II. The morbidity and survival when implantation occurs at this point in disease progression is far from satisfactory and greatly impacts the financial and ethical ramifications of the procedure [7, 19]. If a referring physician could identify and refer patients who are both refractory to conventional therapy but not critically ill (INTERMACS 3 or higher), it is hypothesized that LVAD therapy may have a marked improvement in survival and quality of life outcomes [20].

The best practices for the use of an LVAD were recently published as a set of guidelines by The International Society for Heart and Lung Transplantation (ISHLT) [21] and the American Heart Association (AHA) [22]. While these guidelines are an essential first step toward implementation of this technology, they have fundamental limitations:

(1) They are based on consensus and aggregated experience, and therefore cannot be personalized to an individual patient

(2) They do not capture the values and needs of individual patients and their caregivers

(3) They lack effective methods of communicating the risks and benefits of LVAD use to patients, which is essential to achieving shared decision making

These limitations provide the motivation for the current research reported here. It is built on the premise that decision support tools are needed to augment the current guideline-based rationale for determining when and if to implant an LVAD in an advanced heart failure patient.

1.3 FRAMEWORKS FOR DECISION SUPPORT TOOLS

Requirements of an effective and safe clinical decision support tool have been a topic of much interest, particularly with the government incentive to establish and use electronic medical records [23]. For physicians, there is a clinical decision support guidance called the “Five Rights of CDS” (Table 1) [24]. For patients, there is the frequently updated and well-validated International Patient Decision Aid Standards (iPDAS) (Table 2) [25]. These two sets of guidelines define the frameworks to evaluate the existing tools for LVAD decision making.

Table 1. Five Rights of CDS [24]

1. The right **information**,
2. To the right **person**,
3. In the right intervention **format**,
4. Through the right **channel**,
5. At the right time in **workflow**.

Table 2. iPDAS (v3) Checklist (excluding screening-specific requirements)

Dimension / *details*	Item
Information *Providing information about options in sufficient detail for making a specific decision*	1. The decision support technology describes the health condition or problem (intervention, procedure or investigation) for which the index decision is required
	2. The decision support technology describes the decision that needs to be considered (the index decision)
	3. The decision support technology describes the options available for the index decision
	4. The decision support technology describes the natural course of the health condition or problem, if no action is taken.
	5. The decision support technology describes the positive features (benefits or advantages) of each option
	6. The decision aid describes negative features (harms, side effects or disadvantages) of each option.
	7. The decision support technology makes it possible to compare the positive and negative features of the available options.
	8. The decision support technology shows the negative and positive features of options with equal detail (for example using similar fonts, order, and display of statistical information).
Probabilities *Presenting outcome probabilities*	1. The decision support technology provides information about outcome probabilities associated with the options (i.e. the likely consequences of decisions)
	2. The decision support technology specifies the defined group (reference class) of patients for which the outcome probabilities apply.
	3. The decision support technology specifies the event rates for the outcome probabilities (in natural frequencies).
	4. The decision support technology specifies the time period over which the outcome probabilities apply.
	5. The decision support technology allows the user to compare outcome probabilities across options using the same denominator and time period.
	6. The decision support technology provides information about the levels of uncertainty around event or outcome probabilities (e.g. by giving a range or by using phrases such as “our best estimate is…”)
	7. The decision support technology provides more than one way of viewing the probabilities (e.g. words, numbers, and diagrams).
	8. The decision support technology provides balanced information about event or outcome probabilities to limit framing biases.

Table 2 (continued).

Clarifying and expressing values	1. The decision support technology describes the features of options to help patients imagine what it is like to experience the physical effects.
	2. The decision support technology describes the features of options to help patients imagine what it is like to experience the psychological effects.
	3. The decision support technology describes the features of options to help patients imagine what it is like to experience the social effects.
	4. The decision support technology asks patients to think about which positive and negative features of the options matter most to them.
Decision Guidance *Structured guidance in deliberation and communication*	1. The decision support technology provides a step-by-step way to make a decision.
	2. The decision support technology includes tools like worksheets or lists of questions to use when discussing options with a practitioner.
Development *Using a systematic development process*	1. The development process included finding out what clients or patients need to prepare them to discuss a specific decision
	2. The development process included finding out what health professionals need to prepare them to discuss a specific decision with patients
	3. The development process included expert review by clients/patients not involved in producing the decision support technology
	4. The development process included expert review by health professionals not involved in producing the decision aid.
	5. The decision support technology was field tested with patients who were facing the decision.
	6. The decision support technology was field tested with practitioners who counsel patients who face the decision.
Evidence	1. The decision support technology (or associated documentation) provides citations to the studies selected.
	2. The decision support technology (or associated documentation) describes how research evidence was selected or synthesized.
	3. The decision support technology (or associated documentation) provides a production or publication date.
	4. The decision support technology (or associated documentation) provides information about the proposed update policy.
	5. The decision support technology (or associated documentation) describes the quality of the research evidence used.

Table 2 (continued).

Disclosure and transparency	1. The decision support technology (or associated technical documentation) provides information about the funding used for development.
	2. The decision support technology includes author/developer credentials or qualifications.
Plain Language	1. The decision support technology (or associated documentation) reports readability levels (using one or more of the available scales).
DST Evaluation	1. There is evidence that the decision support technology improves the match between the features that matter most to the informed patient and the option that is chosen

1.3.1 Physician-oriented Tools for LVAD Decision Making

The primary decision support tools for physicians evaluating patients for LVAD implant are risk scores that predict post-operative mortality. The most commonly used scores are summarized in Table 3. Of these, only one risk score considers the format of delivery to physicians, the Seattle Heart Failure Model (SHFM), which has an online and downloadable calculator to present the risk of mortality at 1, 2, and 3 years. All other scores must be calculated by physicians on their own. Almost every risk score is derived from a small, clinical trial cohort, except the SFHM which was developed from a cohort of 1,125 ambulatory patients, who were less sick than the patients being considered for LVADs. Performance of these models can be good in their specific patient populations (ROC AUC 0.89 for DTRS) but does not maintain this rate of success with validation in the sickest patients, reflective of the current LVAD candidates (ROC AUC 0.58 for DTRS) [26].

Table 3. Physician-oriented tools for LVAD decision making

Tool	Outcome predicted	Development cohort	Validation cohort and performance	Pro	Con
Heartmate II Risk Score [27]	90-day and 1-year mortality	Derived from HMII BTT and DT trial patients (N = 1,122) randomly divided into derivation cohort (n = 583)	Validation cohort (n = 539) AUC ROC: 0.64 ROC AUC in LVAD patients from ROADMAP trial: 0.71 at 3 months, 0.62 at 12 months[28]	Developed with contemporary device therapy era patients Discriminates survival in sicker, INTERMACS profiles 1-2, patients	Performed less well than the MELD score (0.66) Did not discriminate well in the less sick. [29]
Seattle Heart Failure Model (SHFM) [30]	1-, 2-, and 3-year mortality in patients with moderate HF	Derived in a cohort of 1125 ambulatory heart failure patients	Prospectively validated in 5 additional cohorts, n= 9,942 ambulatory heart failure patients ROC AUC: 0.73 ROC AUC in LVAD patients from ROADMAP trial: 0.71	Applicable for both non-MCS and MCS patient cohort Easily accessible on the web Longest follow up of any prediction scores	Validation cohort was not reflective of advanced HF population for whom MCS is being considered Overestimate survival when used to stratify advanced HF patients[31]
Destination Therapy Risk Score	90-day in-hospital mortality for pulsatile DT LVADs	Derived from patients consenting to be a part of the DT registry for the HeartMate XVE LVAD (n=222)	Original validation ROC AUC: 0.89	Good delineation between high and low risk groups in 1,124 patients enrolled in the HMII DT and BTT trials	Lacks applicability to the newer generation, CF-LVADs devices ROC AUC: 0.54 and 0.58 for the HeartMate II BTT and DT groups, respectively.[26]
Model for End-stage Liver Disease (MELD) for MCD[32]	Correlated with bleeding risk and predicts 6-month mortality	Derived from n=231 patients being considered for transhepatic portosystemic shunts and later for liver transplant candidate categorization	Bleeding validated in single center MCS population, n=211 Mortality validated in INTERMACs registry, n=324	In single center cohort, MELD predicted hospital stay, postoperative device infections, RV failure, and renal failure. In both cohorts, the MELD score was predictive of operative and 6-month mortality.	Limited utility in less ill cohort Not developed for MCS patients Does not include serum creatinine, which can indicate malnourishment and vastly underestimate renal dysfunction

1.3.2 Patient-oriented Tools for LVAD Decision Making

There are two published patient decision support tools in development for LVAD patients, summarized in Table 4. Both tools are funded by grants from the Patient-Centered Outcomes Research Institute (PCORI). They both were developed using the guidelines from iPDAS, derived from single-center feedback and using multi-center clinical trials for validation. The results of both validation trials have not yet been published, so final outcomes in terms of benefit and impact of the tools are currently unknown.

Table 4. Patient-oriented tools for LVAD decision making

Tool	Development cohort	Validation cohort	Components	Pro	Con
Colorado Patient Decision Support, PCORI [33]	Semi-structured interviews with 24 patients, 20 caregivers, and 24 clinicians to assess readability, bias, and usability	DECIDE-LVAD clinical trial of 6 US hospitals, n=168 patient-caregiver dyads	Paper tool: educational information, pictographs of risks, self-directed patient values exercise, patient and caregiver perspectives Video	Knowledge at 1 month increased by 31% for patients and 6% for caregivers Values-treatment concordance at 1 month was shown with both patients and caregivers, where treatment choice aligned with stated values Satisfied with the length of the decision aids	Baseline study process for clinical trial as a whole was lengthy Video is 24 minutes long without search or skip functionality
Baylor, Deciding Together, PCORI [34]	15 LVAD candidates, 15 patients, 15 caregivers, 15 LVAD decliners and 11 clinicians involved in LVAD care were interviewed (n=71, total)	Alpha tested through cognitive interviews (n=5) and acceptability tested with LVAD patients (n=10), candidates (n=10), and clinicians (n=13) Multi-site randomized trial across 5 cardiovascular hospitals planned	Paper tool, educational, risk information in pictographs, patient values exercise, patient narratives Video (online)	Patients, caregivers and clinicians reported they would recommend the aid to patients considering treatment options for heart failure.	1 hour to review entire decision aid

1.4 NEED FOR IMPROVED DECISION MAKING

1.4.1 Physician support tools

The current tools available for physicians only meet the first 'right' of the CDS framework. Their utilization for LVAD implant decision making is limited by the accessibility of the tool and the workflow habits of the decision-making team [35]. The amount of information required to be manually entered to calculate the risk scores (demographics, labs, history, family support, clinical parameters, etc.) further reduces the likelihood of use in real-time decision making.

In this research thesis, I address the issues of model development with the *right data* (Aim 1) and presentation of the data in the *right format* and channel (Aim 2). Ongoing research by our colleagues addresses the *right person* to use the data and the *right time* in workflow the of support tool use [35].

1.4.2 Patient support tools

The risk information presented in current heart failure decision aids is limited to average probabilities for an aggregate population, for example, based on a clinical trial. However, recent utilization of machine learning and data mining, in combination with the growth of clinical data registries, has made it possible to develop patient-specific prognostic models. Our group has previously used these methods to develop personalized models for predicting LVAD outcomes, including mortality [36], recovery [37], and adverse events [38, 39].

Both tools currently in development for patients incorporate video or online components but use paper instruments for the primary education and have a patient self-guided response section. The response section is not used for analysis or shared with physicians but is intended to stimulate conversation with patients. There is no interactive component to the tools or feedback to the physician.

This study aims to develop well validated predictive models with the latest clinical information and statistical techniques that can be shared with patients (Aim 1) and address their needs in an understandable and unbiased visual format (Aim 3).

2.0 AIM 1: BAYESIAN NETWORK MODELS

The studies conducted in Aim 1 cover a range of modeling outcomes that demonstrate the versatility and utility of Bayesian networks to address the clinical decision making for LVADs.

In Aim 1.1 outcomes for all-cause mortality are modeled for four different time points. This represents the most common outcome cited when physicians and patients discuss risks after LVAD implant. The use of multiple time points allows different factors that impact risk to be highlighted, from the pre-implant patient surgical history and end-organ function for early outcomes to patient age and frailty for longer-term outcomes.

In Aim 1.2 the most commonly occurring adverse event is addressed: gastrointestinal (GI) bleeding. While GI bleeding is not associated with high patient mortality, it is closely related to the occurrence of other adverse events (e.g., re-hospitalization, infection, and stroke), and it significantly impacts quality of life. Recurrent GI bleeding is of the most concern to physicians, because it indicates inadequate medical management. In this aim, recurrent GI bleeding at any time after implant was modeled based on the patient pre-implant status and the patient status at the time of initial bleed. This is the first study to predict recurrent GI bleeding in this patient population.

Finally, in Aim 1.3 the most pernicious adverse event for patients with and LVAD was addressed: ischemic stroke. Ischemic stroke risk is driven by the device-patient interaction, particularly the blood-material interface of the pump and the resulting change in hemodynamics.

To consider the important role pump design plays in effecting the risk of ischemic stroke, models were made to measure the risk of ischemic stroke for each pump type: axial and centrifugal. The latter analysis revealed different sets of predictive factors to be important for each pump type.

The three predictive models developed in Aim 1 cover a spectrum of issues facing patients with LVADs and a range of data elements that can be used for model development with Bayesian networks.

2.1 A BAYESIAN MODEL TO PREDICT MORTALITY FOLLOWING LEFT VENTRICULAR ASSIST DEVICE THERAPY

2.1.1 Introduction

Aim 1.1 sought to develop Bayesian-based prognostic models of mortality for multiple time points following implantation of a continuous flow LVAD, using the Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS). Although various risk stratification models to predict mortality post LVAD have been proposed over the years, they all have limited applications in 'real life' decision making [40], due to their derivation from small data sets, limited number of variables or isolated to a specific pump in a study population [41-43]. Accurate predictions of outcomes after LVAD implantation depend on complex and dynamic interplay of multiple pre-operative variables that may not be captured by traditional multivariate modeling. Bayesian network (BN) modeling can account for dynamic, non-linear interactions *between* clinical and non-clinical variables and their influence on patient outcomes. In this way, they mimic complex human decision making, while drawing their diagnostic

algorithms from thousands of patients. Moreover, these models can predict outcomes at different time points post-LVAD by recognizing the time-varying importance of relevant variables. BN algorithms have been developed to predict mortality, gastrointestinal bleeding, and right ventricular failure in LVAD population [36, 39].

2.1.2 Methods

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 de-identified patient data for implantations undertaken between April 2006 and December 2016 (n=20,216). Modeling was performed using pre-implant patient information from January 2012-December 2015, for adult (over 18 years of age) patients receiving a primary continuous flow LVAD or LVAD and right ventricular assist device (RVAD) in combination (n = 10,277). We chose this time frame to include current generation, continuous flow LVADs with least amount of missing data and derived from over 160 clinical sites in the United States [7]. Total artificial heart recipients and RVAD-only receipts were excluded from this study. Patients who received device exchanges (n=800) were included in the study, with total time on pump calculated across the multiple implants. Patients who recovered while on LVAD support or received heart transplants were included and indicated as "alive" in modeling outcomes up until that time point and censored for subsequent time points. Mortality post-LVAD implantation was noted at the following times: 1 month, 3 months, and 12 months.

Data pre-processing

The INTERMACs data set includes over 400 pre-implant variables, with varying levels of data completion. BN construction requires no missing data in the training set. Therefore, preprocessing was required, in which the missing data elements were categorized into 2 sets: those missing in specific patterns (*missing, not at random*) and those that were 'truly unknown' (*missing at random*). An example of data missing, not at random, was if a patient did not complete a quality of life questionnaire because they were too sick, then the answers for all the questionnaire response variables were filled in as 'not applicable'. A missing at random example was, if a patient did not perform a 6-minute walk test and no reason why was documented, then the result of that test (distance walked) was classified as missing rather than left blank. Variables with over 40% missing were excluded from the analysis (n=42). Additionally, variables with less than 1% positive responses (e.g., Previous Dor procedure, done in only 8 patients) were removed from the analysis (n=16).

Some variables in INTERMACS capture information across a series of binary 'Yes / No' answers. To reduce the fields and improve the predictive power of variables, some fields were collapsed into multilevel variables. For example, INTERMACS has two variables for every comorbidity: Contraindication limiting transplant (Yes/No) and Contraindication, but not limiting transplant (Yes/No). These were collapsed into Contraindication: Yes, Yes-limiting transplant, or No. In this way, the number of variables for modeling was reduced. Fields for past medical interventions were combined into total counts of events, while keeping the individual binary information. For example, a patient with a coronary artery bypass grafting (CABG) and dialysis during their hospitalization was captured as CABG- Yes, Dialysis- Yes, and Total Event Count- Two. Variables with many levels were broken into subsets to identify important features.

For instance, Primary Diagnosis (a 31-level variable) was divided in to Ischemic Etiology, Restrictive Myopathy, Dilated Myopathy, and Congenital disease. After variable pre-processing, 203 pre-implant variables were used in the model construction.

Modeling cohorts

BN classifiers were derived for each time point of interest using a training dataset consisting of 80% of the data records selected at random (using Weka test/train split function.) The remaining 20% of data was held aside as the test set for the final model validation. This resulted in 6 sets of data across 3 pre-specified time points, three training sets—one for each model—and three test sets. The three training sets were each processed for feature selection independently.

Discretization of Continuous Variables

Bayesian modeling requires that all variables be categorical, therefore continuous variables must be discretized. In this study, four different methods of discretization were explored: expert binning (cut points determined for VAD implant guidelines, established risk tools, and normal ranges), supervised binning (MDL method in Weka), equal frequency binning, and equal width binning. Using training data, information gain was measured for each variable using each method, and results were compared. Choosing the method that yielded the highest information gain for each variable, a hybrid approach of expert binning, equal frequency, and equal width binning was used in to discretize the variables. This was performed for all 3 models.

Feature selection

To select variables for inclusion in the model, information gain was run in a 10-fold cross validation on the training data, with the recurring top variables being selected for model inclusion. Cut off for selection was information gain > 0.003 for all three time points. This resulted in a set of 29, 26, and 31 variables for the 1, 3, and 12-month models, respectively.

Bayesian Analysis

BNs process individual patient data in a dynamic and non-linear fashion to predict probable outcomes. The selected features from the training sets were used to learn both Tree Augmented Naïve Bayes (TAN) and Naïve Bayes (NB) graphical models using GeNie software (BayesFusion, Pittsburgh, PA). Each model was optimized by running 10-fold cross validation and removing or adding variables that either had low diagnostic value (as calculated in GeNie) or were on the cusp of the information gain cut off. At all three time points the NB models had superior performance, as measured by the area under the receiver operator characteristics curve (ROC AUC). The final NB models had 28, 26, and 21 predictive variables for the 1, 3, and 12-month models, respectively. Variables were grouped into three categories: demographics/patient status, medical history, and test results (laboratory, exercise and imaging).

Final Validation

Models were validated using the three test sets, which had not been used in the prior model learning. ROCs were plotted in R. In addition, we also report accuracy, sensitivity, and specificity (assuming a 50% threshold) of the Bayesian models' performance.

2.1.3 Results

A total of 10,277 patients met the inclusion criteria (Figure 4). The majority were between 50 and 69 years of age (n = 6,174; 60%); 78% (n = 8,044) were male; 3,811 patients (35%) received the LVAD as DT, and 5,528 patients (54%) were listed as BTT. Ischemic disease was listed as the cause for cardiomyopathy in 4,637 patients (41=5%). At the time of implantation, 16% (n = 1,671) were categorized as INTERMACS profile 1, 35% (n = 3,548) as INTERMACS profile 2, and 32% (n = 3,318) as INTERMACS profile 3. In the training sets (n = 8,222), 1-month mortality was 5% (n = 426), 3-month mortality was 9% (n = 776), and 12-month mortality was 18% (n = 1459). In the test sets (n = 2,055), mortality was at 6% (n = 114) at 1-month, 10% (n = 200) at 3-month, and 19% (n = 390) at 12-month post LVAD implantation.

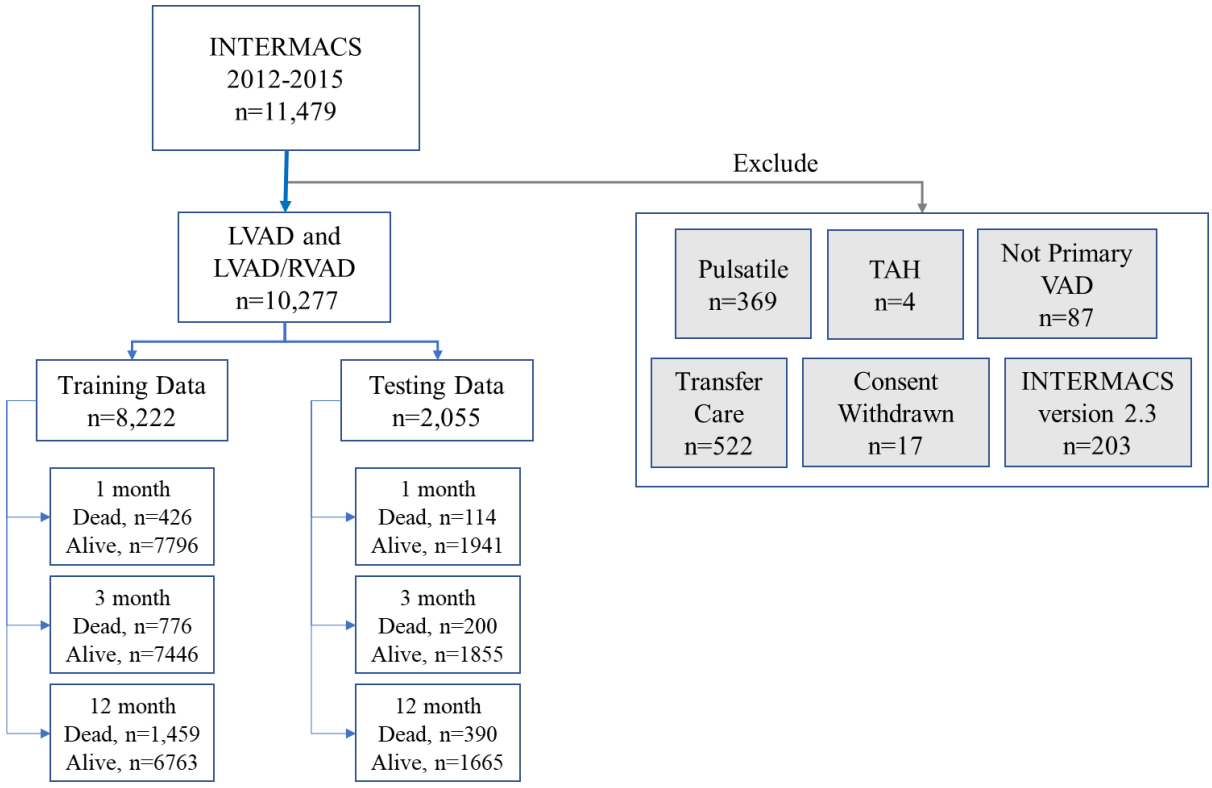


Figure 4. Model cohort selection

Models and test validation:

Bayesian models for 1, 3, and 12 months post-LVAD are illustrated in Figures 6,7, and 8. Variables are color-coded according to 3 categories: demographics/patient status, medical history and results. ROCs, accuracy, sensitivity, specificity, and AUC ROC are summarized in Figure 5 and Table 5. Accuracy ranged between 76% and 87%, and ROC AUC ranged between 70% and 71%.

Table 5. Model test validation performance

	1-month	**3-month**	**12-month**
Accuracy	87%	82%	76%
Sensitivity	30%	33%	33%
Specificity	90%	87%	86%
ROC AUC	70%	71%	70%

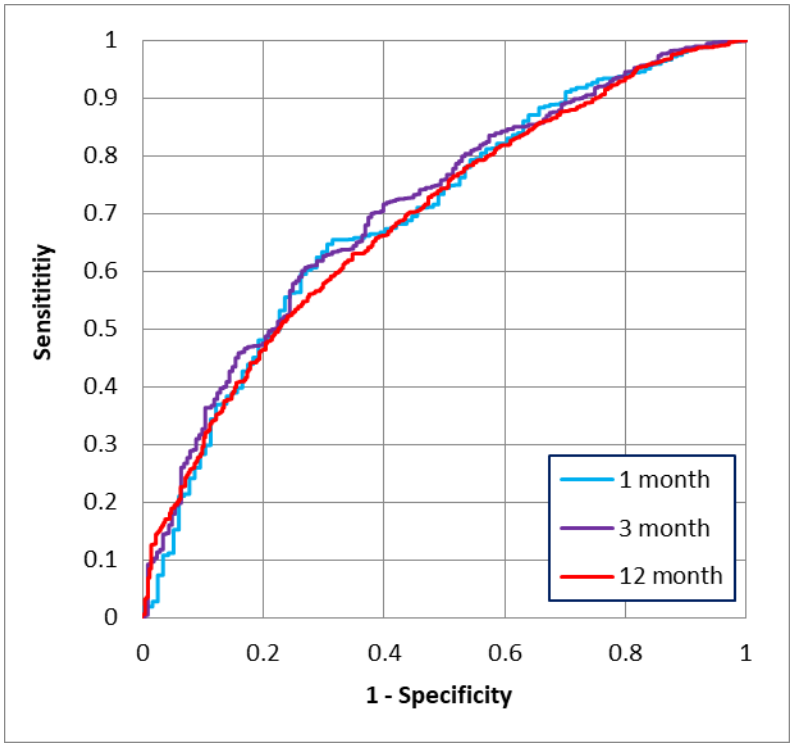


Figure 5. ROC curves for Bayesian models to predict various time points post LVAD implantation (0.70 at 1 month, 0.71 at 3 months, 0.70 at 12 months)

Mortality at 1-month post LVAD: This NB model contains 28 variables directly connected to the outcome (Figure 6). Although the order of influence changes as variables are observed or specified (i.e., while calculating the risk for a specific patient), the variables most predictive of early post-LVAD mortality are concomitant RVAD implant, total number of events during the implant hospitalization, low platelet count, high bilirubin levels, high aspartate aminotransferase (AST) level, and low INTERMACS profile.

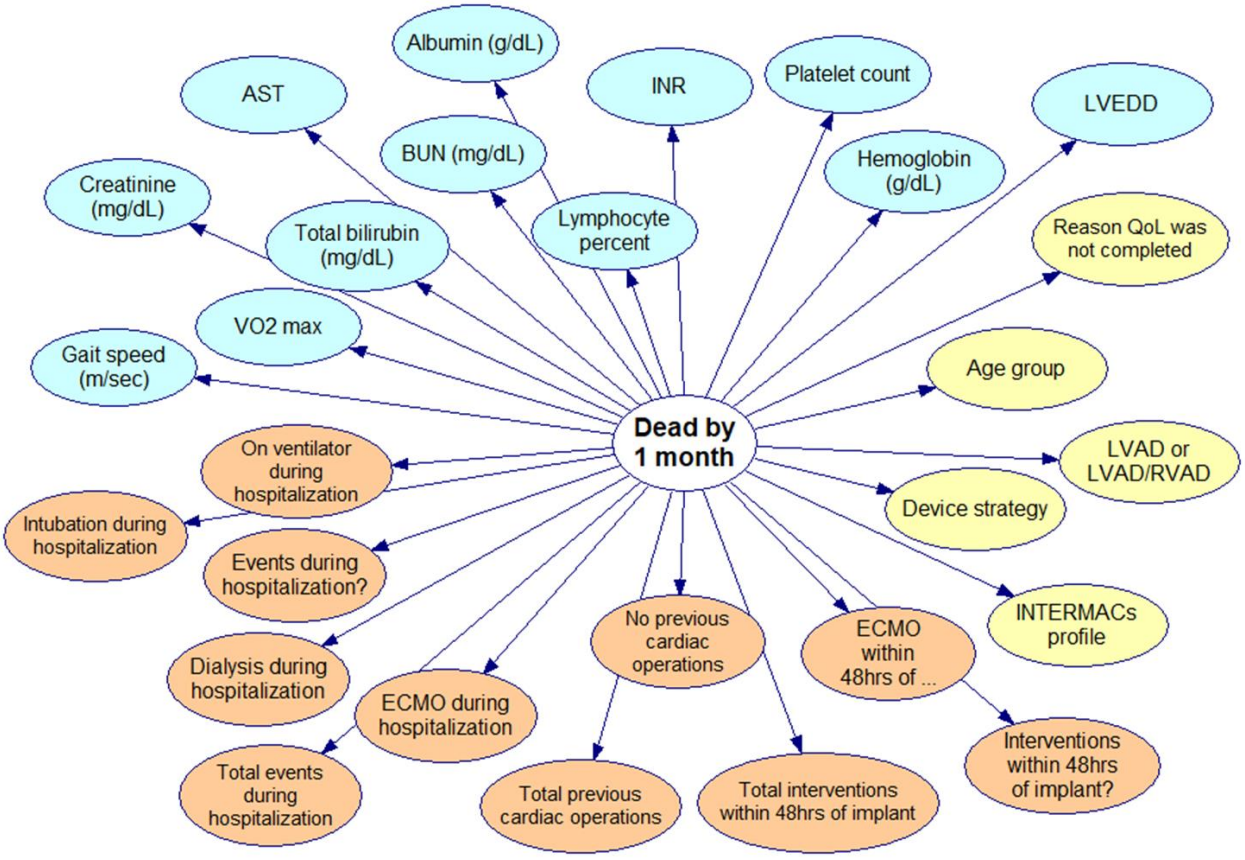


Figure 6. Bayesian model for predicting mortality 1-month post LVAD implantation. Variables are color coded: demographics (yellow), medical history (orange) and test results (blue).

Mortality at 3 months post LVAD: The NB model for mortality at 3 months post LVAD had 26 variables, with concomitant RVAD implant, older age, elevated blood urea nitrogen, low hemoglobin and lower INTERMACS profiles being highly predictive of higher mortality risk (Figure 7).

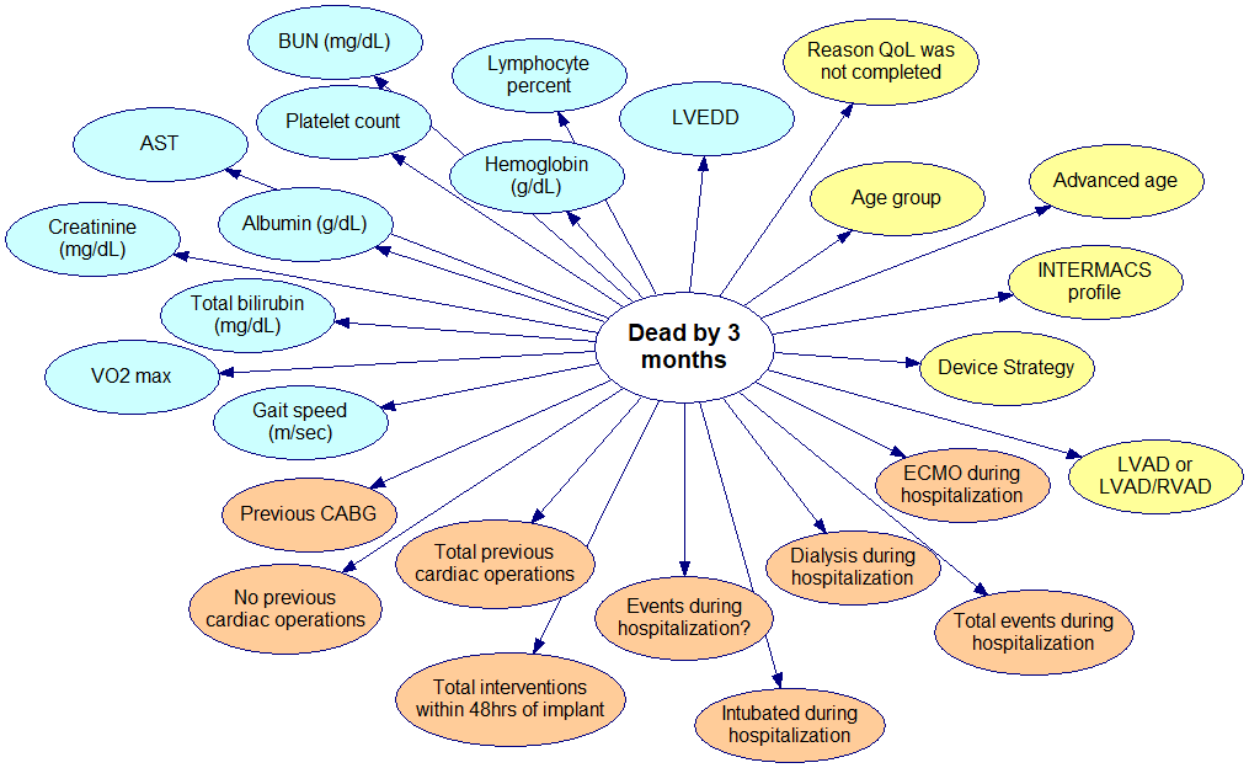


Figure 7. Bayesian model for predicting mortality 3 months post LVAD implantation

Mortality at 12 months post LVAD: The NB model for mortality at 12 months post LVAD had 12 variables, with older age, elevated blood urea nitrogen, low hemoglobin, DT device strategy, and concomitant RVAD implant being highly predictive of mortality (Figure 8).

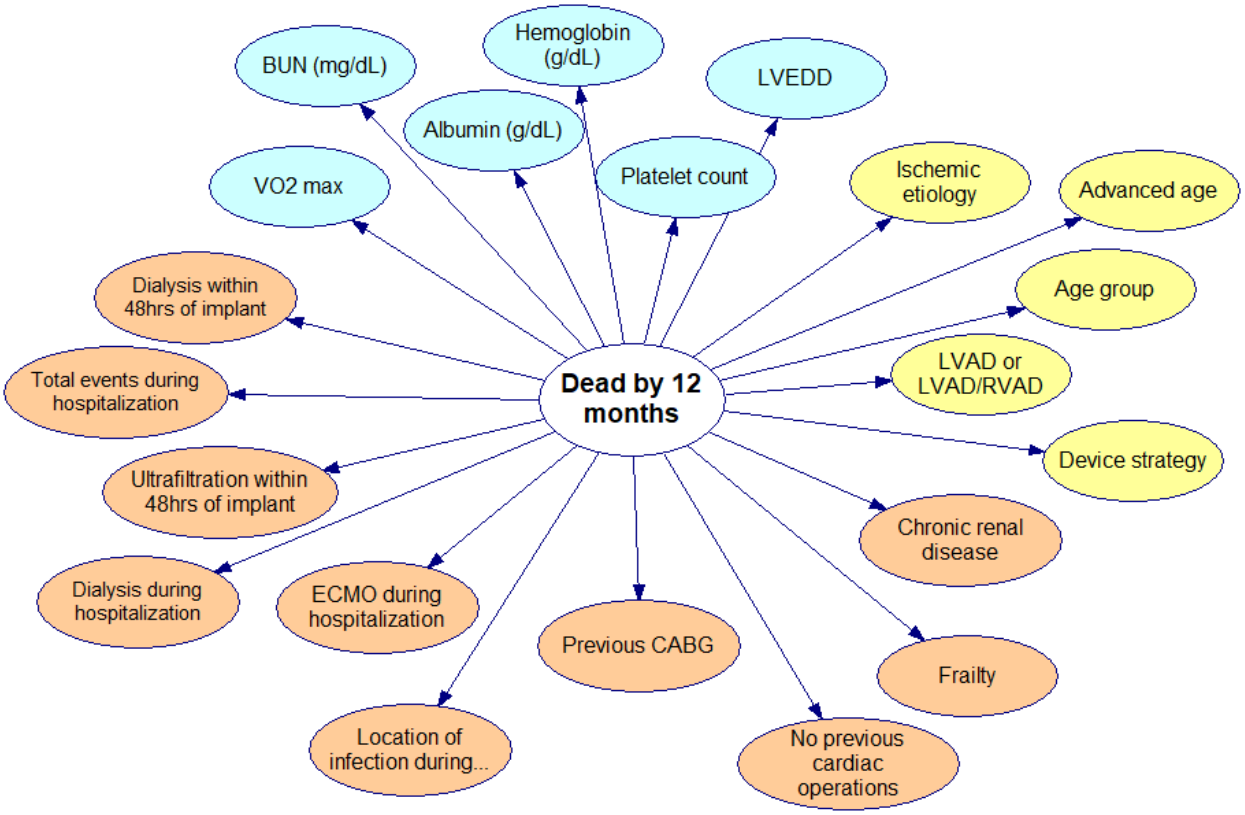


Figure 8. Bayesian model for predicting mortality 12 months post LVAD implantation

Unique variables across time points:

There were several variables that impacted risk of mortality across all time points such as old age, dialysis during index hospitalization, previous cardiac operations, albumin, platelet count, blood urea nitrogen. Similarly, there were several highly predictive variables that predicted short term mortality which were distinct from those predicting 12-month risk of death. These included lower INTERMACS profile, pre-operative ventilator dependence and hepatic function (indicated by AST and bilirubin levels) affecting 1-month mortality while ischemic etiology, history of chronic renal disease and frailty contributed more to 12-month mortality.

2.1.4 Discussion

Appropriate patient selection is key to optimal outcomes after LVAD therapy. There is a critical need for an accurate predictive model that is derived from a comprehensive database across multiple clinical sites, incorporates the impact of a large variety of clinical variables to account for the heterogeneity of end stage HF patient, and is up to date with the evolving technologic innovation in mechanical circulatory support devices. In other words, a successful predictive tool would mimic human decision making, while drawing on data from tens of thousands of patients who have undergone LVAD implantation. BN algorithms can provide the necessary tools to achieve this, as demonstrated in our analysis.

These analyses revealed a variety of risk factors from disparate categories (e.g., demographics, medical history, and laboratory test results) that influence post LVAD survival. Many of the variables that were found to be predictive in these models have previously been recognized as high risk factors in separate analysis [27, 44, 45]. Rather than trying to combine a multiplicity of factors by using a weighted summation, Bayesian models provide a dynamic incorporation of many variables, yielding a more robust ROC value than previously published scores [46]. The 90-day and 12-month HMRS stratifications had AUC of 60% and 57%, respectively [27], whereas the Bayesian 90-day and 12-month predictions exhibited AUC of 71% and 70%, respectively. BN analyses can show how clinical variables impact the predicted class value (mortality) without requiring that every patient variable be entered to give a prediction. This is an advantage over existing risk scores, which are rendered unusable if any of the parameters are not known.

In the present analysis, there were several variables found to have significant impact on the predicted mortality at different time points after LVAD implantation. 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-patient variable was the number of LVAD implants performed at a site annually, which has been shown to impact outcomes [27]. The final BN models included both non-modifiable/historical variables (such as patient age and surgical intervention history) and modifiable variables (such as nutritional assessment and renal function). Long-term mortality post-LVAD implant is likely more influenced by post-operative adverse events (such as stroke, infections, or right ventricular failure) than pre-operative variables, which is reflected by a slight drop in the ROC for the 12-month mortality model.

The ability to recognize the impact of different variables in predicting mortality at various time points post-LVAD implant is important, given that many high-risk variables (e.g. acute renal failure) that could impact short term mortality may reverse with time and be less relevant in predicting long term outcomes [47]. Although there are some high-risk variables that impact both short and long-term risk of mortality, their depth of impact may change over time. Extrapolating data from 90-day models to predict one-year mortality as was done in HeartMate Risk Score (HMRS) neglects this change in variable importance, but is overcome by using multiple, independent predictive models.

These BN mortality models demonstrated a remarkable improvement over existing models with respect to accuracy, specificity and ROC. The models in this study have an ability to (1) learn from prior probability, (2) apply to the most recent patient mix and device technology, and (3) be more tolerant to missing data elements when calculating predictions. In addition, BNs

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.

Limitations:

We acknowledge that this study has several important limitations, including missing data pertaining to the independent variables. Although the INTERMACS database is large and representative, it suffers from sparsity of many data elements. This prompted us to exclude some variables which may have been relevant predictors. Additional limitations include inherent retrospective bias (all patients were already chosen to receive an LVAD) and only FDA approved VAD devices were included in registry. 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.

2.1.5 Conclusion

The BN mortality models show great promise as reliable and accurate risk stratification tools for clinical decision making. The potential utility of the models is to assist the medical team in decision making with patients for whom the merits or contraindications to LVAD implantation are not immediately clinically apparent. Accordingly, we hope that CORA will promote the appropriate and perhaps judicious use of LVAD therapy by providing clinicians and patients a more informed decision regarding potential short-term and long-term outcomes.

2.2 PREDICTING PREDISPOSITION TO AND RECURRENCE OF GI BLEEDING IN PATIENTS WITH CF-LVADS

2.2.1 Introduction

Gastrointestinal (GI) bleeding is one of the most frequently occurring adverse events in patients who have continuous flow left ventricular assist devices (CF-LVADs), with an incidence reported between 18% and 40% [48, 49]. Typically recurrent, it substantially impacts the patient's quality of life through frequent readmissions, prolonged hospitalizations [50], and a potentially higher risk of infection and thromboembolic events [51]. In addition, the associated blood transfusions may result in allosensitization [52] which impacts the patient's transplant candidacy[53, 54] and can pose a long term risk of post-transplant rejection.

The etiology of GI bleeding in patients with CF-LVADs has been studied extensively and is likely multifactorial. Non-pulsatile blood flow [55], which may lead to vascular stiffening [56, 57] or arteriovenous malformation (AVM)[58], unfolding of von Willebrand factor from the shear stress of the impeller that leads to increased susceptibility to degradation and an acquired von Willebrand syndrome [59], elevated Angiopoeitin-2 levels due to coagulation factors from blood-metal interface leading to AVM [60], and the need for anticoagulation and antiplatelet therapy[61] have all been associated with elevated risk of GI bleeding.

Pre-implant clinical risk factors that are known to be associated with increased incidence of GI bleeding include older age, elevated creatinine, and pre-operative right heart failure [52, 57]. Given the heterogeneity of GI bleeding etiology and the effect of both pre- and post-LVAD implant factors, multivariate modeling or focusing on pre-implant variables alone may not be adequate for identifying patients who are at risk for GI bleeding. There is a need to better

understand the factors that impact the risk of GI bleeding, which may provide insights as to how GI bleeding may be mitigated or prevented. In this study, we use Bayesian network modeling on a large, retrospective data set to construct two predictive models to better characterize the risk factors and causes for GI bleeding (Figure 9).

2.2.2 Methods

Data Source

The data for this study was derived from the Interagency Registry for Mechanical Assisted Circulatory Support (INTERMACS), funded by the National Heart, Lung and Blood Institute, National Institutes of Health, Department of Health and Human Services.

Inclusion criteria for this study were the use of a CF-LVAD as the primary implant between years 2010 and 2015 in patients over 18 years old. Patients who received biventricular ventricular assist devices (BiVADs) were included. Patients with temporary RVAD support alone were excluded. Total artificial heart implants and pulsatile LVAD implants were excluded. Patients were excluded from the model if they died or were transplanted within the first 30 days of implant.

Endpoints

Incidence of GI bleeding was determined using INTERMACS event data and definitions. A GI bleeding event was identified as an *upper GI bleed*, *lower GI bleed*, or *positive occult stool (location unknown)*. To assess GI bleeding that was caused by the LVAD implant, initial GI bleeding was defined as a GI bleeding event occurring more than 2 weeks after implant. GI bleeding events occurring before 2 weeks were not counted. Recurrent GI bleeding was identified as an additional GI bleeding event occurring 2 or more weeks after the initial bleed.

This outcome is used as the study endpoint because a patient with a single GI bleeding event with no recurrence can be considered effectively managed and the GI bleed is less likely to be LVAD related. Specifically, our goal is to address CF-LVAD-associated reasons for GI bleed, which may result from the chronic dis-regulated angiogenic state that leads to multiple bleeds. Focusing on recurrent GI bleeding allows for prediction of successful (non-hemorrhagic or non-recurring) versus unsuccessful (recurrent GI bleeding) outcomes. There was no maximum time limit set between initial and recurrent bleeding events.

Model scope

Two models to predict recurrent GI bleeding were constructed in this study (Figure 9.) The first model used patient pre-implant health information to predict the risk of recurrent GI bleeding after CF-LVAD implantation. The *pre-implant predisposition model* used mostly non-modifiable patient characteristics. The modeling data set included 13,082 patients, with 1,439 (11%) having recurrent GI bleeding.

The second model used post-operative factors at the time of a patient's first GI bleeding event to predict the risk of a second GI bleeding event. This is referred to as the *post-implant risk model* for recurrent GI bleeding events. The goal of this model was to provide insight on the LVAD-related factors associated with recurrent bleeding. The model used medications at the time of the initial GI bleeding event (e.g., anticoagulation), labs (e.g., INR) and therapeutic interventions within one week of the initial GI bleeding event. Of the 3,139 patients with a single GI bleeding event at least two weeks after CF-LVAD implant, patients without data within a one-week window of their initial bleed were excluded (n= 1,612). The remaining 1,527 patients were used in the post-implant risk model building, with 49% having a recurrent GI bleed.

Model development and validation

INTERMACS data was processed by discretizing continuous variables using the supervised binning class-attribute interdependence maximization (CAIM) method, which was iterated 5 times[62]. Missing data was not imputed but was captured categorically as "missing".

Feature selection was performed using information gain and hill climbing, each with 10-fold iterations, to select the primary variables impacting recurrent GI bleeding. After creating a subset of variables using information gain (gain > 0.003) and the most frequently selected variables from hill climbing, a Tree Augmented Naïve (TAN) Bayes model was created and validated by 10-fold cross validation using GeNie software (BayesFusion, Pittsburgh, PA). Variables in the model were evaluated for their impact on the prediction by diagnostic value. Diagnostic value is a measure of the influence the variable has on the model prediction, based on the expected gain in cross-entropy. Variables with the lowest diagnostic value were removed from feature selection and the model was re-learned and validated with 10-fold cross validation. Variables were iteratively removed and added until the model performance, defined by the receiver operating characteristic area under the curve (ROC AUC), no longer improved.

Model validation was performed using more recent patient data from INTERMACS, covering new implants done in 2015-2016, as well as data from previous patients who had not experienced recurrent bleeding by 2015 and who were still alive on their original CF-LVAD. The pre-implant predisposition model had 3,351 patients and the post-implant risk model had 1,236 patients in the validation cohorts.

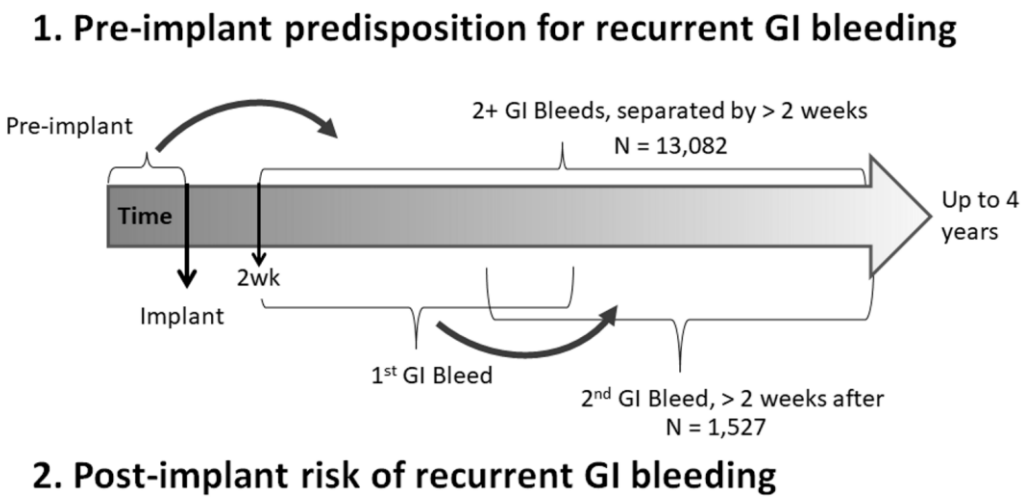


Figure 9. Schematic of Patient Selection for Models.

2.2.3 Results

Incidence of GI Bleeding

Of the 13,082 patients implanted with a primary CF-LVAD between 2010 and 2015, 3,505 patients (27%) had 7,426 episodes of GI bleeding. The mean number of GI bleeding events per patient was 2.1 with a range of 1 to 45 per patient. GI bleeding events were experienced by patients from 0.5 to 64 months after implant, with the mean time to the first GI bleeding event of 6.6 months and median time of 2.4 months. The mean time between the first and second GI bleed was 4.3 months and median time of 1.7 months (range 0.5 to 61 months).

Pre-implant characteristics of patients with and without recurrent GI Bleeding

Of the 13,082 patients, 1,439 (11%) experienced recurrent GI bleeding. Compared to the non-recurrent bleeding cohort (patients with 0 or 1 GI bleeding event), patients who had recurrent GI bleeding events were more likely to be older (age 60-79), on an axial flow pump (HeartMate II), and implanted as a destination therapy (Table 6).

Table 6. Clinical Characteristics of Patients by GI Bleeding Recurrence

Patient Information		**Non-Recurrent GI Bleeding (0-1 events), n = 11,643**		**Recurrent GI Bleeding (2+ events), n = 1439**		
		n	**%**	**n**	**%**	***p-value***
Pump Type	Centrifugal	2069	18%	140	10%	*< 0.001*
	Axial	9572	82%	1299	90%	*< 0.001*
Device	LVAD	11320	97%	1413	98%	*0.032*
	BiVAD	323	3%	26	2%	*0.032*
NYHA Class	III	2112	18%	245	17%	*0.298*
	IV	8608	74%	1085	75%	*0.230*
Strategy	DT	4803	41%	829	58%	*< 0.001*
	BTT	6749	58%	600	42%	*< 0.001*
Age	80+	75	1%	14	1%	*0.153*
	70-79	1450	12%	307	21%	*<0.001*
	60-69	3613	31%	651	45%	*<0.001*
	50-59	3243	28%	354	25%	*0.009*
	40-49	1734	15%	89	6%	*<0.001*
	30-39	952	8%	21	1%	*<0.001*
	19-29	576	5%	3	0%	*<0.001*
Gender	Male	9152	79%	1121	78%	*0.542*

Data shown as the total in each category and percentage of total, with comparison between groups measured using a two-way z-test. Acronyms: LVAD, Left Ventricular Assist Device; RVAD, Right Ventricular Assist Device; DT, Destination Therapy; BTT, Bridge to Transplant (includes patients listed and not-yet listed). In this dataset, all axial flow pumps are Heartmate II (Abbott) and all centrifugal flow pumps are HVAD (Medtronic).

Recurrent GI bleeding Predisposition Model and Results

Out of 261 pre-implant variables that were used in feature selection, 18 were identified as the top predictors of predisposition for recurrent GI bleeding (Figure 10). Final cross-fold validation ROC AUC was 69% (Figure 11). In the validation data set from INTERMACS

(n=3,351 patients), 13% of patients experienced recurrent GI bleeding events. Model performance with this test data was ROC AUC of 68% (Figure 11).

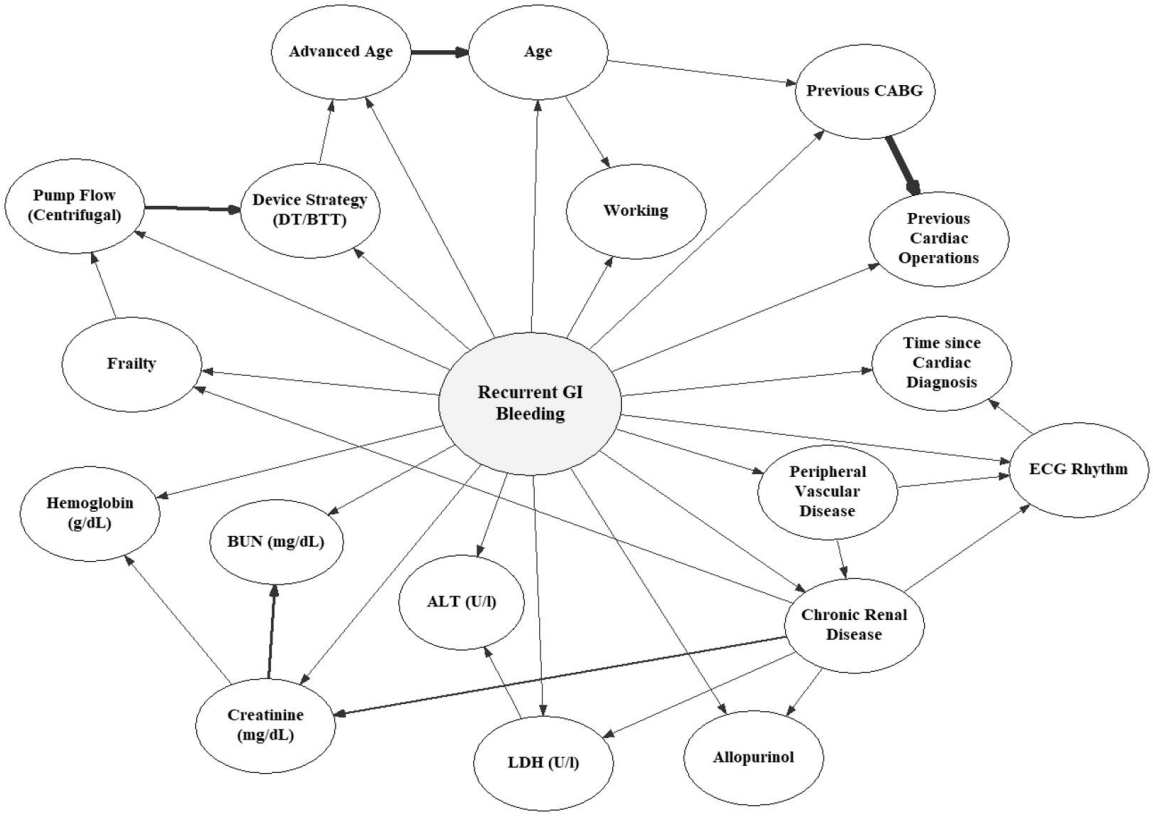


Figure 10. Predisposition to Recurrent GI Bleeding Model. Arrows indicate a relationship between variables, with arrow thickness indicating strength of relationship.

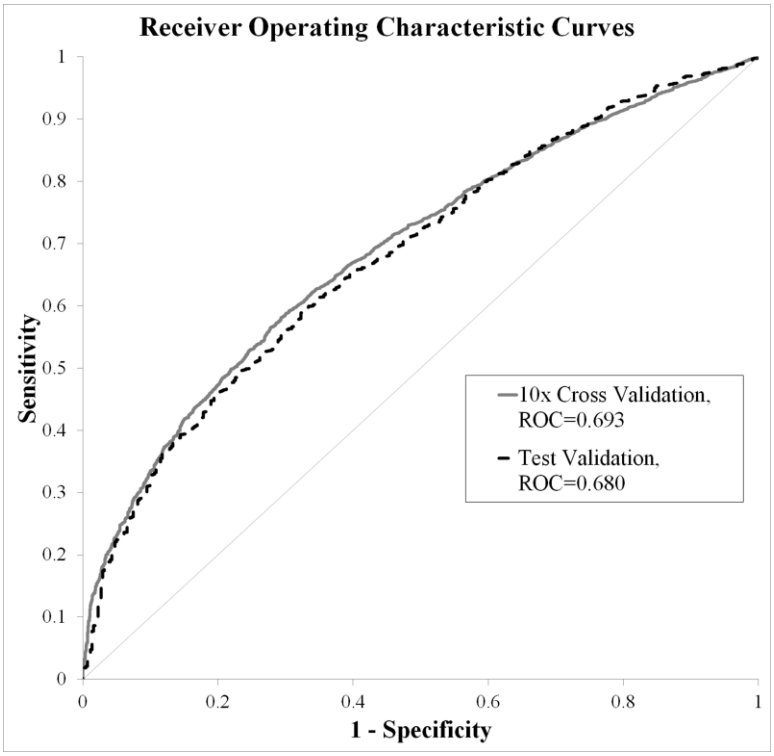


Figure 11. Receiver operating characteristic curves for predisposition to GI bleeding events, by cross-fold and test validation

The most predictive variables driving recurrent GI bleeding risk were: age, previous cardiac operations, anemia (low hemoglobin), destination therapy device strategy, axial flow pump, and elevated blood urea nitrogen (BUN). A summary of all the variables, their diagnostic value, and effect on risk is captured in Table 7.

Table 7. Summary of Variables Predicting Predisposition to Recurrent GI Bleeding

Variable	**Diagnostic Value**	**Increase or Decrease Risk**
Age	0.05	Increase with age
Device strategy	0.019	Increase when DT
Previous CABG	0.018	Increase
Previous Cardiac Operations	0.016	Increase
Advanced Age	0.014	Increase
Hemoglobin (g/dL)	0.008	Decrease
BUN (mg/dL)	0.007	Increase
Pump Flow	0.007	Increase when axial flow
Creatinine (mg/dL)	0.006	Increase
LDH (u/L)	0.006	Decrease when LDH increases
ECG Rhythm	0.006	Increase with Atrial Fibrillation
Frailty (INTERMACS definition)21		
0.005	Increase	
Peripheral Vascular Disease	0.005	Increase
Time Since Cardiac Diagnosis	0.005	Increase with time since diagnosis
ALT (U/L)	0.005	Decrease when ALT increases
Allopurinol	0.005	Increase when used
Working	0.005	Increases if patient not working
Chronic Renal Disease	0.004	Increase

Post-implant Risk of Recurrent GI Bleeding Model and Results

Variables used in the post-implant risk of recurrent GI bleeding event model included independent patient variables (e.g., patient age), medications at the time of the initial GI bleeding event and lab values and adverse events experienced by the patient within one week of the initial GI bleed. Out of 92 variables that went into feature selection, 16 were selected for the post-implant TAN risk model (Figure 12). Final cross-fold validation ROC AUC was 61% (Figure 13). More recent data from INTERMACS (n=1,236 patients), which was not used in the model learning, was used to validate model performance. In this test set, 39% of patients experienced a recurrent GI bleeding event. The test validation had a performance of ROC AUC of 60% (Figure 13).

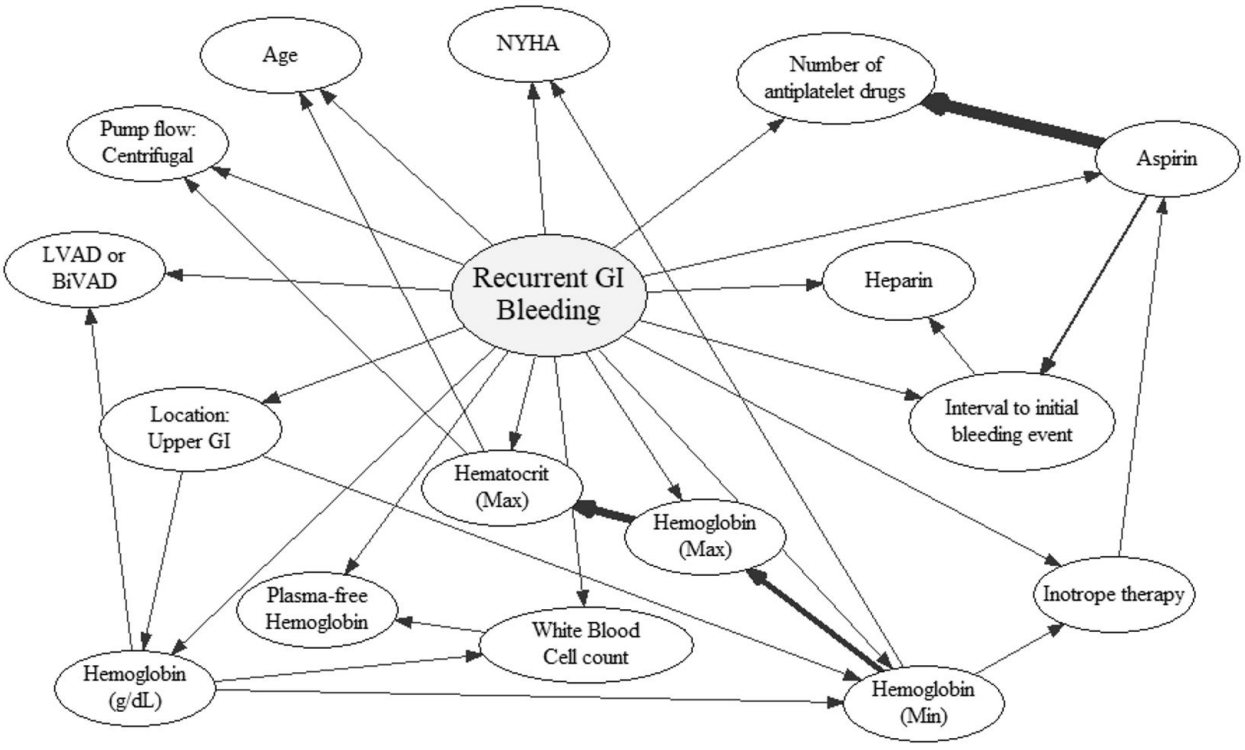


Figure 12. Tree Augmented Naive Bayesian network of post-implant recurrent GI bleeding event risk. Arrows indicate a relationship between variables, with arrow thickness indicating strength of relationship.

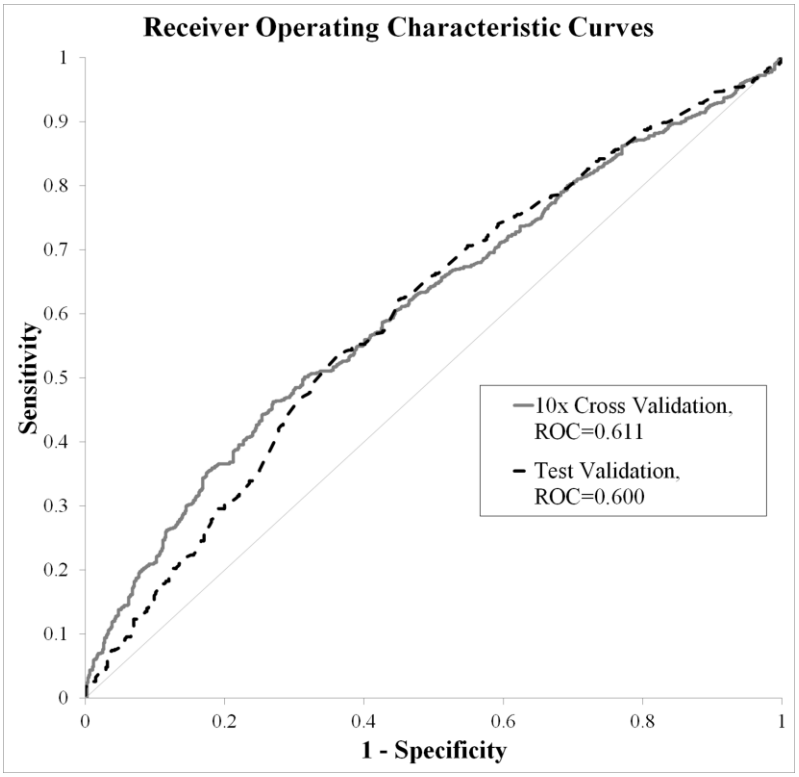


Figure 13. Receiver operating characteristic curve for post-implant risk of recurrent GI bleeding events, cross fold and test validation

Of the 16 model variables, the most predictive were: hematocrit and hemoglobin, age, and plasma free hemoglobin (Table 8). Four of the 16 variables were treatments, possibly modifiable by clinicians. The rest were lab values from the time of the initial bleed (6), independent variables (4), and time and location of the bleed (2).

Table 8. Variable in post-implant reoccurrence model and impact on risk

Variable	Diagnostic Value	Increase or Decrease Risk	Category
Hematocrit (Max)	0.042	Decreases	Lab
Hemoglobin (Max)	0.014	Decreases	Lab
Age	0.014	Increases	Pre-implant, Independent
Min Hemoglobin	0.011	Decreases	Lab
Plasma-free Hemoglobin	0.01	Increases	Lab
Interval to initial bleeding event	0.008	Decreases with time	Independent
White Blood Cell Count	0.007	Decreases	Lab
Hemoglobin (Min)	0.007	Decreases	Lab
Hemoglobin at bleed event	0.007	Decreases	Lab
Heparin	0.007	Decreases	Treatment
Inotrope Therapy	0.005	Decreases	Treatment
Aspirin	0.004	Decreases	Treatment
Antiplatelet Count	0.004	Decreases	Treatment, summary
Pump flow: Centrifugal	0.004	Decreases	Pre-implant, Independent
Location: Upper GI	0.003	Increases	Independent
LVAD or BiVAD	0.003	BiVAD decreases	Pre-implant, Independent

2.2.4 Discussion

We present the first risk models for recurrent GI bleeding in patients with CF-LVADs, using Bayesian networks to analyze both pre- and post-implant risk factors. The pre-implant predisposition model performed with a ROC AUC of 0.68, while the post-implant risk model was less successful with a ROC AUC of 0.60. Both models identified features of high risk patients and provide insight into the complex pathophysiology of GI bleeding associated with CF-LVAD use.

The greatest predictor of pre-implant predisposition for recurrent GI bleeding events was patient age. This is in line with findings from previous studies of GI bleeding risk factors[57, 63-65]. Old age is also a predictor of spontaneous AVM formation in elderly patients without heart failure[66].

In addition to age as an objective measure, INTERMACS captures the subjective physician assessment of a patient being of "advanced age", which is independent from actual patient age. Initially used to indicate whether a patients' age prevents them from receiving a heart transplant, the definition was expanded to include any concern about implanting an LVAD that a physician may have due to the patient's age[67], such as frailty. This indicator is an important factor in the pre-implant predisposition model; for example, patients who are considered "advanced age" in the 60-69 years old group have higher risk of recurrent GI bleeding (20% risk) compared to patients 60-69 years of age who are not "advanced age" (14% risk) or patients who are "advanced age" and 70-79 or 80+ years (18% risk for both). The inclusion of physician intuition increases the utility and performance of the predisposition model.

A similar factor indicating patient overall health is patient work status. Patients who work either full or part time at the time of the CF-LVAD implant have a lower risk of recurrent GI bleeding. This is probably due to these patients being more likely to be younger and in better health than those who are not working. This relationship with working and better outcomes has also been seen in patients with heart transplants[68].

Patient age is a related factor with other variables in the predisposition model, such as device strategy (DT) and type of pump (HeartMate II). DT patients are typically older than BTT patients and the only device approved for DT at the time of this data collection was the HeartMate II, which is an axial flow device. This relationship between age and device strategy

has also been noted in the literature [57]. However, these additional variables do help differentiate risk in the predisposition model, for example: a 65-year-old patient who is BTT and listed for transplant has a 7% risk of recurrent GI bleeding if on a centrifugal flow pump, compared to a 12% risk for a similar patient on an axial flow pump. Similarly, the 65 years-old listed as DT on an axial flow pump has a 20% risk of recurrent GI bleeding.

The type of pump flow has been associated with factors that may impact GI bleeding risk. Centrifugal flow pumps have less hemolysis [69], possibly due to their lower rotations per minute and/or lower blood-pump contact area. This lower hemolysis may decrease angiogenesis and therefore AVM-related GI bleeding [59]. However, other studies suggest that the pump design differences contribute less to the risk than the confounding age and device strategy factors [59, 60].

Hemoglobin/hematocrit (Hgb/Hct) levels were influential factors in both the pre- and post-implant risk models for recurrent GI bleeding. Low pre-implant Hgb was associated with a higher risk of post-implant recurrent GI bleeding events. While pre-implant Hgb alone may not be a causal factor, the etiology behind the low Hgb is the likely factor driving recurrent GI bleeding risk, as it may indicate persistent, low grade GI bleeding [70].

Similarly, low post-implant maximum Hgb/Hct, the highest Hgb/Hct levels between the patients' last follow up and the time of the first GI bleeding event, were associated with an elevated risk of GI bleeding recurrence. Another potential explanation for the low maximum Hgb/Hct is subclinical hemolysis due to the CF-LVAD created shear stress and blood-metal interface [71]. This is further supported by high plasma-free hemoglobin as an additional risk factor. Hemolysis is a surrogate for turbulent flow beyond the design parameters of the device;

such turbulence facilitates degradation of vWF and is thought to further impair vWF-regulated angioneogenesis with increased risk of AVM formation [59].

The post-implant model identified anticoagulants (aspirin, heparin) and inotropes as predicting a lower risk of recurrent GI bleed. Whether these treatments were maintained between the first and second GI bleed event and at what doses were not assessed, therefore few conclusions can be made as to their role in preventing recurrent GI bleeding. One explanation may be successful adjustment of anticoagulation regimens: if anticoagulation contributes to a first GI bleeding event, it can be adjusted to prevent a recurrence of bleeding. Inotrope use may be connected to the pulsatility-related mechanism of GI bleeding; single center evidence has shown that inotropes, specifically epinephrine, can increase cardiac pulsatility and decrease GI bleeding risk [72]. However, the association of inotrope therapy with decreased recurrent GI bleeding in our model may be confounded by very sick patients on inotropes dying before they have a recurrent GI bleed. Similarly, patients on BiVADs have lower risk of recurrent GI bleeds, likely due to their diminished overall survival[73].

These models can be used to identify patients at risk for recurrent GI bleeding and allow for more careful planning for and management after CF-LVAD implant. For example, patients with low pre-implant Hgb may benefit from capsule endoscopy assessment prior to implant surgery to rule out pre-existing sources of GI bleed.

The pre-implant predisposition for GI bleeding model adds to the tool kit of risk assessments that physicians can use when making decisions about the care and management of patients receiving LVADs [27, 39, 74, 75]. The post-implant risk model highlights the potential CF-LVAD related causes of GI bleeding. While the present ROC AUC is modest, this is an important step to improving prediction and understanding of GI bleeding in CF-LVAD patients.

These predictive models will be made available to clinicians to evaluate as part of the Cardiac Outcomes Risk Assessment (CORA) decision support tool, which is available for demonstration use at www.app.mycora.org.

Limitations:

The data collected in INTERMACS for GI bleed does not separate AVM related bleed from other sources of GI bleed. This is a major hindrance in being able to classify the etiology of GI bleeding, particularly with how it relates to the CF-LVAD implant. Future work will use data from individual clinical sites that specifies AVM etiology to hypothesize physiologic reasons for the elevated AVM bleed risk. Another limitation is the low ROC AUC for the post-implant recurrent GI bleeding risk model. This may be due to the open-ended time interval used for recurrent GI bleeding. New models that examine recurrent GI bleeding over specific intervals (e.g., within 3, 6, or 12 months) could exclude patients who passed away or were transplanted and may improve predictive performance.

2.2.5 Conclusions

The important risk factors for recurrent GI bleeding can be identified for patients before they receive an LVAD implant and after an initial bleed occurs. The primary predictors for bleeding in both models are patient age and hemoglobin levels. Subclinical bleeding and possible hemolysis from CF-LVAD function may increase risk of GI bleeding in CF-LVAD patients. Physicians can use these models to identify high risk patients to monitor them for bleeding, as well as consider the best pump type to implant. Further work is required to identify the origin of the GI bleed, be it AVM or other causes, and determine the influence of specific risk factors on the pathophysiologic mechanisms driving GI bleeding.

2.3 RISK FACTORS FOR ISCHEMIC STROKE AFTER CF-LVAD IMPLANT BY PUMP TYPE

Abstract: The risk of stroke continues to be a major adverse event after CF-LVAD implantation, limiting the utility of CF-LVADs. Ischemic stroke risk is directly related to factors arising from the pump-person interaction of the CF-LVAD, but these factors may differ by pump design. Using a Bayesian Network machine-learning approach, we predicted pre-implant risk for ischemic stroke in patients with axial or centrifugal flow pumps at 3 months after LVAD implant. Features of high risk patients on axial flow pumps were elevated c-reactive protein, invasive interventions during the CF-LVAD hospitalization and myocardial infarction. The features of high risk patients on centrifugal flow pumps were smaller patients, not using diuretics or antihypertensive medications. Common factors to both pump types were old age and elevated blood pressure. The performance of the risk predicting Bayesian model was a ROC AUC of 61% for axial and 66% for centrifugal flow pumps.

2.3.1 Introduction

Stroke is one of the most devastating adverse events affecting patients who receive a continuous flow left ventricular assist device implant (CF-LVAD). It is associated with high mortality and morbidity [76, 77], reduced patient quality of life, and impaired candidacy for

heart transplant [78]. Due to the high level of adverse effects, the risk of stroke is one of the main reasons CF-LVADs are not recommended for use in the less-sick heart failure patient population [79].

Strokes occurring in CF- LVAD patients can have either a hemorrhagic or ischemic etiology, with incidence of each type reported from 0-16% and 4-17.1%, respectively [80]. The causes of and treatment for each stroke pathology are different [81]; therefore the factors impacting post-CF-LVAD risk may also be different. For this reason, the present study focuses solely on causes and risk factors for ischemic stroke.

Ischemic stroke falls into the category of CF-LVAD adverse events that arise from the pump-patient interface [82]. The interaction of blood with the metal interface, potential blood damage from the high-speed rotors, change to continuous blood flow, and increased potential for infection are all pathology-effecting factors that arise from the use of the CF-LVAD. Because of this, it makes sense that the device type, surgical technique and associated medical management may impact the risk of resulting adverse events, like stroke. In fact, recent data has indicated that pump type does play a role in stroke incidence and risk mitigation [83, 84]. However, analysis of the pre-operative predictive risk factors for ischemic stroke do not often distinguish between the two main pump types being used clinically [76, 84].

The goal of this study is to identify and compare pre-operative patient features associated with an elevated risk of ischemic stroke after axial or centrifugal CF-LVAD implant. Differences between the risk factors associated with each pump type can help identify the causative factors for ischemic stroke and be used in decision making for selecting appropriate CF-LVAD candidates.

Bayesian models were used to create the risk predictions in this analysis due to their ability to handle the interaction of many related pre-operative variables.

2.3.2 Methods

Data set and definitions

The data for this study was derived from the Interagency Registry for Mechanical Assisted Circulatory Support (INTERMACS), funded by the National Heart, Lung and Blood Institute, National Institutes of Health, Department of Health and Human Services under Contract No. HHSN268201100025C. More information is available at: www.intermacs.org. IRB approval was obtained through the hospitals participating within INTERMACS.

Ischemic stroke was defined as a neurologic adverse event that was an ischemic/embolism type of cerebrovascular accident (CVA), using INTERMACS nomenclature.

Patient cohort

Inclusion criteria for this study were the use of a CF-LVAD as the primary implant and age over 18 years old. Patients who received Bi-VAD (left and right VADs) were included. Total artificial heart implants and pulsatile LVAD implants were excluded. The time frame for implants was between 2012 and 2016. Patients were censored for transplant, explant, or non-ischemic stroke related death before 3 months post-CF-LVAD implant.

Data pre-processing

Pre-implant patient data was split into two patient populations: patients receiving a primary axial flow pump and patients receiving a primarily centrifugal flow pump. Outcomes for each data set were occurrence of an ischemic stroke within 3 months of CF-LVAD implant.

Data was processed by discretizing continuous variables using equal width binning for each of the two patient populations. This method was determined by comparing the information gain of each variable after discretization by three different methods: supervised binning, equal width binning, or equal frequency binning and selecting the method that had the highest information gain. This was done independently for the axial flow and centrifugal flow models. The number of bins was determined by comparing the Naïve Bayes model performance with continuous variables split into 2 to 10 bins and selecting the version with the highest receiver operating characteristic area under the curve (ROC AUC). In the axial flow model, continuous variables were discretized into 10 equal width bins, and, in the centrifugal flow model, continuous variables were discretized using 7 equal width bins.

No imputation was performed for missing data.

Variable feature selection and model training

Data for each time point was divided into two parts: a training data set comprising 80% of the data, and a test set of 20%. Splits were made randomly in Weka. Training data was used for feature selection and model structure and parameter learning, while test data was only used in the final performance validation.

Feature selection was performed using information gain and hill climbing (Weka) to select the variables most related to ischemic stroke risk for each pump type. Information gain threshold was set at gain > 0.003.

Using the resulting feature selected variables, a Tree Augmented Naïve Bayes (TAN) model and a Naïve Bayes (NB) model were created (GeNie, BayesFusion, Pittsburgh, PA) to classify the outcomes for each pump type. The initial models were validated by 10-fold cross validation and each variable was evaluated for diagnostic value. Diagnostic value is a measure of the influence the variable has on the model prediction, based on the expected gain in cross-entropy. The lowest ranked variables were removed, and the model was re-learned and then validated with 10-fold cross validation. Variables were iteratively removed and added until the model performance, defined by ROC AUC, no longer improved. This procedure was performed independently for the axial and centrifugal flow patient populations. Outcomes from the TAN and NB models were compared to select the best performing model.

Model validation

The test set of data, comprising 20% of the initial patient data set, was used for performance validation of the final axial and centrifugal flow ischemic stroke models. The test data sets were also used to measure the performance of a recently published ischemic risk score [76].

2.3.3 Results

Patient cohort

Out of 13,593 patients who received CF-LVADs between 2012 and 2016, 937 (7%) patients experienced an ischemic stroke at some point after implant. Of these ischemic strokes, 32% were fatal. Of the patients who experienced ischemic stroke, 47% had them by 3 months after implant (Figure 14). When considering timing by pump type, 45% of axial pump and 57% of centrifugal pump ischemic strokes occurred by 3 months after implant.

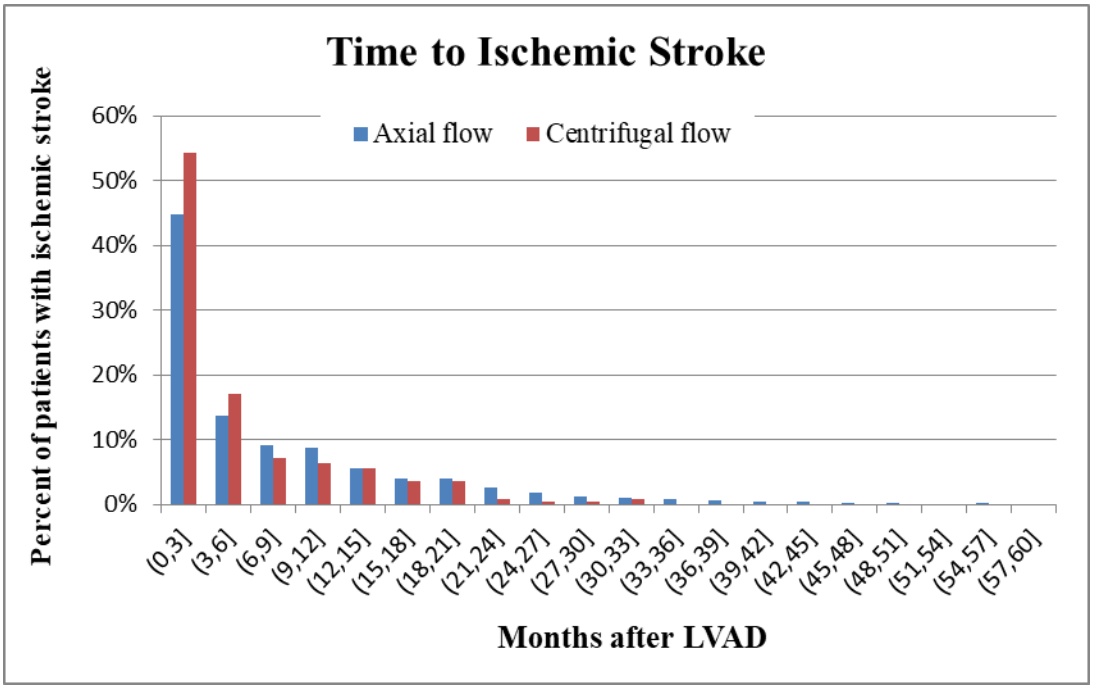


Figure 14. Time to ischemic stroke, by pump type. Percent shown is of total patients who had an ischemic stroke, by pump group.

The average patient centrifugal flow patient with an ischemic stroke was younger (56 vs 60 years of age, p-value < 0.0001), bridge to transplant (92% vs 35%, p-value < 0.0001) and more likely to be INTERMACS profile 2 (38% vs 27%, p-value = 0.0040). All other descriptive factors had no statistically significant differences between pump groups (Table 9).

Table 9. Comparison of Patient Features by Adverse event of Ischemic Stroke

		Patients with Ischemic Stroke after CF-LVAD				
Characteristic		**Axial Flow, n = 685**		**Centrifugal Flow, n=252**		**p-value**
Age	Mean (StdDev)	60	(11)	56	(10)	*< 0.0001*
Gender	Male	175	100%	58	100%	0.4715
Race	White	490	71%	163	64%	0.2543
	African American	145	21%	57	23%	0.6643
	Other	51	7%	33	13%	0.0084
Device Strategy	BTT	241	35%	233	92%	< 0.001
	DT	435	64%	19	8%	< 0.001
	Other	9	1%	0	0%	0.0688
Blood Type	O	306	45%	119	47%	0.5518
	A	265	39%	82	33%	0.1313
	B	85	12%	35	14%	0.5592
	AB	22	3%	14	6%	0.0994
NYHA	II	1	0%	1	0%	0.4605
	III	109	16%	46	18%	0.4105
	IV	534	78%	193	77%	0.7833
INTERMACS Profile	1	122	18%	41	16%	0.5996
	2	187	27%	95	38%	0.0040
	3	249	36%	82	33%	0.3381
	4	106	15%	24	10%	0.0260
	5	13	2%	7	3%	0.4099
	6	6	1%	1	0%	0.4515
	7	2	0%	2	1%	0.2965

Data is shown as the total in each category and percentage of total, with comparison between axial and centrifugal flow pump patient groups using two-way z-test. DT, Destination Therapy; BTT, Bridge to Transplant (includes patients listed and not-yet listed).

Axial flow model

Between 2012 and 2016, 9,159 patients received axial pump CF-LVADs, 307 (3.4%) of whom had ischemic stroke within 3 months after implant. The axial pump model structure and parameters were constructed using a training set of 7,327 patients, 243 of whom (3.3%) had ischemic stroke within 3 months.

Of the 247 pre-implant variables that went into feature selection, 53 were identified as potential predictors using information gain and hill climbing methods. These were used to build both TAN and NB models in GeNie. Variables were assessed for diagnostic value, low value variables were removed, and a new model was learned. This was done iteratively until performance, measured by ROC AUC from 10-fold cross validation, was optimized. The best performing axial pump model was with NB and included 32 variables (Figure 15). Model performance was measured with a test validation dataset (n=1832). For this axial flow patient population, 3.5% of patients (64 of 1832) had an ischemic stroke by 3 months. The model had a ROC AUC of 0.61 (Figure 16).

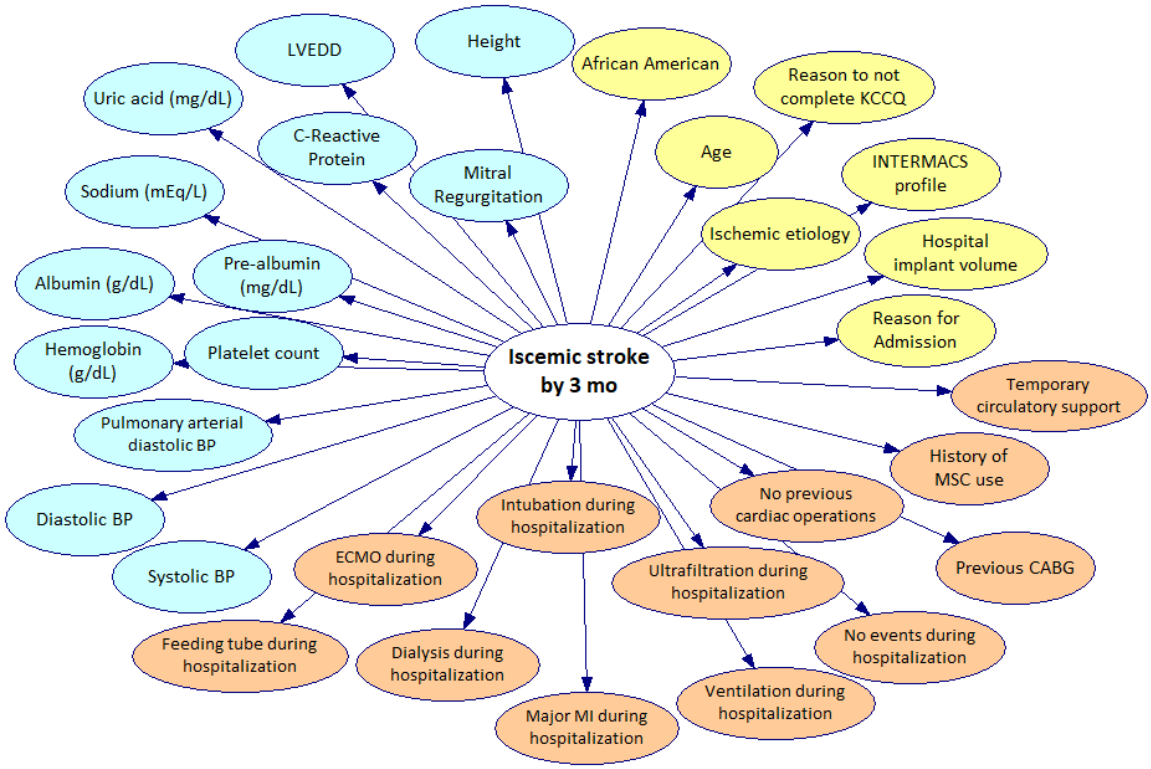


Figure 15. Tree augmented naive Bayesian network of ischemic stroke risk at 3 months with an axial flow pump

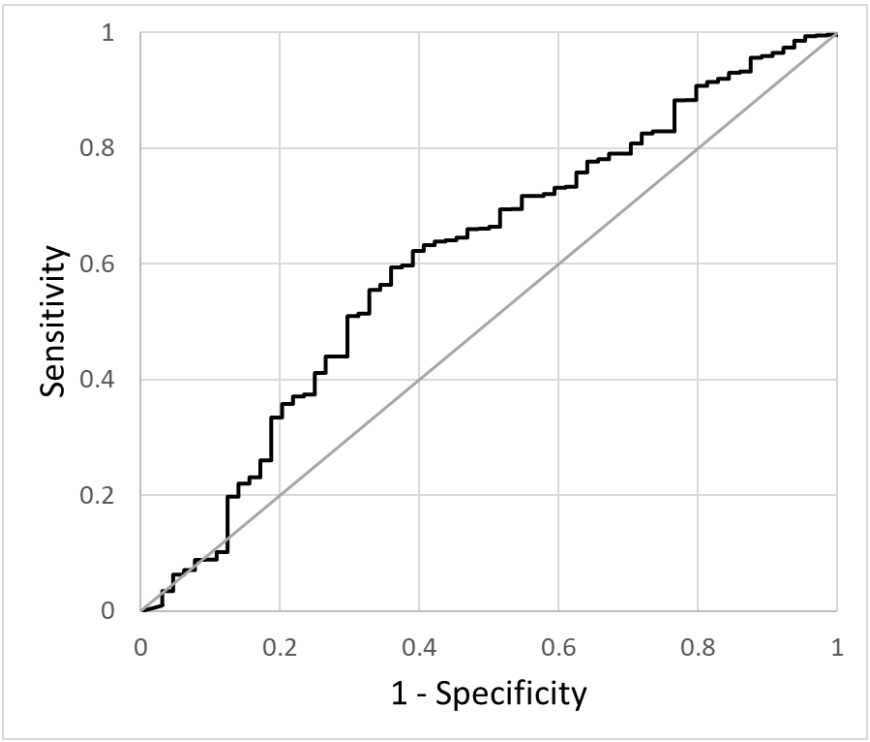


Figure 16. Receiver operating characteristic curve for model of ischemic stroke risk at 3 months with an axial flow pump

Centrifugal flow model

There were 2,909 patients implanted with a centrifugal flow pump between 2012 and 2016, 137 (4.7%) of whom had ischemic stroke within 3 months of implant.

Model structure and parameter learning was performed with a training data set of 2,377 centrifugal flow pump patients, 115 (4.9%) of whom had an ischemic stroke by 3 months.

Out of the 246 pre-implant variables that were used in feature selection, 50 were identified as top predictors using information gain and hill climbing. As with the axial flow pump model, the selected variables were used to build both TAN and NB models in GeNie. Variables were assessed for diagnostic value, low value variables were removed, and performance was measured with 10-fold cross validation until ROC AUC was optimized. The

best performing model was achieved using TAN and included 36 variables (Figure 17). Model performance was measured with a test validation dataset (n=582). In this dataset, 3.8% of patients with a centrifugal flow pump (22 of 582) had an ischemic stroke by 3 months. The model had a ROC AUC of 0.64 (Figure 18).

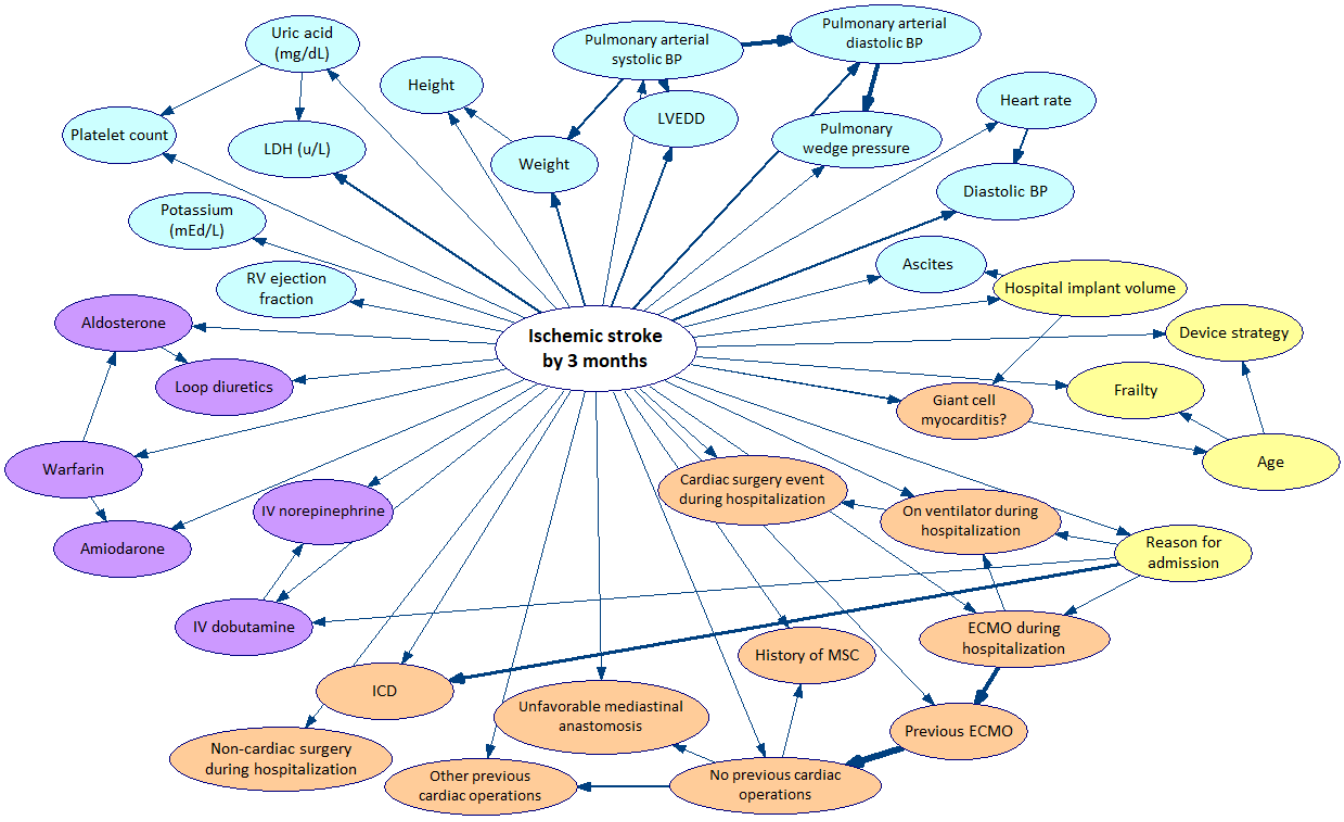


Figure 17. Tree augmented naive Bayesian network of ischemic stroke risk at 3 months with a centrifugal flow pump

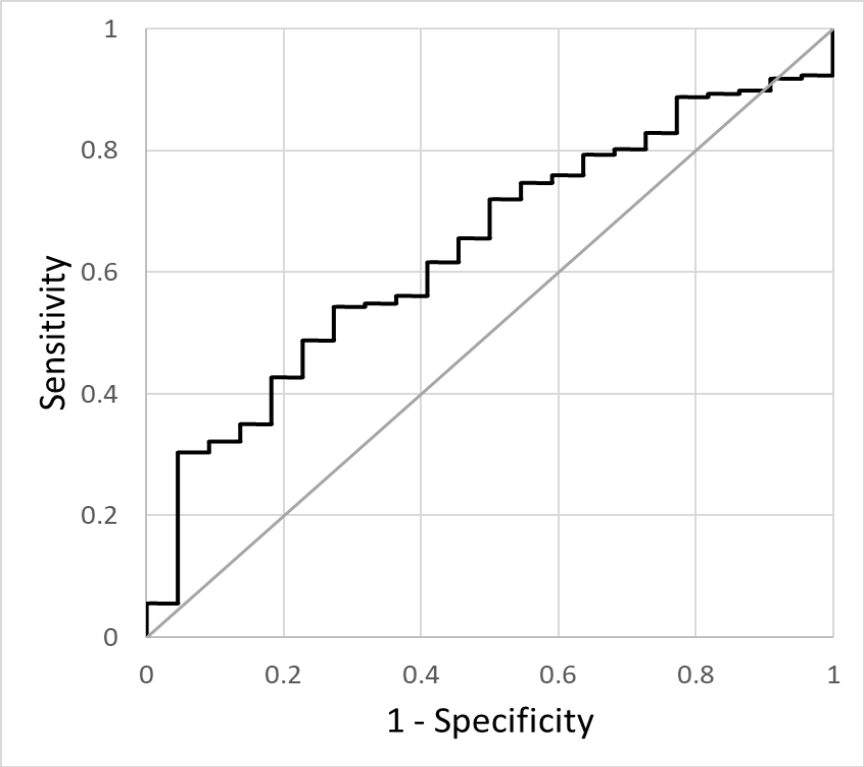


Figure 18. Receiver operating characteristic curve for model of ischemic stroke risk at 3 months with a centrifugal flow pump

Comparison of Axial and Centrifugal Flow Model Variables

Key variables in each of the predictive models are captured in Table 10.

Top predictors for the axial pump ischemic stroke model were: elevated C-reactive protein levels, elevated uric acid, previous use of temporary circulatory support, small left ventricular end diastolic diameter, and being too sick to take the Kansas City Cardiomyopathy Questionnaire (KCCQ). Both admission due to myocardial infarction (MI) and major MI were predictive of higher stroke risk. Demographic features affecting stroke risk were old age, female sex, and African American race.

The most predictive variables for the centrifugal pump model were: elevated uric acid, destination therapy (DT), hospital implant volume over 50 per year, no use of loop diuretics, and short height. Unique to this model is the risk factor of giant cell myocarditis, presence of ascites, and unfavorable mediastinal anastomosis. The only demographic factor affecting stroke risk of patients with centrifugal flow pumps was old age.

Out of the 67 features in the two models, 13 are in both models and 41 are unique. The main differences between the two are more medication variables in the centrifugal flow model (norepinephrine, loop diuretics, warfarin, aldosterone, dobutamine, and amiodarone). In the axial flow model, there are more events during hospitalization risk factors (dialysis, intubation, feeding tube, ultrafiltration, and major MI).

Table 10. Variables by diagnostic value in the axial and centrifugal pump patient populations

Axial Flow		Centrifugal Flow	
Variable	**Diagnostic Value**	**Variable**	**Diagnostic Value**
C-reactive protein	0.019	Uric acid	0.025
Uric acid	0.015	Device strategy (DT)	0.023
Temporary circulatory support	0.014	Hospital implant volume	0.017
LVEDD	0.014	Loop diuretics	0.015
Reason for not taking KCCQ	0.013	Height	0.014
Pre-Albumin	0.012	IV Norepinephrine	0.013
On ventilation during hospitalization	0.011	LVEDD	0.013
Admission due to MI	0.01	Giant cell myocarditis	0.012
Intubation during hospitalization	0.01	Previous cardiac operations?	0.012
Major MI during hospitalization	0.009	On ventilation during hospitalization	0.012
Platelet count	0.009	ECMO during hospitalization	0.012
Ischemic etiology	0.008	History of MCS	0.011
Previous cardiac operations?	0.008	Previous ECMO	0.01
Previous CABG	0.008	Weight	0.01
On ECMO	0.008	Potassium	0.01
Sodium	0.008	Pulmonary arterial systolic pressure	0.01
Pulmonary arterial diastolic pressure	0.008	Pulmonary arterial diastolic pressure	0.01
INTERMACS profile	0.007	Age	0.009
Albumin	0.007	Pulmonary wedge pressure	0.009
Diastolic blood pressure	0.007	Frailty	0.008
Hospital implant volume	0.006	Ascites	0.008
Dialysis during hospitalization	0.006	Non-cardiac surgery during hospitalization	0.007
Events during hospitalization?	0.006	LDH	0.007
Ultrafiltration during hospitalization	0.006	Unfavorable mediastinal anastomosis	0.006
Feeding tube during hospitalization	0.006	Admission due to MI	0.006
History of MCS	0.006	Cardiac surgery during hospitalization	0.006
Mitral regurgitation	0.006	Right ventricular ejection fraction	0.006
Height	0.006	Heart rate	0.006
Systolic blood pressure	0.006	Warfarin	0.005
Hemoglobin	0.005	Platelet count	0.005
Age	0.004	Current ICD	0.004
African American	0.004	Aldosterone	0.004
		Diastolic blood pressure	0.004
		IV Dobutamine	0.003
		Amiodarone	0.003

LVEDD, Left Ventricular End Diastolic Diameter; KCCQ, Kansas City Cardiomyopathy questionnaire; MI, Myocardial infarctions; CABG, Coronary arterial bypass graft; ECMO, Extracorporeal membrane oxygenation; MCS; Mechanical circulatory support; DT; Destination therapy; IV, intravenous; LDH, Lactate dehydrogenase; ICD, implantable cardioverter defibrillator.

Comparison of models to contemporary risk score

To compare the models' performance to a published risk score, we evaluated the same test data sets with the INTERMACS stroke score [76]. For the both the axial flow and centrifugal flow pumps, the risk score is not significantly associated with the rate of ischemic stroke (Figure 19). Evaluating the risk score by ROC, centrifugal ROC AUC is 54.8% and axial is 58.8% (Figure 20).

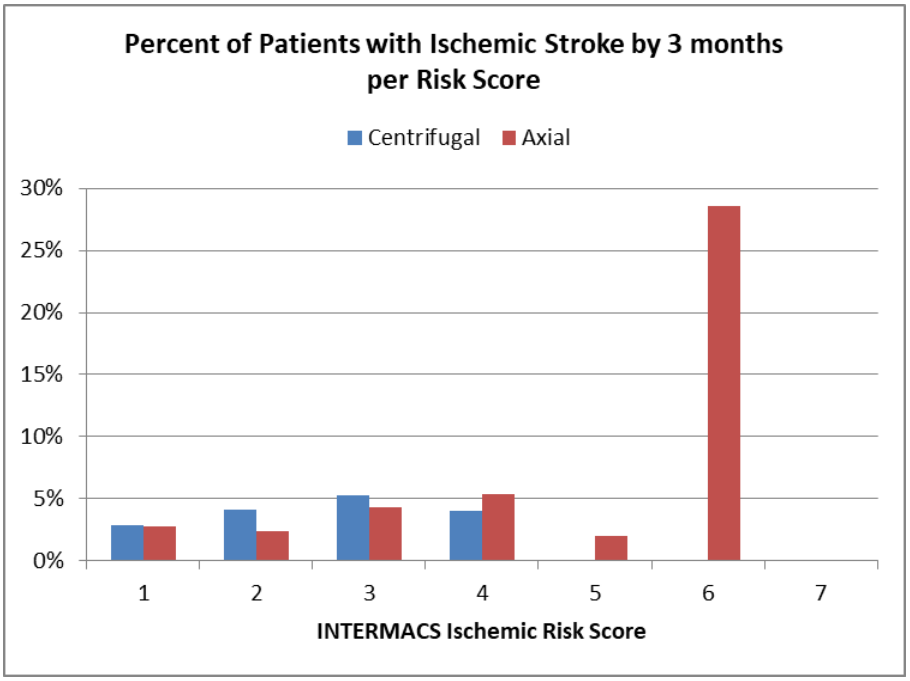


Figure 19. INTERMACs ischemic stroke risk score patient stroke incidence discrimination

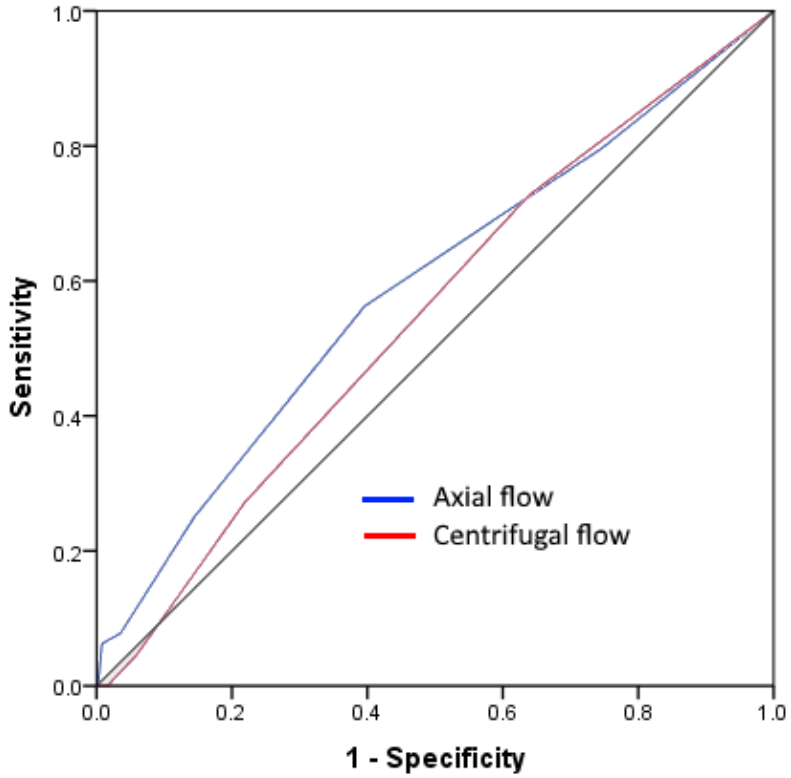


Figure 20. Receiver operating characteristic curves for axial and centrifugal flow pumps using the INTERMACS ischemic stroke risk score

2.3.4 Discussion

Patient factors affecting risk of ischemic stroke can be identified before CF-LVAD implant, and the factors driving ischemic stroke risk type differ by type of device being implanted.

The two device types being compared in this study were axial and centrifugal flow pumps. At the time of this data collection, only axial flow pumps had been approved for destination therapy (DT). Patients who are DT are usually older or have co-morbid conditions, therefore it was not surprising that the axial flow patients were significantly older than the

centrifugal pump patients. Age is associated with increased ischemic stroke risk in both the predictive models generated in this study and in the literature for both VAD and non-VAD populations [85].

Despite patients being a higher age over all, axial flow pumps had a lower overall risk of ischemic stroke by 3 months compared to centrifugal flow pumps. This difference has been noted in other studies, where there were more events per patient year of ischemic stroke in HVAD (the only centrifugal pump approved at the time of this study) than in HeartMate II pumps (the only axial flow pump approved at the time of this study) [85]. This difference in incidence was not seen with hemorrhagic stroke.

Pre-implant blood pressure has been identified as a modifiable factor that can affect risk of ischemic stroke [86], and can be successfully managed with hypertensives [87]. In the most recent INTERMACS report, a systolic blood pressure over 120mmHg was one of the key components of the predictive ischemic risk score [76]. Blood pressure is included in the axial flow predictive model in the form of systolic blood pressure, diastolic blood pressure, and pulmonary arterial diastolic pressure (PADP). Interestingly, of the three measures, PADP had the largest impact on risk prediction. In the centrifugal pump patient model, the blood pressure measures that drive risk are: Pulmonary arterial systolic pressure (PASP), PADP, pulmonary wedge pressure, and diastolic blood pressure.

Admission to hospital due to myocardial infarction (MI) was a driver of ischemic stroke risk in both pump models, with major MI during hospitalization also being a factor for risk in the

axial pump model. MI causing an anterior infarction has been associated with increased risk of developing ventricular thrombi, which may dislodge to cause stroke [88]. Additionally, the hemodynamic change and inflammatory response to infarction may factor into subsequent stroke risk [89, 90].

Uric acid has been widely studied as a risk factor associated with acute stroke in the non-LVAD population [91-93], though whether it is an independent risk factor or a marker of atherosclerotic disease is not fully understood [94]. It is a predictive factor in both the axial and centrifugal pump models. Uric acid is also connected to platelet count and LDH in the centrifugal pump TAN model, suggesting that it is an indicator of associated disease and not just a factor on its own.

Elevated C-reactive protein (CRP) was the primarily predictor for ischemic stroke in patients with an axial flow pump. CRP is a marker of inflammation and has been shown to be a strong predictor of ischemic stroke, MI, and death [95].

In both pump models, interventions during the hospitalization before CF-LVAD implant drive the risk of ischemic stroke. This includes dialysis, intubation, feeding tube, ultrafiltration, and major MI in the axial pump model, ventilation and ECMO in both models, and cardiac or non-cardiac surgery in the centrifugal pump model. All the interventions are invasive and carry the risk of tissue inflammation and infection, as well as indicating an overall poorer health at the time of hospitalization for the patient.

Use of loop diuretics and aldosterone are associated with lower ischemic stroke risk in patients with centrifugal flow pumps. This is in line with the connection between hypertension and stroke risk, were patients not being managed with these are at greater risk of becoming hypertensive after CF-LVAD implant. Use of amiodarone and IV dobutamine were also

associated with lower ischemic stroke risk. Atrial fibrillation has been highly associated with ischemic stroke risk [81, 96], though was not a selected feature in either risk model. The use of an anti-arrhythmic such as amiodarone may decrease stroke risk by minimizing occurrence of arrhythmias in these patients. IV norepinephrine was associated with increased ischemic stroke risk, potentially due to its vasopressor effect.

Device strategy is a significant predictor for outcomes in the centrifugal flow pump model, with DT being the strategy associated with higher risk. Centrifugal flow pumps were not approved for DT at the time of this study, so the 2% of patients with them that were DT were technically off label. This rare off-label use is associated with ischemic stroke risk, however that may be due to other factors that caused the patients to be on off-label in the first place, such as very small body size and older age (the DT-approved axial flow pump is larger and may not be tolerated as well in very small patients.) Patient size by height and weight was also a predictor of centrifugal pump ischemic stroke, with smaller patients having higher risk.

Gender was not a predictor in either model, despite it being commonly reported as a risk factor in other literature [76, 97]. In one study, the relationship between the higher risk of females having ischemic stroke was characterized in context of their smaller size, which was a predictive factor in our centrifugal pump model, and their risk of thromboembolism due to the use of oral contraceptives or hormones [98].

Comparing the predictive strength of our resulting models to the current published risk score, we demonstrated superior performance. The current predictive performance of our models

at 61% and 64% ROC AUC for the axial flow and centrifugal pump models, respectively, shows moderate predictive power.

This study was limited by the large amount of missing data, particularly for lab values, which ranged from (5-72% missing). Data is manually entered into INTERMACS and may be subject to errors in entry. However, this study is the largest of its kind to compare outcomes by pump type and the first to derive independent predictive models for ischemic stroke. Future work will explore additional time points, including early (within 2 weeks of implant) stroke and late stroke 12 months, as well as risk factors for hemorrhagic stroke, the less common but deadlier of the stroke types in LVAD patients.

2.3.5 Conclusion

By using a Bayesian approach, we explored pre-implant factors that are predictive for ischemic stroke and their relation to the type of pump. Patients on centrifugal flow pumps have a higher ischemic stroke risk, but selection of the right patients can mitigate this increased risk. Factors driving overall risk include blood pressure, which can be pre-operatively managed, and incidences of invasive interventions. These models may be utilized to identify optimal candidates for LVAD implantation that have a lower risk of ischemic stroke.

3.0 AIM 2: VERIFICATION OF MORTALITY MODELS AND PHYSICIAN USE CASE FOR LVAD DECISION SUPPORT

3.1 RETROSPECTIVE EVALUATION OF MORTALITY MODELS AT SINGLE IMPLANT CENTER USING COMPLETE PATIENT DATASET

3.1.1 Introduction

Heart failure is a chronic, progressive condition that affects over 6 million Americans. It is characterized by a decline in function of the heart to pump enough blood to perfuse the body [1]. As the condition progresses, treatments may escalate from dietary modification and oral medications to intravenous drug delivery and surgical interventions, such and mechanical heart-assist pumps and heart transplantation [2]. Heart transplant is the gold standard treatment for end stage heart failure; however, donor heart supply is limited and not all patients are eligible for transplant, due to their age, comorbid conditions, or lifestyle choices. As an alternative, advanced heart failure patients may receive a durable left ventricular assist device (LVAD) as a bridge to transplant (BTT) or as a destination therapy (DT) [10].

LVADs can increase quality of life and improve patient survival [7, 8], but also require significant changes in daily life, investment of time and money, and introduce risks of major adverse events [9]. These tradeoffs underscore the importance of careful patient selection, for which predictive models can serve as an important component of risk assessment.

We recently published models to predict post-LVAD mortality at 1, 3, and 12 months after implant [75] using the data from the Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS), the largest registry of retrospective LVAD patient data in the United States [7]. The models were developed using Bayesian analysis and validated with a subset of registry data that was withheld from the model derivation. While use of the large registry dataset provides a robust model, it obscures institution-dependent differences in patient selection, care, and outcomes. Use of a personalized decision support tool in a 'real world' clinical setting is necessary to understand its applicability at individual institutions.

The INTERMACS registry includes a large population (n = 20,216) of LVAD patients but suffers from missing data and entry errors. Because data is entered manually by LVAD coordinators and nurses at each participating site, there are inevitable errors such as misplaced decimal points, incorrect units, and skipped fields. No data checks are imposed on the data entry process, therefore any information that is unusual or out of range is not flagged. When we developed the Bayesian models with these data, out-of-range or illogical data entries were censored; however, missing data was left as-is, instead of being imputed, to minimize over-fitting. The extent to which these issues affect the performance of the Bayesian predictive models is unknown; therefore, a carefully checked and evaluated dataset from a single clinical site was used to measure model performance.

This study was undertaken to establish the performance of our Bayesian models for LVAD mortality at a single institution with a complete, retrospective patient data set. The goal of this work was to prove the utility of the models for use in prospective patient risk assessment.

3.1.2 Methods

Data acquisition and cleaning

We acquired site-specific INTERMACS data for 100 consecutive patients who received a CF-LVAD at Allegheny General Hospital (AGH) between 2014 and 2015. A data sharing agreement was established between Carnegie Mellon University (CMU) and AGH to assure the security of protected health information in this study. This study was approved by CMU and AGH's review boards for biomedical research (IRBs).

The time-period was selected to include records with at least 1 year of follow up data. The data was organized into three categories: Pre-Implant, Post-Implant, and Event. Missing or illogical data (outside of feasible range or conflicting with other entries) was manually identified and checked by a data coordinator. Data elements that were designated as "unknown" or "missing" were addressed by reviewing all available patient medical records. In cases where the data could not be found, the data field was denoted as "not recorded." All units for continuous variables were also checked. Once all 100 patients were verified by the coordinator at AGH, the data set was sent to CMU for analysis.

Data pre-processing

Pre-implant continuous data were binned into groups as previously described [75]. Mortality outcomes were assigned to each patient using the Event data for each of the three time points: 1, 3, and 12 months post LVAD.

Model validation

The complete AGH data sets were used to measure the Bayesian mortality model performance for each time point, using test validation in GeNie (BayesFusion, Pittsburgh, PA).

3.1.3 Results

Data cleaning revealed 9% of all pre-implant information (2704 out of 28500 possible fields) was missing or out of range in the patient records. After data cleaning, this was reduced to 4% (1184) fields that were confirmed as not recorded.

The patient cohort at AGH was similar to the overall INTERMACs population in terms of patient age and gender (Table 11.) The main statistical differences between cohorts were the proportion of INTERMACS profile 2 and 3 patients. This difference indicates a sicker patient population in the AGH cohort than the INTERMACS patients overall.

Table 11. Patient cohort comparison

		AGH Patients (n=100)		INTERMACS Patients (n = 10,277)		
Characteristic		**n**	**%**	**n**	**%**	**p-value**
Age	Mean (std)	56.2	(12.7)	56.9	(13)	*0.59197*
Gender	Female	27	27%	2225	22%	*0.19706*
	Male	73	73%	8044	78%	*0.20408*
NYHA	I	0	0%	12	0%	*0.72786*
	II	2	2%	88	1%	*0.2187*
	III	55	55%	1850	18%	*< .001*
	IV	26	26%	7816	76%	*< .001*
	Unknown/ Not documented	17	17%	511	5%	*< .001*
INTERMACS	1	20	20%	1671	16%	*0.3125*
	2	48	48%	3548	35%	*0.0048*
	3	14	14%	3318	32%	*0.0001*
	4	15	15%	1340	13%	*0.56192*
	5	0	0%	230	2%	*0.13104*
	6	2	2%	58	1%	*0.0601*
	7	1	1%	41	0%	*0.34722*
	Unknown		0%	71	1%	*0.40654*
Ischemic Etiology	No	48	48%	5640	55%	*0.16758*
	Yes	52	52%	4637	45%	*0.16758*
Device Strategy	BTT Likely	67	67%	5261	51%	*0.00164*
	BTT Unlikely	5	5%	267	3%	*0.13362*
	DT	25	25%	4658	45%	*< .001*
	Other	3	3%	91	1%	*0.02642*

NYHA, New York Heart Association class; INTERMACS, Interagency Registry for Mechanically Assisted Circulatory Support; BTT, Bridge to transplant; DT, destination therapy.

One month after implant, 4 (4%) of the 100 AGH patients had died. Then NB mortality model correctly predicted 3 out of the 4 deaths (75%) and predicted 87 out of 96 alive patients (91%), using a threshold of 50%. The ROC AUC was 78%. This is better performance than the original model validation of 70% ROC AUC (Figure 21).

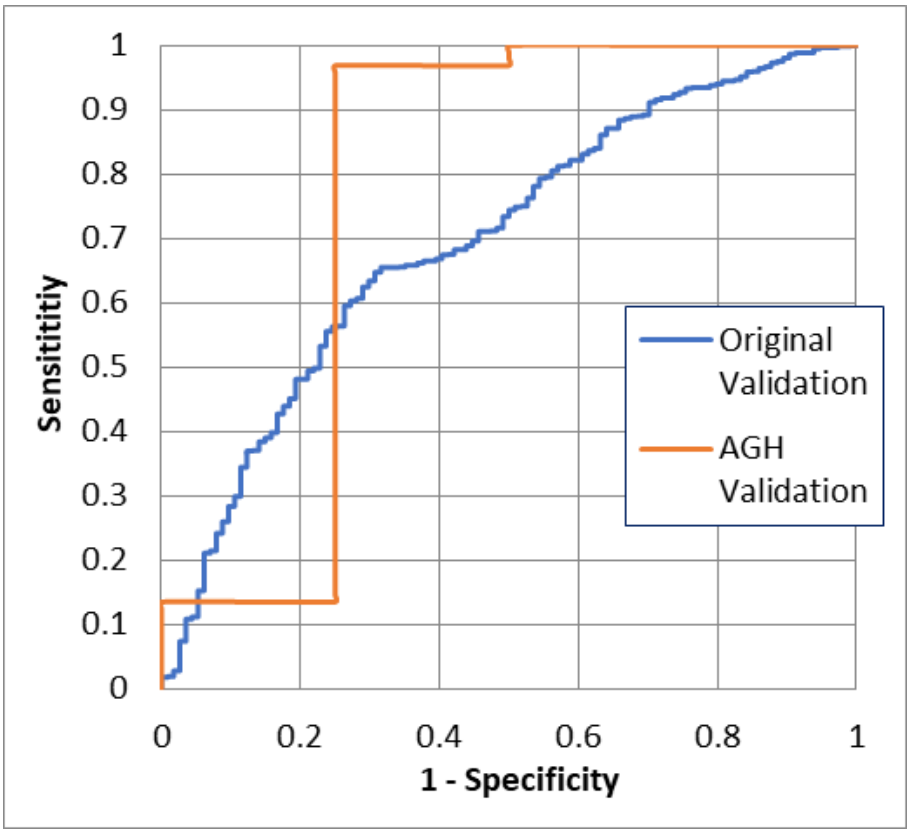


Figure 21. ROC curves for 1-month mortality from original and AGH-specific validation

At three months after implant, 8 (8%) of the 100 patients had died. The NB mortality model correctly predicted 4 of the 8 deaths (50%) and 83 of the 92 living patients (90%), using a determination threshold of 50%. The ROC AUC for the model performance was 76%. This is superior to the original model test validation of 71% (Figure 22).

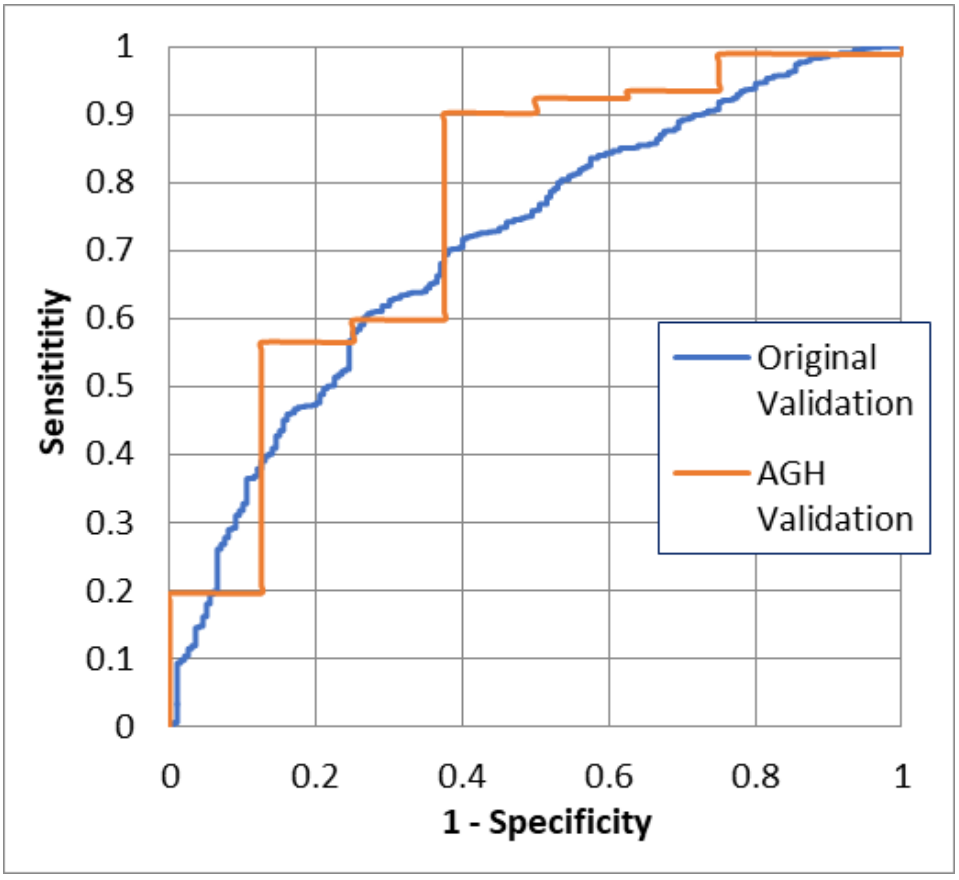


Figure 22. ROC curves for 3-month mortality from original and AGH-specific validation

By twelve months after implant, 18 (18%) of the 100 patients had died. The NB mortality model correctly predicted 6 of the 18 deaths (33%) and 73 of the 82 living patients (89%), using a determination threshold of 50%. The ROC AUC for the model performance was 75%, which was better that the original model validation of 69% (Figure 23).

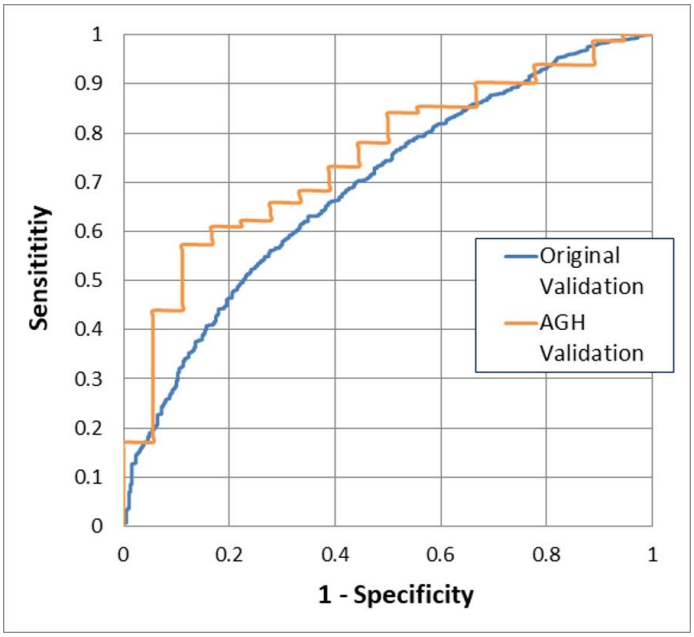


Figure 23. ROC curves for 12-month mortality from original and AGH-specific validation

3.1.4 Discussion

We had previously reported ROC AUCs of 70%, 71%, and 69% for Bayesian mortality predictions at 1, 3 and 12 months post-LVAD implant with a validation cohort from INTERMACS. All three mortality models performed better in the AGH patient dataset than in the INTERMACS validation cohort. The AGH patients had similar demographics to the patients in the model learning dataset, however there were significantly more patients with severe heart failure, as indicated by the percentage of patients with INTERMACS 2 classification.

One of the explanations for the better performance with AGH patient data is the greater proportions of severe heart failure patients. Recent analysis by our group has demonstrated that the Bayesian mortality models perform better in the more severe heart failure populations (Kanwar et al, in preparation). The INTERMACS profile 1 patient group had ROC AUCs of

71% for each of the 1, 3, and 12-month time points and the profile 2 patient group had ROC AUCs of 74%, 75%, and 70% for the time points. These are the same or better performing than the validation with all patients. The difference in performance may be attributed to the greater proportion of data available for the sicker patients. Bayesian models are derived using prior probabilities and thus are more accurate when applied to patient populations that comprise a greater percentage of the derivation cohorts. Another reason for the difference may be that the factors that increase a patient's risk of dying (such as recent cardiac surgery, advanced age, and dialysis) are easy to capture in the dataset, while it is much harder to identify and quantify features that predict a patient's good health and survival.

Before using the Bayesian mortality model predictions in clinical practice at an implant center, it is essential to verify their performance on that center's specific patient population. This is especially important given the influence of institutional experience on outcomes. This is illustrated by the Heartmate II Risk Score, which includes institution implant volume as a statistically significant predictor for mortality outcomes[44]. Additionally, an assessment of implant center volume on one-year mortality of destination therapy (DT) patients found that low volume centers had a higher mortality rate [99]. Similar relationships have been reported for transplant graft survival[100] and right heart failure-associated mortality[101]. Since AGH is an experienced, high volume implant center, the models may perform better there than in a lower implant volume institution.

The data cleaning step at AGH did not create a significant difference in missing data, with the majority (56%) of missing data elements identified as not recorded. The Bayesian method of modeling is robust to missing information when making predictions, and this is shown to be true by the resulting ROC AUCs. Whether having no missing data would improve the

model performance remains unknown. However, it is unlikely that any institution can have a value for every possible patient variable, making these models attractive for real world use.

The models assessed in this analysis have been made available at app.myCORA.org, as part of the Cardiac Outcomes Risk Assessment (CORA) decision support tool for physicians (Figure 24.) This tool will now be prospectively evaluated with the multidisciplinary team at AGH to measure its impact on patient selection and decision making. Predictive models for post-LVAD adverse events are being developed to add to the CORA tool (e.g., ischemic stroke, major bleeding) and will be evaluated for performance with the same single center, retrospective validation methodology.

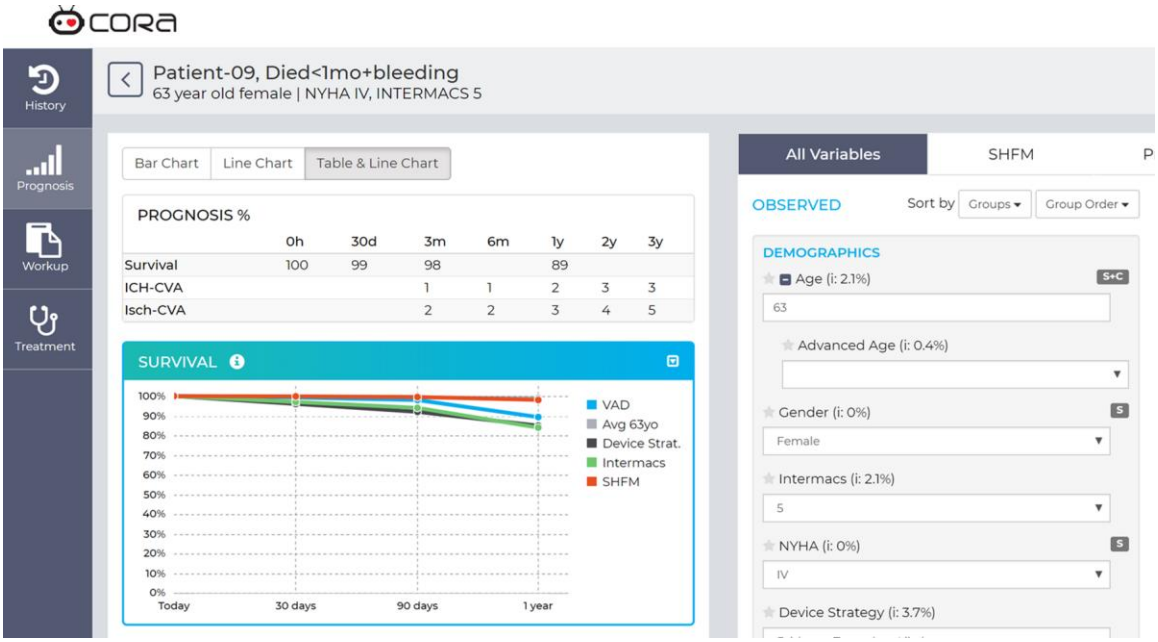


Figure 24. Screenshot of the app.myCORA.org web tool

3.1.5 Conclusion

By validating the model set at a single clinical site, performance can be demonstrated for the patient population served at that site and for the unique surgical and medical management style of the clinicians. This exercise is imperative to confirm the utility of the mortality models for clinical decision making. Future work will be to prospectively test the model performance in the AGH multidisciplinary team meeting setting, to evaluate utility in real life decision making.

3.2 PILOT TESTING THE MYCORA PHYSICIAN USER INTERFACE

If a tree falls in the woods, but no one is there to hear it, does it make a sound?

If a decision support tool gives accurate predictions on patient outcomes, but no one takes the time to use it, does it help medical practice?

3.2.1 Introduction

Design of a clinical decision support tool must be carefully considered for the tool to be used in and improve upon medical practice. The government Medicare and Medicaid electronic health record (EHR) incentive program has clinical decision support (CDS) as one of its core focus areas [24]. This prompted a guidance document for a CDS framework. Called the "CDS Five Rights", it states that CDS interventions should provide [102]:

1. the right information (evidence-based guidance, response to clinical need)
2. to the right people (entire care team – including the patient)
3. through the right channels (e.g., EHR, mobile device, patient portal)
4. in the right intervention formats (e.g., order sets, flow-sheets, dashboards, patient lists)
5. at the right points in workflow (for decision making or action)

In short, for a CDS to be effective, it must be relevant to those who use it to facilitate the right decision for the right patient at the right time.

While much research has been done in the field of LVAD decision making to address the first CDS tenet using the right information [26, 30, 44], the tenets of the right people, channel, format, and point in workflow have been largely neglected. This study addresses the issues of determining the right channel and intervention format for the myCORA decision support tool use by physicians.

This pilot study tested the usability of the current myCORA decision support interface with physicians to inform the design of a large online study for quantitative measurement of usability, interpretation, and content quality.

3.2.2 Methods

Pilot testing was performed with one of the CORA clinical collaborators. The participant was asked to access the new myCORA interface on their own computer and to share their screen via Skype. The participant was asked to think aloud as they responded to questions about the interface and explain what they were doing and why. A preliminary script of questions was used to guide the participant through initial interface thought and two exercises, using patient information already in the myCORA tool and then entering information for a past patient. The screen of the user was video recorded with MouseFlow, a mouse tracking software, and audio recorded to augment note-taking.

Responses were analyzed for themes in response in two main categories: layout and content.

Mouse tracking heat maps were used to show the areas the users spent the most time one, where they had the most clicks, and where their attention was primarily focused.

3.2.3 Results

Access

Pilot user accessed the myCORA app through her personal laptop computer with the Safari internet browser.

Initial Feedback

Responses followed a natural flow from left to right. The most time was spent on reading and making sense of the content of the prognostic graphs, followed by scanning the variable input options. Themes covered in the initial feedback are summarized in Table 12.

Table 12. Responses to myCORA layout

Layout Theme	**Example response**
Page layout	"Icons on the left are easy to read, nice, understandable"
Data presentation	"Table is busy. Can see that it's percentages but would like a better visual."
Model information	• 'Avg age 60', interpreted as a healthy control patient • "Why are the x-axes different? This is misleading, because I naturally compare them to each other"
Wording	• "Wording confusing for 'patient does not want transplant'– should be 'does patient want a transplant, yes/ no?'" • "What do 'scenarios' mean?" (in the model legend)
Interaction	• Variables are disappearing (moving from unobserved to observed) • Trying to change variable sorting to influence overall, used second drop-down option, which is dependent on the first

First exercise (pre-entered data)

When asked to interpret the prognosis graph information, respondent answered correctly and rapidly, using the information on the prognosis table at the top of the screen.

When asked to name three un-observed variables for the demonstration patient, participant reported three titles of the variable groups, as opposed to individual variables, e.g., "laboratory values" instead of "creatinine". When asked to say what variable had the most influence on the survival outcome, participant knew to look for ranking of variables by their

influence, but originally selected the wrong dropdown box. After prompting by the interviewer, she successfully sorted by influence and reported the variable. At first, she paused to give the variable group title, but realized that the group titles had been removed when the sorting method changed.

Second exercise (own patient data)

When asked to create a new patient, participant knew to use the arrow at the back button to the main patient screen and found the "Add New" option quickly. Data for the initial patient information was entered without hesitation or question of why those items were chosen.

Asked to enter patient information to determine outcome, the participant entered 13 variables, only changing one during the process. She scrolled down the un-observed variable list, leaving in it "Group" mode. As each variable was selected it moved to the 'observed' list and effectively disappeared. This was met with frustration as some of the variables had dependents. Models were not seen to the left because she was answering variables down the page, out of view of the models. Thus, no change in model outcome was seen during data entry.

Responding to model outcome, participant correctly interpreted the graph but was surprised at the survival prediction outcome. Participant did not seek to look at the variable influences until prompted. When looking at the most influential variable, concern about outcome was somewhat alleviated. Reponses about the content are summarized in Table 13.

Table 13. Response to myCORA model content

Content Theme	**Example response**
Mortality prediction	• "That is very poor prognosis, which is surprising because the parameters I entered are fairly common for our VAD patients" • "Why is SHFM included here? What is its purpose?"
Other model predictions	Did not view or comment on

Post-exercise feedback

When asked how this tool would be used in her practice, participant said she would most want to use it to evaluate ambulatory heart failure patients on medical therapy who need intermittent IV inotropes. Regarding when in the relationship with the patient (e.g., at time of diagnosis, during LVAD evaluation, immediately before surgery) this would be, she said it was different for every patient – a patient receiving a diagnosis in the ICU on ECMO wouldn't be appropriate, but a patient earlier in their disease progression might be.

In terms of usability, the participant indicated that though she was overwhelmed with the information at first (the summary table of prognoses, in particular) she felt more comfortable navigating the tool after some time using it. At one point when going to sort variables by influence the participant said, "I remember how to do this!" Prolonged use of the tool did not impact her perception of the model content, with concerns about the predicted patient prognosis being too dire persisting throughout the exercise.

In closing, the participant said she liked how myCORA worked, but would not want to use it with patients until she better understood and believed the predictive models. It was suggested that model validation with her site's contemporary data would help achieve this.

Mouse tracking analysis

On the patient selection screen, mouse movement tracking revealed most attention spent looking through patient names and taking note of patient features, predominantly gender and NYHA level (Figure 25). The mouse action also shows tracing across a row of patient information.

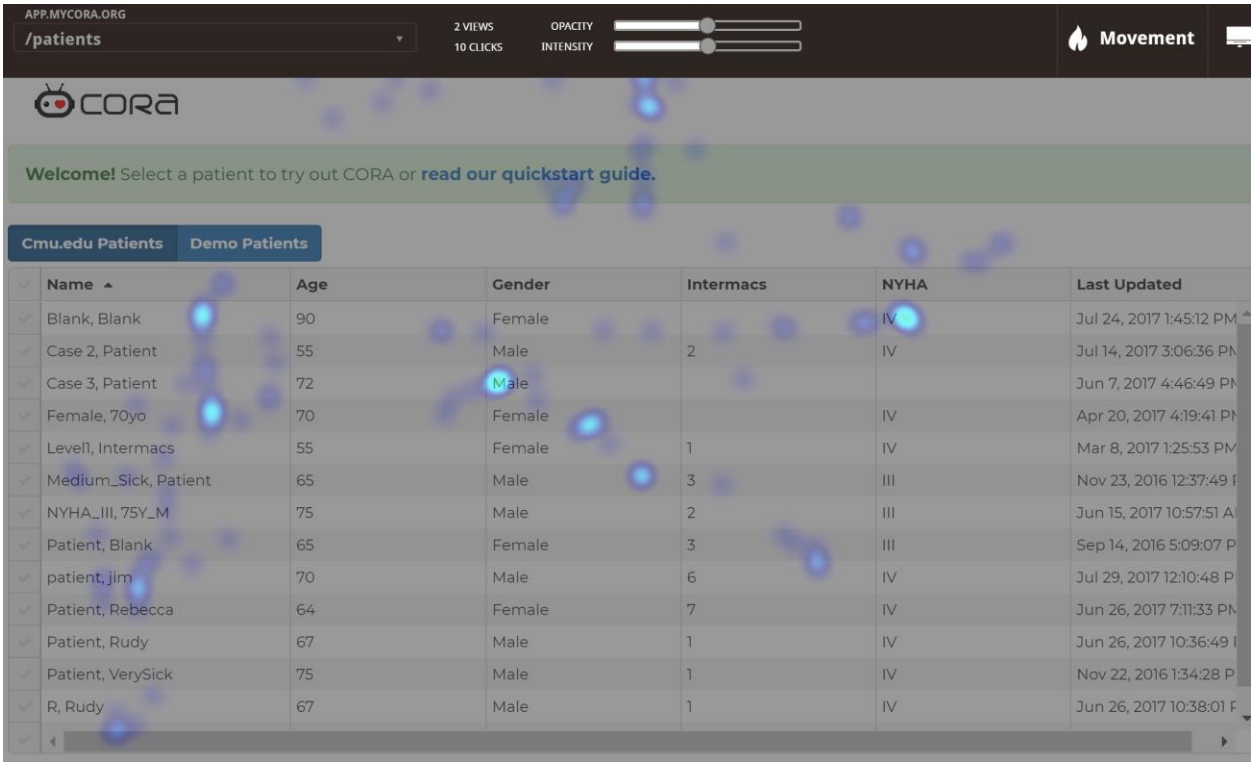


Figure 25. Heat map of mouse activity on patient selection screen

Attention tracking, which is extrapolated from mouse clicks, movements, and scrolling, shows that the patient list is the area of main focus, with least focus on the navigation options at the bottom of the page (Figure 26).

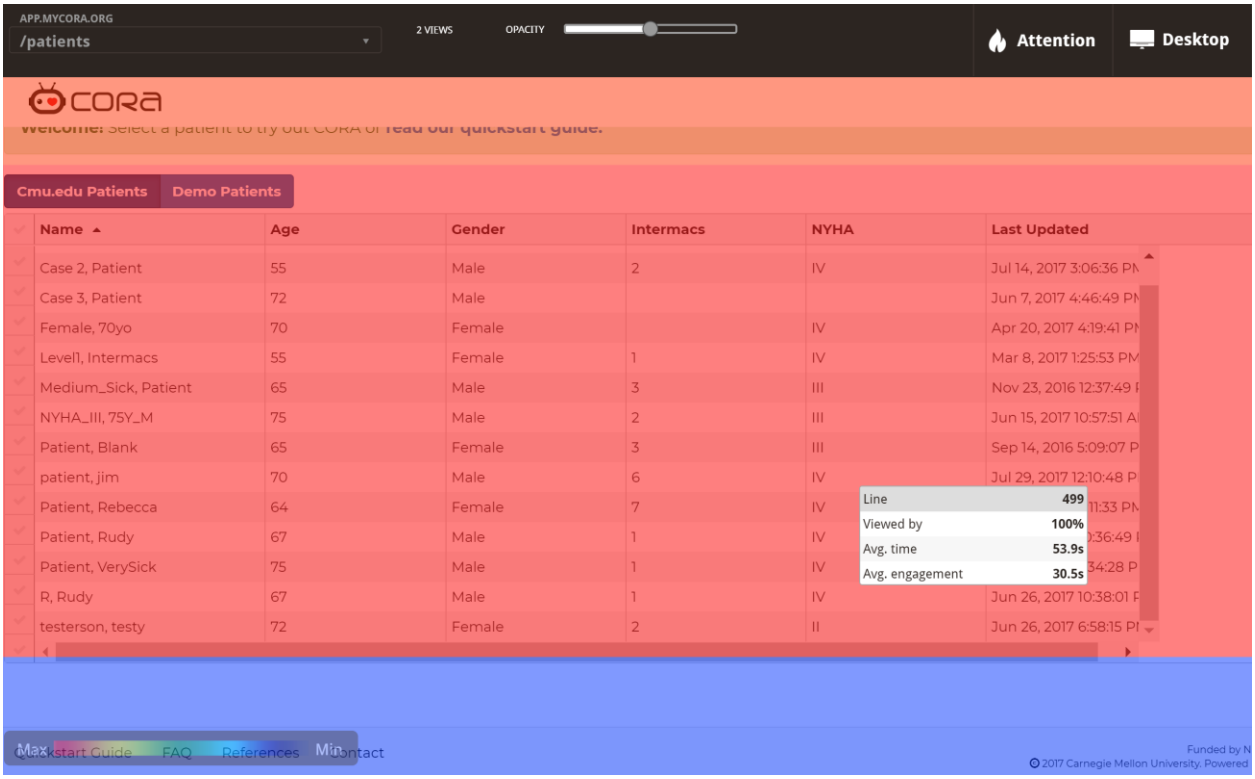


Figure 26. Attention heat map of patient selection screen

On the patient prognosis page during the user exercise, movement was concentrated around the survival model and observed variables (Figure 27). Because some of the observed variables had dependent dropdown options, movement was high over the additional fields to click on. Mouse activity decreased over the models, with no activity over the last three models.

Attention was primarily on the survival model and decreases going down the screen (Figure 28). There are 33 unobserved variables that fall below the last line of models, which received the minimum attention.

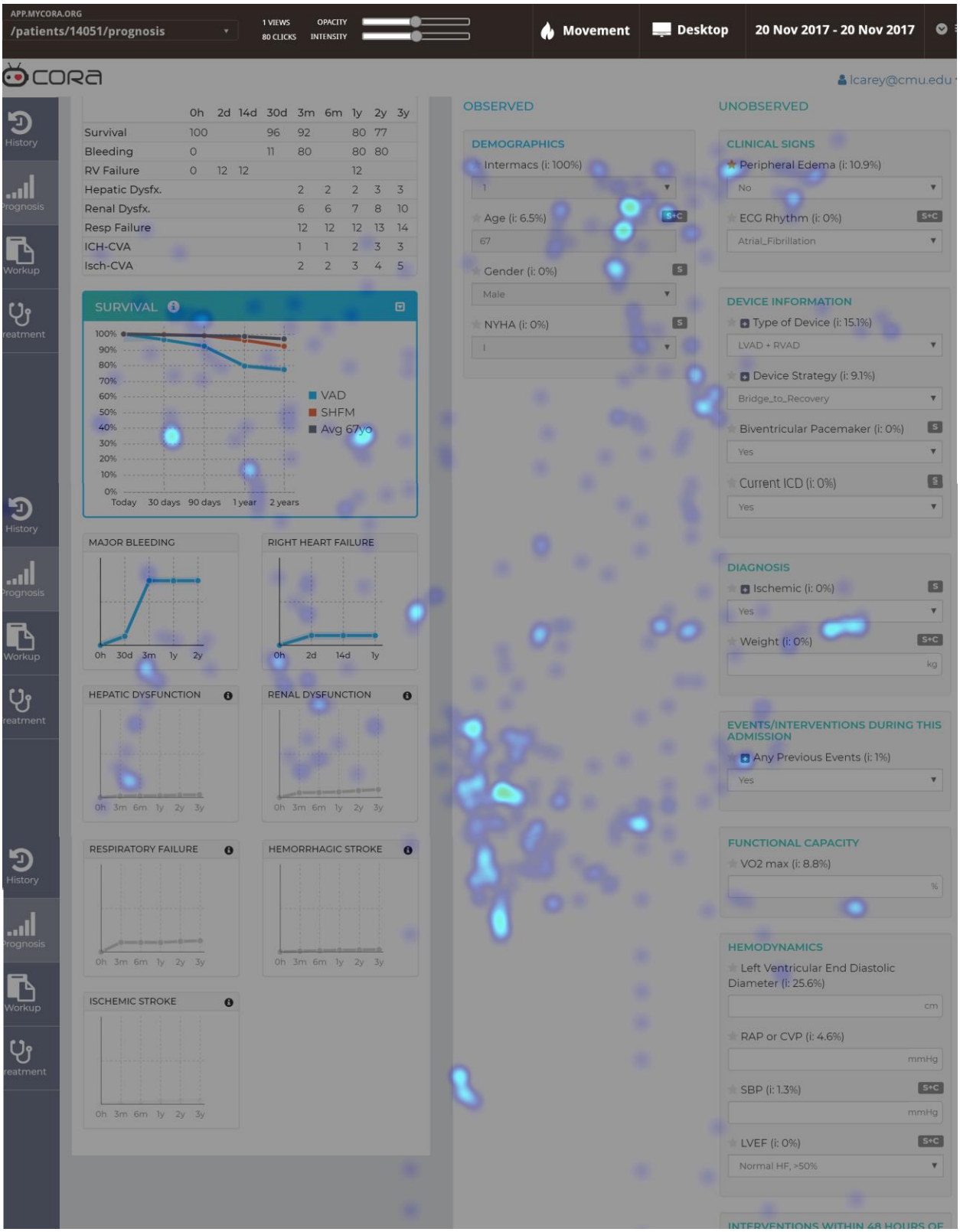


Figure 27. Heatmap of mouse activity on patient prognosis screen

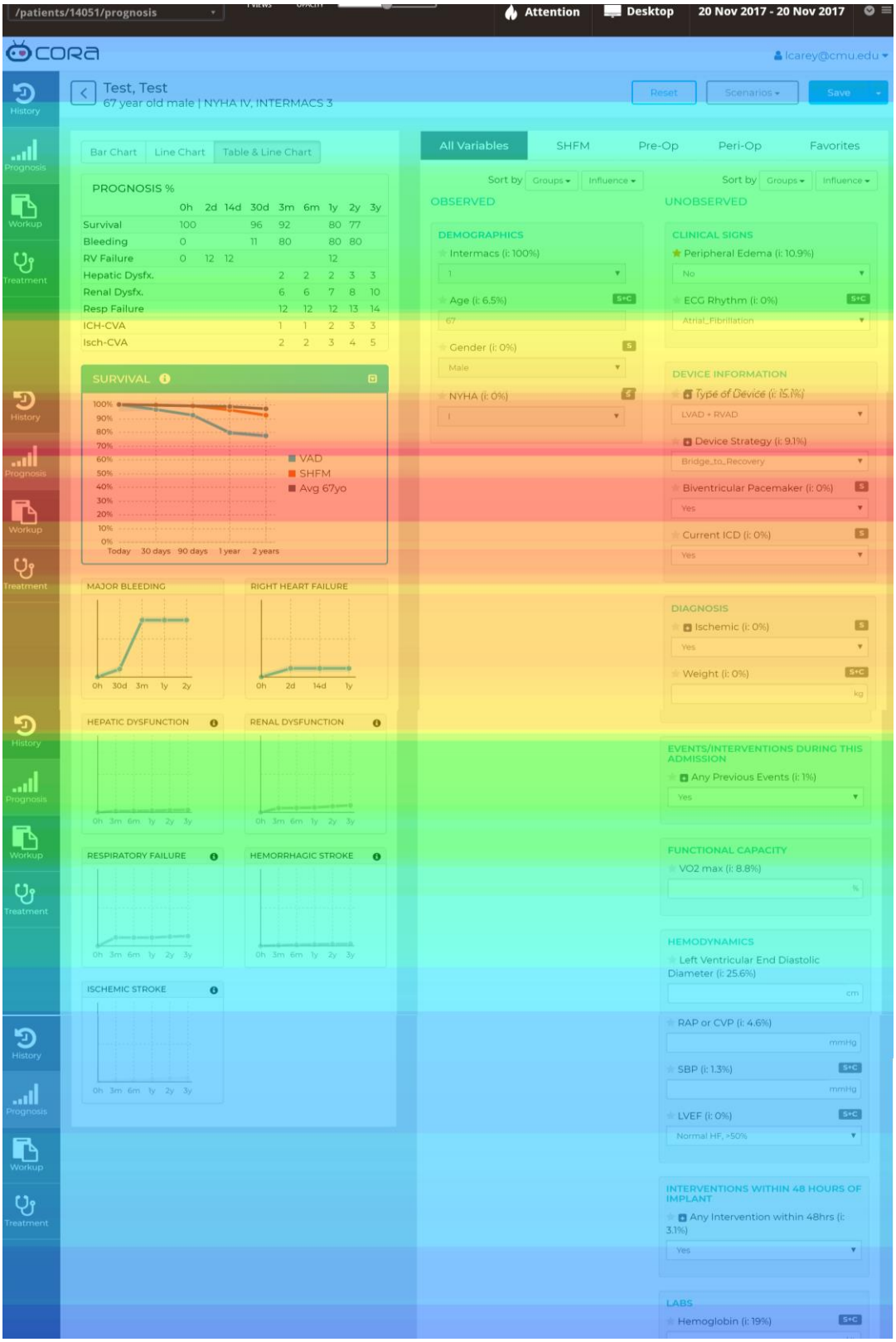


Figure 28. Attention heat map of patient prognosis screen

3.2.4 Discussion

The myCORA user interface was shown to be visually appealing but with issues in usability and model content. Main areas of use-issues were the sorting options for variables, movement of variables from un-observed to observed columns, and entering information for dependent variables. Main issues for models were lack of information about the legend labels, survival prognosis and use of SHFM, interpretation of static/in development models.

Analysis of mouse movement and attention indicated a focus on survival at the most important information, with little attention on options at the top and bottom of each page. Ways to minimize the amount of scrolling necessary to view and enter information should be explored.

The optimal time in patient disease progression to initiate use of myCORA is difficult to pin down and may vary from user to user. In this pilot, the participant wanted to use CORA to evaluate less sick patients, with the reasoning that very sick patients would receive an LVAD because there is no other treatment option. This is in contrast with feedback from collaborators who indicated they would want to use this tool to convince very sick patients that an LVAD would not be beneficial and to pursue palliative care instead. For the development of the larger user testing, a visual of patient disease progression will be incorporated to allow users to more easily conceptualized and indicate when they think the tool would be of most interest.

Overall tool use improved over time, with speed of actions improving and the participants knowing where to look for information. This pilot did not include using a ‘quick start’ guide or any educational information about the tool prior to the exercises. Introduction of a quick start guide or educational videos will be included for the large user testing to speed up the process of learning and comfort with the support tool.

To be confident with the model predictions, the participant indicated that she needed more information about the model and its performance. Future work will include validation of the model on each collaborating site's specific data (as performed with patient data from Allegheny General Hospital.) A question for the user feedback survey will be added to ask what evidence they would want to see to feel confident using the prognostic tools.

Though the myCORA online application has sections beyond prognosis, the main focus of the user was on the prognosis screen, which is the landing page after patient selection. Additional work will be needed to assess the utility of the other myCORA components (History, Treatment, and Workplan.)

The user in the pilot testing was familiar with the Bayesian modeling used to derive the prognostic results and receives financial support from the CORA grant. Her feedback may be positively biased given her familiarity and involvement with the CORA research project. However, she may also have spoken more freely given her familiarity with the interviewer. Results from this pilot trial will not be included with the following user exercise testing but instead will be used to inform and clarify the user exercise prompt.

3.2.5 Conclusion

The myCORA user interface for physicians has a pleasant visual layout but needs to be improved in terms of usability. In terms of content, models need more information about their derivation and the information being presented. MyCORA has promise to be a useful tool in physician decision making, pending layout and content improvements, and may be used in multiple points in the patient disease progression. Future work includes generating feedback from a larger audience and validating models with data from collaborating clinical sites.

4.0 AIM 3: KEY FEATURES OF PATIENT BEHAVIOR IN CONTEXT OF MEDICAL DECISION MAKING

4.1 FACTORS AFFECTING HEALTHCARE ENGAGEMENT BY PATIENTS WITH SEVERE HEART FAILURE: AN INVESTIGATION USING MACHINE LEARNING

ABSTRACT

Background: The decision to receive a durable left ventricular assist device (LVAD) to treat end stage heart failure involves understanding and weighing the risks and benefits of a highly invasive treatment strategy. These patients may have experienced a long, slow decline in health leading up to their first contemplation of receiving an LVAD. Consequently, they may exhibit a spectrum of cognitive impairment. Decision support tools can potentially help these unique patients with the LVAD decision process, but the content and presentation of information should be tailored to effectively engage these patients.

Methods and Results: A survey study of 57 heart failure patients was performed to understand their attitudes towards their health care engagement, measured by: their medical knowledge, interaction with physicians, confidence with technology and data visualization, and questions they have about their health. The survey responses were analyzed using traditional descriptive statistics and machine learning (Bayesian search, k-means clustering, and latent dirichlet allocation text analysis). Descriptive statistics showed a positive patient response to

health engagement (65%, n=37 satisfied with their involvement), interest in accessing their health record (74%, n=42) and using a prognostic tool (56%, n=32). Machine learning identified a strong relationship between the patients' numeracy and their interest in participating in their healthcare decisions. Text analysis of an open-ended question indicated an interest in education about the technical details of the LVAD (26%, n=15), a desire for personalized survival information (21%, n=12), and hesitancy to discuss their healthcare wants aloud with a non-physician staff person (25%, n=14).

Conclusions: While most patients reported interest in engaging in their healthcare, there was a subset of patients who were less interested in engaging in their own treatment decisions and less confident in understanding both health information and data visualizations. Design of a decision support tool for LVAD patients should consider a spectrum of ability and desire to understand health information and data visualization.

4.1.1 Introduction

Shared decision making between patients and their healthcare provider is a well-recognized goal throughout healthcare [6, 103, 104]. It is especially important in situations where the consequences of treatment decisions are complicated, uncertain and severe [105]. An example of such a scenario is the patient's decision to receive a durable left ventricular assist device (LVAD) - a decision that involves a highly invasive treatment strategy with complex trade-offs that affect patient survival and quality of life over an extended time.

LVAD therapy is one of the limited treatment options for patients with end-stage heart failure. Initially used to bridge patients to a heart transplant, LVADs are now also offered as a

destination therapy to patients who are ineligible for a transplant. Use of an LVAD can extend survival and increase quality of life [7, 8], but also requires significant changes in daily life as well as investment of time and money from the patient and their caregiver, and presents a heightened risk of severe adverse events such as stroke and infection [9]. Accordingly, multidisciplinary heart failure teams cooperate in educating prospective LVAD candidates and their caregivers on the associated risks, responsibilities and lifestyle changes as they are discerning their treatment options [18]. Clinical decision aids can facilitate patient and caregiver education and decision making in this process. Many decision aids have been designed to guide patients throughout the progression of heart failure, including preventative care[106], management acute chest pain[107], and durable implanted devices such as cardiac defibrillator[108] and LVAD therapy[109]. Information may be shared in multiple formats, including printed brochures, online text, graphics, videos [110], or combinations thereof [34, 109]. Risks are commonly conveyed to patients as percentages or probabilities [111, 112]. The format for presenting this information must be carefully considered, as it is important to provide a complete and interpretable picture of both the risks and benefits while not overwhelming a patient [111, 113]. Risk information is commonly presented as average probabilities for aggregate populations. However, recent advances in machine learning and data mining, in combination with the growth of clinical data registries, have made it possible to develop patient-specific prognostic models. Our group has used these methods to develop personalized models for predicting LVAD outcomes, including mortality [36], and adverse events [38] [39].

The goal of this study was to understand these patients' attitudes toward their engagement with health information and medical decision making. While patient response to

various data and risk presentation methods has been extensively studied [111, 114-116] and patient-physician relationships have been explored [117, 118], this study explores the interaction of both domains in the end-stage heart failure patient population. The unique features of this patient population is that they are typically suffering a long, slow disease progression that leaves them physically and emotionally exhausted by the time of LVAD decision making[17]. They may also be cognitively impaired from their disease[119]. The results of this study are intended to inform the design of the first personalized prognostic decision aid for patients considering LVAD therapy.

4.1.2 Methods

A paper-based survey was developed by researchers at Carnegie Mellon University (CMU) that included domains related to: interaction with their cardiologist, interest and comfort with medical information related to their condition, and familiarity with visualization of quantitative data. (The survey is included in the Appendix, as supplemental material.) The survey protocol was approved by the institutional review boards (IRB) at CMU, Duke University, and Allegheny General Hospital. All study participants provided written informed consent and were not compensated for participating in the study.

Patient Cohort and Data Collection

Patients were enrolled at Duke University Medical Center (n=22) and Allegheny General Hospital (n=35) between May 2015 and April 2016. Patients referred to the advanced heart failure program in the outpatient setting for LVAD evaluation (either bridge-to-transplant or destination therapy) who were age ≥ 18 years and New York Heart Association (NYHA) class II

– IV were included. Pediatric patients < 18 years and those unable to provide consent due to mental or physical inability were excluded. Surveys were either conducted by a study coordinator (n=35) or by the patient themselves in a private setting on the hospital premises (n=22), prior to the patient's clinic visit. All patients who began the survey study made it to completion (n = 57).

Descriptive Analysis

Descriptive statistics were compiled and analyzed in Microsoft Excel and SPSS.

Free-text response analysis

To discover common themes among responses to the final, free-text question*: "If you could imagine a computer wizard that could answer all your questions, what would you ask?*" we used the Latent Dirichlet Allocation (LDA) algorithm [120]. LDA is an unsupervised probabilistic graphical model for topic discovery. This model can be used to cluster a set of documents into groups that discuss a common topic. We used the Python implementation of LDA provided by gensim (RaRe Technologies) to analyze the responses. We performed LDA for 5000 iterations, determining the number of iterations by computing the difference in topic distribution after each iteration and stopping the algorithm when the difference became negligible. The number of topics was chosen by generating groups with LDA for 2, 3 and 4 topics, then assessing the coherency of the results for each group. Starting with the 4-topic LDA model, we found that distributions for most responses (45 out of 57) were captured entirely by topics 1 through 3, and the remaining responses contained a negligible proportion (of the order of 10^{-16}) of topic 4. This indicated that the distribution of the responses could be captured

effectively with less than four topics. When testing the 2-topic model, manual inspection showed that related documents were not clustered. For example, question responses *"How long will I live and what can I do?"* and *"Survival rate of heart transplant post heart transplant and lifestyles"*, which both are related to survival, were sorted into two different topics. The first response contained a high proportion of topic 2 (0.89) while the second one contained a high proportion of topic 1 (0.93). When using a 3-topic model, these responses were classified into the same topic. Thus, we concluded that the LDA model with 3 topics was the best performing model and was used for this analysis. In addition, we manually coded a fourth group of responses that were not captured by LDA, which were either blank or entered as "No".

Bayesian Analysis

Survey data was analyzed by a Bayesian Search method using GeNie 2.1 Academic (Bayes Fusion, Pittsburgh, PA). Missing data elements were classified in their own category, *missing*, and were not imputed. Model background knowledge was organized such that follow-up or dependent survey questions were secondary to initial or stand-alone questions. The number of inter-dependencies between variables (nodes) was limited to 20, and the number of parent nodes was limited to 8. The network structure was learned over 20 iterations, with a sensitivity of 10% and a prior link probability of 0.1%. Results were visualized as a directional nodal network, with the arcs between nodes representing the influence between responses, the arc pointing to the dependent variable, and their thickness indicating the strength of the association [121]. Variables without any relationships were shown as nodes without any arcs.

Cluster Analysis

Cluster analysis was performed on the dataset by assigning a numeric value to each answer response for the related questions identified by the Bayesian Search method. Distances between responses were computed using the mean absolute difference, which were then negated to produce a symmetric matrix of patient response similarities. Kernel Principal Components Analysis (kPCA) was then applied, using the similarity matrix as the kernel, to find a two-dimensional distribution of responses [122]. This was visualized by producing a series of scatter plots, in which the responses to questions were encoded by a color scale.

4.1.3 Results

We surveyed 57 patients using a 44-item questionnaire. Patients were 82% male, predominantly NYHA class III, with an average age of 60 (range 29-79). The respondent cohort is summarized in Table 14.

Table 14. Patient demographics

Total Patients, n = 57		n	Percent
Gender	Male	47	82%
NYHA Class	I	0	0%
	II	8	14%
	III	28	49%
	IV	20	35%
Age	Range	25-78	
	Mean	60	
Administered by	Self	22	39%
	Research Coordinator	35	61%

Medical Knowledge and Interaction

Responses related to medical knowledge and interaction are summarized in Table 15.

Most respondents (61%) reported being at least somewhat familiar with LVADs, thought their condition was severe enough to need a heart transplant (77%), and would accept one (93%). Most patients (89%) knew their cardiologist, felt comfortable discussing their physical and emotional state with their medical team, preferred communicating with their doctor in person, but would use an email or message system if it was available. When interacting with their doctor, most (74%) reported spending over 60 minutes talking about their condition and were satisfied with the duration of their interaction. In terms of learning more about their health, most (74%) patients expressed the desire to view their medical records but had not requested to see them. The majority of patients (65%) reported feeling that they had control over their treatment options.

Table 15. Answers to Medical Knowledge and Interaction questions

How familiar are you with VADs?	**n**	**Percent**
Never heard of them	4	7%
Somewhat familiar	33	61%
Very familiar	17	31%
Do you think your condition is so severe that you need a heart transplant?		
Yes	44	77%
No	13	23%
Would you accept a heart transplant?		
Yes	53	93%
No	4	7%
Do you know your HF cardiologist?		
Yes	50	89%
Yes, but forgot their name	4	7%
No	2	4%
Do you feel comfortable discussing your physical and emotional state with your physicians?		
I am comfortable discussing both my physical and emotional state	50	89%
I am comfortable discussing my physical condition, but not my feelings or emotions	4	7%
I am generally uncomfortable asking questions about my physical and emotional state	2	4%
Which method makes you feel most comfortable asking questions of your doctor?		
In person	50	88%
Over the phone	5	9%
Text message	2	4%
If you could communicate with your medical team using either an email or messaging system, would you consider using it?		
Yes	33	58%
Maybe	12	21%
No	12	21%
About how much total time have you spent speaking with your doctor about your condition?		
Less than 15 min	2	4%
15-30 min	7	12%
30-60 min	6	11%
Over 60 min	42	74%
Do you feel you spent adequate time, or wish you could spend more with your doctor?		
I am satisfied	40	71%
I was satisfied at first, but later remembered questions I wish I asked	3	5%
I was not able to ask all the questions of my doctor, but the staff (nurses, coordinators, etc.) were able to fill in my missing questions	6	11%
I was not able to ask all the questions of my doctor, but the staff (nurses, coordinators, etc.) were able to fill in my missing questions	7	13%

Table 15 (continued).

If you had access to your electronic health records, would you look at them and try to understand it?		
Yes, I am eager to look at my records	42	74%
No, I am not really interested in my records	4	7%
No, I don't think I would understand my records	8	14%
No – "I don't want to know"	3	5%
Have you ever requested access to your medical records?		
Yes, very informative	8	14%
Yes, but couldn't understand them	4	7%
No, used a chart (myChart)	2	4%
No	43	75%
Which of the following best describes how you feel about your involvement in your treatment?		
I feel like I have control over what treatments I received and when	37	65%
I have no say whatsoever, the doctors just do what they want and never ask me	3	5%
I feel like I'm *too involved*… the doctors can't make decisions on their own, without asking me	1	2%
None of the above	16	28%

Technology and Visualization Preferences

Responses related to technology and data visualization are summarized in Table 16. Many patients (46%) used a smartphone and/or a computer every day (44%), and most (54%) did not use a tablet at all. Most patients were comfortable understanding bar graphs (63%), line graphs (61%), and pie charts (71%). Few patients (40%) were confident interpreting survival curves. While most patients (68%) had not used a decision support tool in any context, most thought a roadmap of their healthcare progression would be useful (53%). They also believed that a website or computer program with their prognosis would be useful (57%) and were interested in accessing videos of patients telling stories of their implant experiences (65%).

Table 16. Answers to Technology and Visualization Preference Questions

How frequently do you use you a smart phone?	**n**	**Percent**
Every day	26	46%
Occasionally	9	16%
Never	22	39%
How frequently do you use a computer?		
Every day	23	44%
Occasionally	14	27%
Never	15	29%
How frequently do you use a tablet?		
Every day	14	27%
Occasionally	10	19%
Never	28	54%
How comfortable are you understanding Bar Graphs?		
Not at all	10	18%
Somewhat	11	20%
Very	35	63%
How comfortable are you understanding Line Graphs?		
Not at all	9	16%
Somewhat	13	23%
Very	34	61%
How comfortable are you understanding Pie Charts?		
Not at all	7	13%
Somewhat	9	16%
Very	40	71%
How comfortable are you understanding Survival Charts?		
Not at all	16	29%
Somewhat	17	31%
Very	22	40%
Have you ever used a decision tool?		
Yes	17	32%
No	36	68%
If you were given a roadmap that shows the progression of your health, and the decision points in your care, would you find it useful?		
Yes	30	53%
Somewhat	13	23%
Not at all	14	25%

Table 16 (continued).

If there was a website or computer program that would show your prognosis, would that be useful?		
Yes	32	57%
Somewhat	15	27%
Not at all	9	16%
If there was a website where you could watch short videos of other patients like you telling stories of their experiences, would that interest you?		
Yes	37	65%
Somewhat	7	12%
Not at all	13	23%

Free-text response analysis with LDA

In response to the question*: "If you could imagine a computer wizard that could answer all your questions, what would you ask?",* the LDA model detected three topics from the data, exclusive of those who chose not to respond: 1) survival, longevity, 2) Non-LVAD heart failure treatment and prevention, and 3) LVAD outcomes and side effects. Of the 57 respondents, 19 gave responses categorized in Topic 4: either blank, "No", "Nothing", or "N/A" (Table 17).

Table 17. Responses to free text question, "*If you could imagine a computer wizard that could answer all your questions, what would you ask?*"

Topic	**N**	**Example responses**
1. Survival, longevity	12	"Is this (LVAD) worth the risk, will I have a longer life?" "How long can I live without a heart transplant?" "How long will I live and what can I do?"
2. Non-LVAD Heart Failure Treatment and prevention	11	"What is the average time one with my health usually stays on milrinone? What is the likelihood someone with my health will receive a heart for transplant in that time?" "Overview multiple courses of possible treatment at each stage and have all info in one place when making decisions"
3. LVAD Outcomes and side effects	15	"How big is the equipment? How do you shower with it?" "Outcome of patients similar in age with heart disease" "What are the side effects [of the LVAD]?"
4. No answer	19	

The majority of completed responses included specific details about the LVAD, risks, and side effects.

Bayesian Search Results

Bayesian analysis revealed several variables to be inter-related. (See Figure 29.) The probability of requesting medical records was positively related to comfort talking with their physician. The time spent with physician was positively correlated with ability to understand visualization of data. Interestingly, *satisfaction* with the time spent with the physician was not associated with the *amount of time* spent with physician. Finally, patients less comfortable discussing their emotional and/or physical health with their physicians were less likely to be interested in accessing their electronic healthcare records.

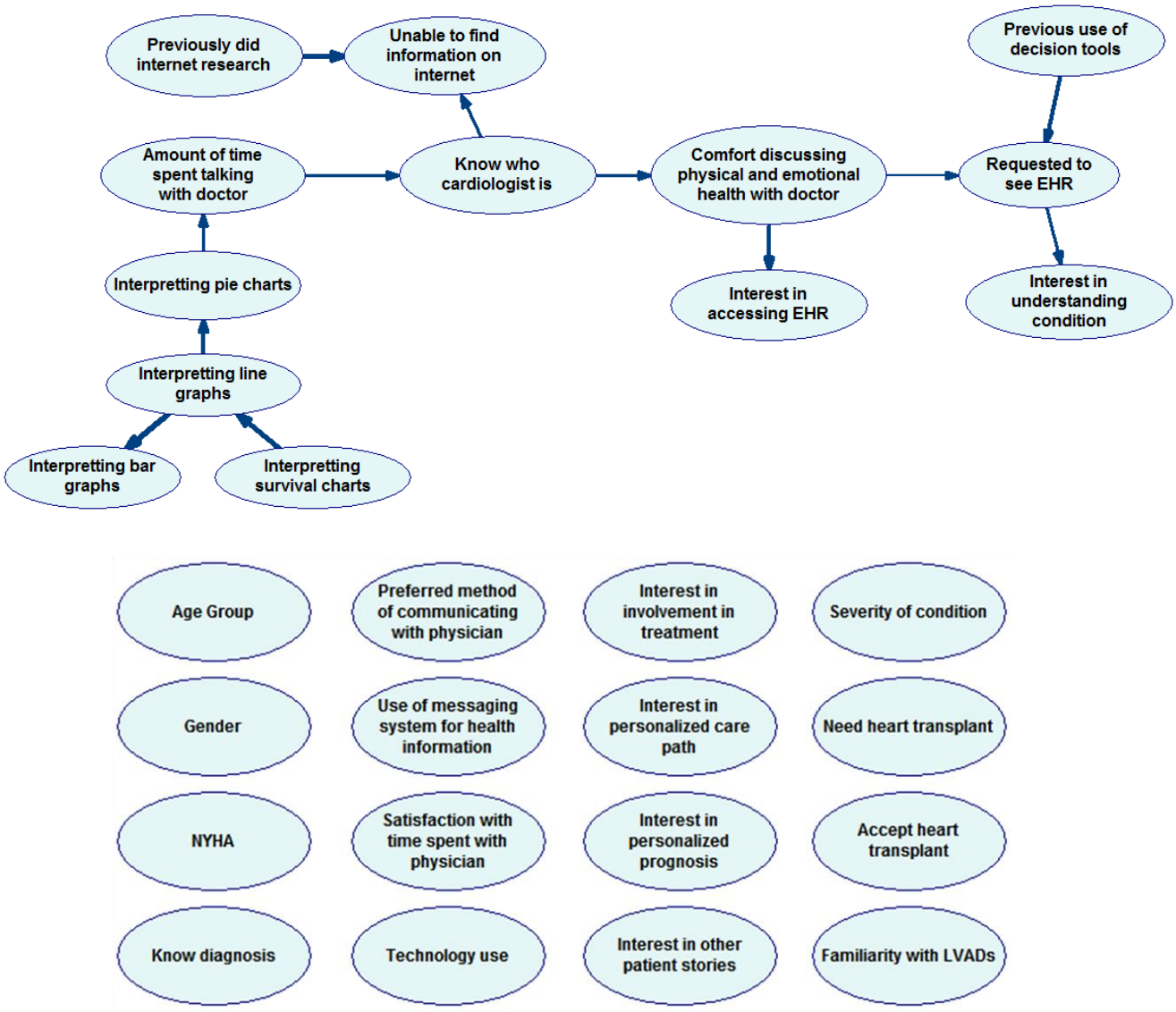


Figure 29. Bayesian model of patient responses. Each question in the survey is represented by a node. The arcs between nodes represent the influence between responses to these questions; and their thickness indicates the strength of the association. Arc direction indicates the directionality of the relationship, with the arc pointing at the dependent node. Unconnected variables (bottom) were not shown to be related to any other variable.

Cluster Analysis

Patients were compared by their responses to the interconnected nodes shown in Figure 29 as well as three independent variables (age, gender, and NYHA class). Their clustering by answer response is shown in Figure 30. Each point is a patient and the color for each point indicates the patient's response to the survey question. Interpretation was performed by visually noting clustering of similar responses.

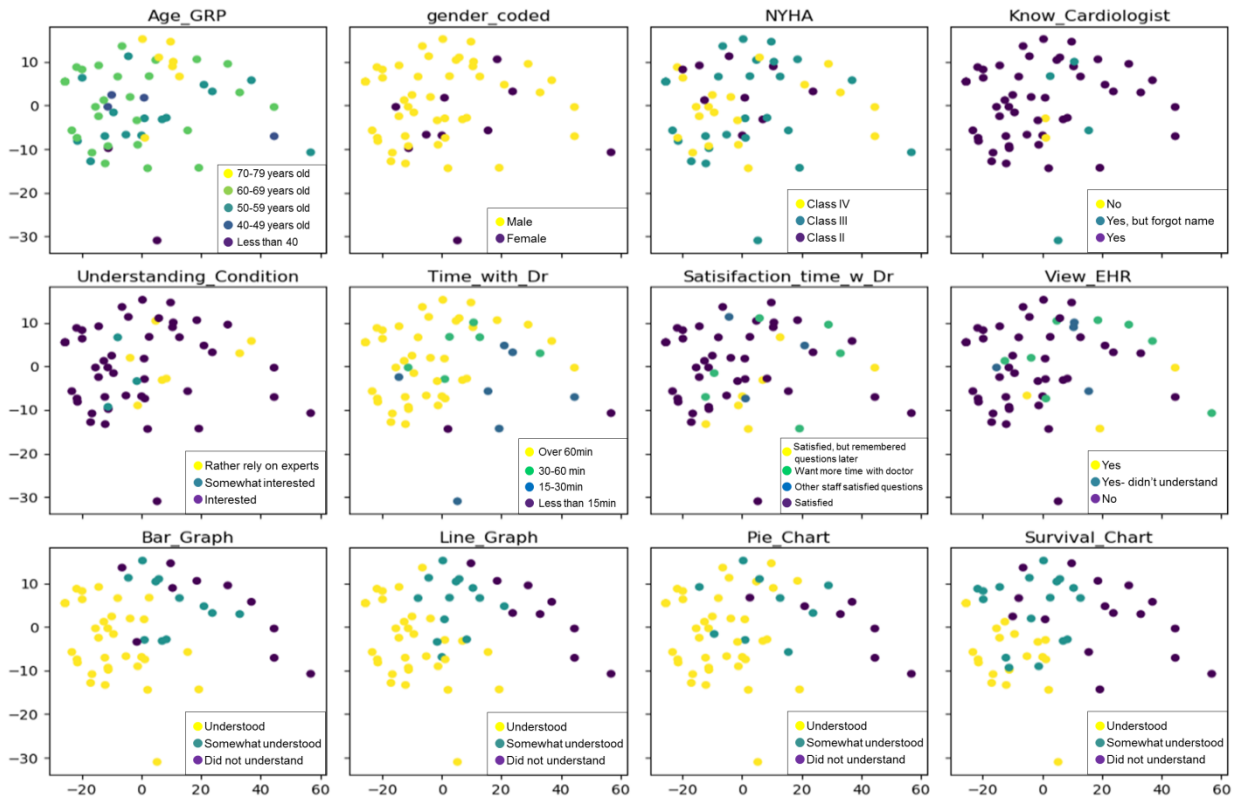


Figure 30. Cluster analysis of patient responses. Each point is one patient, plotted on a unit-less 2D space. The color of each dot corresponds to the patient's answer to the question indicated in the title of each chart.

Clustering shows a clear delineation with patient response to data visualization. (See bottom row of Figure 30.) Patients who reported they understood the visualization method (yellow dots) cluster on the left of each plot; and patients who reported not understanding the visualization cluster on the right (purple dots). This is seen in all four data visualization plots. The right-hand cluster of patients (those who did not understand the visualizations) is also seen in the chart answering: "Have you requested to see your EHR?" (View_EHR) as those who had difficulty understanding their medical records in the past. This overlap of responses indicates that the patients who had not understood their EHR were the same patients who did not understand data visualizations.

4.1.4 Discussion

This study provides insight into the attitudes of advanced heart failure patients towards engagement with their health care, with the goal of informing the development of a patient decision aid that will benefit all potential LVAD patients. The information gained from the descriptive analysis represents an optimistic outlook, with most patients comfortable discussing their health with their doctors (89%), participating in and satisfied with their treatment decisions (65%), and interested in accessing more information to help make their decisions (74%). However, there is a non-negligible minority that does not fit this outlook. To better understand this patient set, we employed machine learning techniques. In this way, we uncover possible relationships that can inform the design of decision support tools to encourage all patients to engage in their healthcare decision making.

The patients who were less interested in healthcare engagement tended to spend less than 60 minutes (cumulative) speaking with their physicians, self-identified as being less comfortable discussing their health with physicians and were less confident in understanding graphs and their own medical records. Engaging these patients in decision making requires understanding and addressing the underlying issues and concerns. Interestingly, in this study patient age, gender, and NYHA status were not associated with patient attitude towards their healthcare engagement. Socioeconomic status has also been suggested as a predictor of patient engagement [123, 124], but was not captured in this study.

The results of both the Bayesian search and clustering analysis indicate numeracy, captured in this study as comfort interpreting different types of graphs, as the key feature connected to a discomfort or disinterest in health engagement. The relationship between numeracy and health engagement in this study is consistent with a prior finding with cancer patients in which overall health engagement was correlated with highest level of patient education [125]. It is also consistent with a study of subjective patient numeracy and satisfaction with physician communication, which found and inverse relationship between the two [126]. Therefore, patient numeracy and education level are important considerations when presenting health information to patients.

The consequence of failing to assure understandability of health data is illustrated by the relationship between patient interest in seeing their medical records and whether they had asked to see their records previously. Patients who had previously asked for their records but could not understand them (7%) indicated little interest in health engagement (47% were uninterested

compared to 12% for patients who had not viewed their records and 4% for patients who had viewed and understood them). Patients may feel discouraged from not understanding their health information and therefore being less likely to try to engage again.

Text analysis showed patients' desire for information on the technical components of the LVAD and a desire for personalized outcomes, primarily survival. The first topic is a part of patient education, which can be effectively delivered through a decision support tool [34, 109]. The second topic, of personalized predictions, is being addressed by our research group through the development of personalized mortality predictions [36].

An interesting finding of this study was that patients may be less likely to answer questions about their wants when asked aloud, as opposed to writing them down anonymously. Of the 19 patients who did not respond to the open-ended survey question, only 5 (23% of 22) had self-administrated the survey versus 14 (56% of 35) patients who had the survey administered by a research coordinator. This suggests that patients were not comfortable verbalizing their desire to know more about their condition or were not comfortable telling someone in a clinical setting. A support tool that can be viewed and interacted privately by the patient may provide an opportunity for engagement when patients may otherwise feel uncomfortable. Examples of treatment success with a non-human interaction have been seen in the mental health space: with the use artificial intelligence to provide therapy to patients with PTSD [127] and use of a chatbot with college students dealing with depression [128].

When responding to questions about their current and potential use of technology, patients in this study indicated that they would find a computer program with their prognosis useful (57%) and were interested in viewing a hypothetical website with videos of patients talking about their experiences (65%). This is encouraging for use of advanced technology in patient education but cannot be used as evidence alone to eschew the more traditional hardcopy decision support information they currently receive. Future work needs to compare various education options with patients to determine the best delivery method for the most relevant information related to their healthcare decisions.

Most patients who took this survey expressed a positive outlook on their healthcare engagement. However, because the trigger for the survey was attending the informational discussion with their physician, the sampling for this survey study may be biased toward healthier and better communicating patients.

4.1.5 Conclusion

The results of this study will inform future development of a decision aid for patients considering an LVAD. The heart failure patients participating in this study who had not understood their health information in the past or who had poor numeracy indicated low interest in, or comfort with, healthcare engagement. Therefore, the decision support tool should aim to accommodate patients at their educational and comfort level, to encourage them to face daunting decisions and to create a safe space to record their worries or questions.

4.2 PROGNOSTIC DATA LITERACY AND EFFECT ON RISK PERCEPTION IN GENERAL AND PATIENT POPULATIONS

4.2.1 Introduction

Patient decision making relies on patient and caregiver understanding of potential risks weighed against the benefits of treatment. For a patient considering a left ventricular assist device (LVAD), a major trade-off is the risk of adverse events versus increased longevity and quality of life [7, 8]. Decision support tools can help navigate this decision making through presentation of educational and statistical risk information [34, 109].

With improvements in data analysis techniques, personalized risk predictions are being developed for use in decision support tools. This adds a new layer of information for patients and their caregivers to process: what is their personal risk as opposed to the average patient risk. However, it is unknown how this individualized information will affect the patient's overall perception of the risk of receiving an LVAD.

How to present risk data and probabilities to a patient population has been extensively studied [111, 112], with recommendations for the combination of both visual and numeric information, use of probabilities instead of percentages, and careful choice of probability denominators. However, these studies are focused on general risk information, and do not directly address the issue of personalized risk information communication. Additionally, the differences in perception of medical risk information between patients and non-patients have not been fully elucidated.

In this study we compare the responses from a general population and a heart failure patient population to see if there differences in response to visualization of risk information. The goal is to design a decision support tool with personalized risk information that is easy to interpret and effective for both patients and their caregivers. To achieve this, this study has three aims: first, compare different visualizations of risk information for ease of interpretation; second, examine the effect data visualization on risk perception of using general and personalized risk information; and third to understand how people interpret dependent risks, such as adverse event risk in the context of survival probability. Differences between patient and general population responses will be analyzed for all three aims.

4.2.2 Methods

All analysis performed in this study was approved by the CMU IRB board as an exempt study.

Participants

In this study, we recruited respondents from Prolific Academic, an online community of questionnaire takers geared toward research studies. The general population recruitment was limited to English speakers over the age of 40 at the time of taking the survey. The patient population recruitment was limited to English speakers who self-reported a diagnosis of heart failure from a medical doctor. Respondents were each paid $1.50 for completing the survey exercise.

The percentage of general population respondents who gave valid, complete surveys was 130 out of 136. Participants were excluded for not finishing the survey or timing out of the survey (taking over 25 minutes to complete). Of the patient respondents, 3 were excluded out of

the 80 responses due to contradictory self-reporting about their health status (e.g., “Yes, I have heart failure” and “No, I do not have heart failure”. Two patients were excluded for not completing or timing out of the survey.

Ease of interpretation of risk information

Respondents were presented with risk information about having a surgery to cure a slowly progressing but deadly disease and then asked questions about the risks of the surgery. Questions measured accuracy of data interpretation as well as perception of risk..

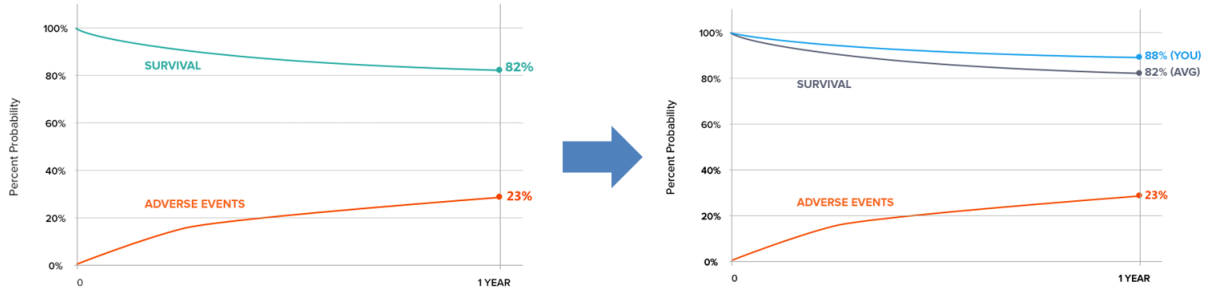
Visualizations

Respondents were randomized to viewing one of three types of visualizations (line, bar, or pictograph) with two types of information presented initially (average survival and average adverse event (AE) information or just average survival alone) and one of two types of a personalized survival probability reveal (above or below the average survival probability by 6%). With these levels, there are 12 possible types of visualizations seen by respondents.

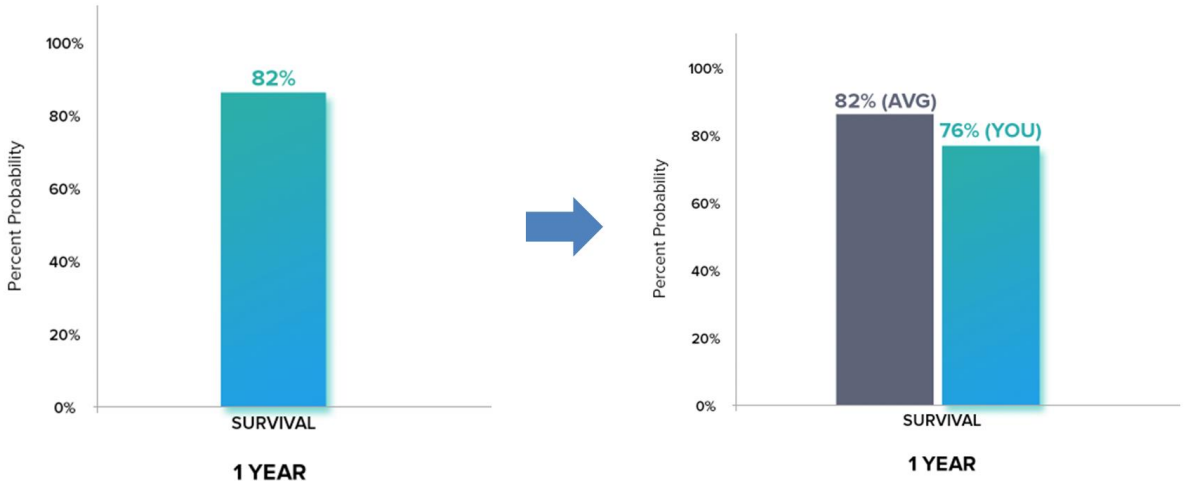
These three methods of displaying data were chosen based on their use in current and developing decision support tools. The myCORA physician decision support tool uses line graphs to show physicians the changing patient risks over time. The PCORI-funded LVAD patient decision support tool at UC Denver uses pictographs to convey risk information [109]. Bar graphs were included because of the high reporting of patient understanding in the CORA patient survey study, and the frequent reporting of bar graphs being most easily understood and preferred by users in the literature [129].

Visuals were developed by a professional designer, keeping the color scheme, font sizes, and textual information displayed consistent from visual to visual. Data used for the visual is roughly based on average risk information for patients receiving an LVAD [7]. Figure 31 has

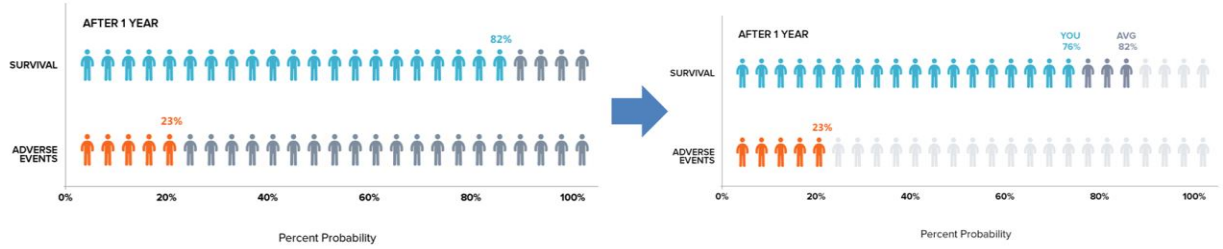
examples of the three visual types, with and without AE information displayed, before and after the personal survival reveal. All images are scaled to half the size respondents would see on their computer.



Example one: Line graph with survival and risk of AE, with reveal of personalized survival probability better than average.



Example two: Bar graph with survival, no AE information, with reveal of personalized survival probability worse than average.



Example three: Pictograph with survival and AE, with reveal of personalized survival probability worse than average.

Figure 31. Examples of visualizations before and after personalized survival reveal

Time to answer

The amount of time in seconds to complete each section was recorded and reported for analysis as time per question. Normalization to the number of questions was done to account for different sets of questions with and without AE risk and before and after the personalized information reveal (Table 18).

Table 18. Number of questions for each participant group and section

Information seen	**Number of questions**
Before reveal: Average survival and AE risk	6
Before reveal: Average survival only	3
After reveal: Personalized survival with average survival and AE risk	5
After reveal: Personalized survival with average survival	3

Time was recorded by time to submit for an entire section. Optional comments sections were not included in this time recording.

Time to answer was analyzed by one-way ANOVAs to compare the amount of time taken by visualization type and by whether AE information was presented. A two-way ANOVA was performed to measure the interaction between visual and AE information. Mean +/- 1 standard deviation was used to report the data.

Accuracy of interpretation

Interpretation questions were used to measure the ability of the participant to correctly infer data from the graph. Three different types of questions were used: 1) Direct reporting: “Out of 100 people who have the surgery, how many will have an adverse event by 1 year?”, where participants needed to state the percent AE shown on the graph; 2) Reporting the inverse: “Out of 100 people, how many will have died by one year?”, where participants needed to subtract the

percent survival from 100% to get the mortality number; and 3) Interpreting the adverse event and survival information together: “What is the probability of being alive without an adverse event by one year?” With this question type, participants needed to subtract the percentage of patients with AE from the percentage of survival.

Participants who did not see AE data were only asked question type #2. Participants who did see adverse event data were asked question types #1 & #3 before the personalized information reveal and asked question type #3 again after the personalized prognosis reveal.

The wording of the questions used two different styles: probabilistic (out of 100) and percentage-based (%). Question type #1 and 2 were probabilistic, while question type #3 was percentage-based.

A correct answer was determined if the response was within +/- 3 of the intended response, to allow for small math errors. A correct answer was assigned a ‘1’ and an incorrect answer was assigned at ‘2’. Accuracy was measured by Mann Whitney-U comparison of distributions for analysis between two groups and the Kruskal-Wallace comparison of distribution between 3 groups. Percent of population correct was used to report the summary data.

Effect of risk presentation on risk perception

Perception of risk was measured by asking the respondents to rank the size of the risk of dying before one year, the likelihood of having an adverse event within one year, and the willingness to the surgery, each on a scale of 0 to 10.

Analysis of responses between two groups was performed in SPSS using the Mann Whitney-U comparison of distributions for analysis between two groups and the Kruskal-Wallace comparison of distribution between 3 groups. Data was visualized as the percentage of

respondents indicating a risk per each category (0-10). Summary data was presented as the percent of respondents reporting a category over 5 (6 or higher).

Effect of average versus personal risk information on risk perception

The change in the perceived size of risk of dying within one year was measured by subtracting the respondent's original risk rating from their risk rating after the personalized survival probability reveal. A negative risk change indicated a decrease in the size of the perceived risk of dying, while a positive value indicated an increased perceived risk.

Similarly, the change in willingness to have surgery was measured by subtracting the original willingness rating from the rating after the personalized prognosis reveal. A negative outcome indicated an increased likelihood to have surgery after the information reveal, while a positive outcome indicated a decreased likelihood of wanting the surgery.

Perception of personalization

Respondents were asked to indicate how personalized the visualizations of information felt to them on a scale of 0 to 10, with 0 being the least personalized and 10 being the most, after the personal prognosis was revealed. Responses were compared using the Mann Whitney-U or Kruskal-Wallace comparisons of distributions, for assessing two and three groups, respectively. Data was reported as the percentage of respondents indicating a risk per each category (0-10).

4.2.3 Results

Participant Features

The patient and general population respondents represented similar ranges of educational level, races, household incomes, and location types (Table 19). The patient group was significantly younger than the general population respondents (mean 33 years of age versus 48)

and had significantly fewer females (31% versus 47%). The general population had significantly more caretakers than the patient group.

Table 19. Demographics of general and patient populations

	Gen Pop, n = 130		**Patient, n = 75**		
	n	**%**	**n**	**%**	***p-value***
Mean (Std)	48	(7.7)	33	(10.4)	*< 0.001*
Less than high school degree	2	2%	1	1%	*0.904*
High school/GED	22	17%	12	16%	*0.865*
Some college	23	18%	16	21%	*0.522*
Associate degree	12	9%	2	3%	*0.073*
Bachelor's	36	28%	22	29%	*0.803*
Master's	25	19%	19	25%	*0.303*
Professional degree	3	2%	2	3%	*0.873*
Doctoral	7	5%	1	1%	*0.150*
Asian	3	2%	3	4%	*0.490*
Black	2	2%	2	3%	*0.490*
Other	1	1%	4	5%	*0.041*
White	122	94%	65	87%	*0.080*
White and Black	2	2%	0	0%	*0.280*
Native Hawaiian	0	0%	1	1%	*0.187*
Female	61	47%	23	31%	*0.023*
Less than $30,000	41	32%	21	28%	*0.596*
$30,000 - $59,000	50	38%	22	29%	*0.187*
$60,000-$99,000	18	14%	16	21%	*0.165*
$100,000-$149,000	5	4%	8	11%	*0.054*
Over $150,000	16	12%	8	11%	*0.726*
Rural	20	15%	7	9%	*0.219*
Suburban	59	45%	24	32%	*0.060*
Urban	51	39%	44	59%	*0.007*
No	113	87%	60	80%	*0.187*
As a caretaker	12	9%	1	1%	*0.026*
As a patient	101	78%	66	88%	*0.067*
As a professional	11	8%	8	11%	*0.603*
No personal interaction	6	5%	0	0%	*0.059*

Ease of interpretation of risk information: Time to respond

Comparing the time to respond per question across all the survey sections between patients and the general population, the patient population is significantly faster, with 15 seconds for patients and 22 seconds for the general population (p-value = 0.0006) (Figure 32).

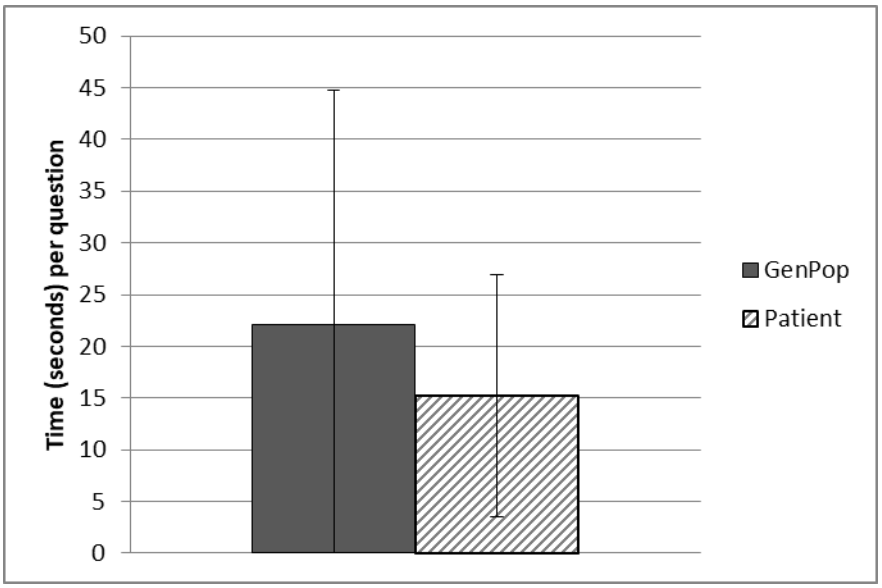


Figure 32. Time to respond by study participant type

Looking at the effect of visualization type, there was no significant difference in the general population, but there was an effect in the patient population (p=value = 0.03), with line graphs (24 seconds) and pictographs (17 seconds) being significantly different (p-value = 0.026) (Figure 33).

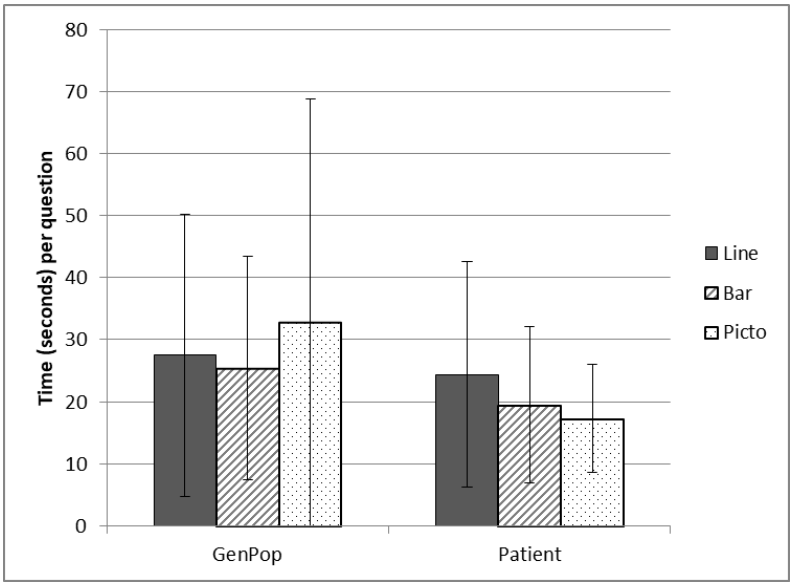


Figure 33. Time to respond per question by participant type and visualization

The response time with and without AE information was compared. The distribution of interpretation questions was different when AE information was presented (3 questions) compared to when it was not presented (1 question), therefore there may have been an effect on timing due to question difficulty. This was seen in the general population group, with a significant difference of a 17 second increase to response time when AE information was included (Figure 34). However, there was no statistical difference in the time to response in the patient group.

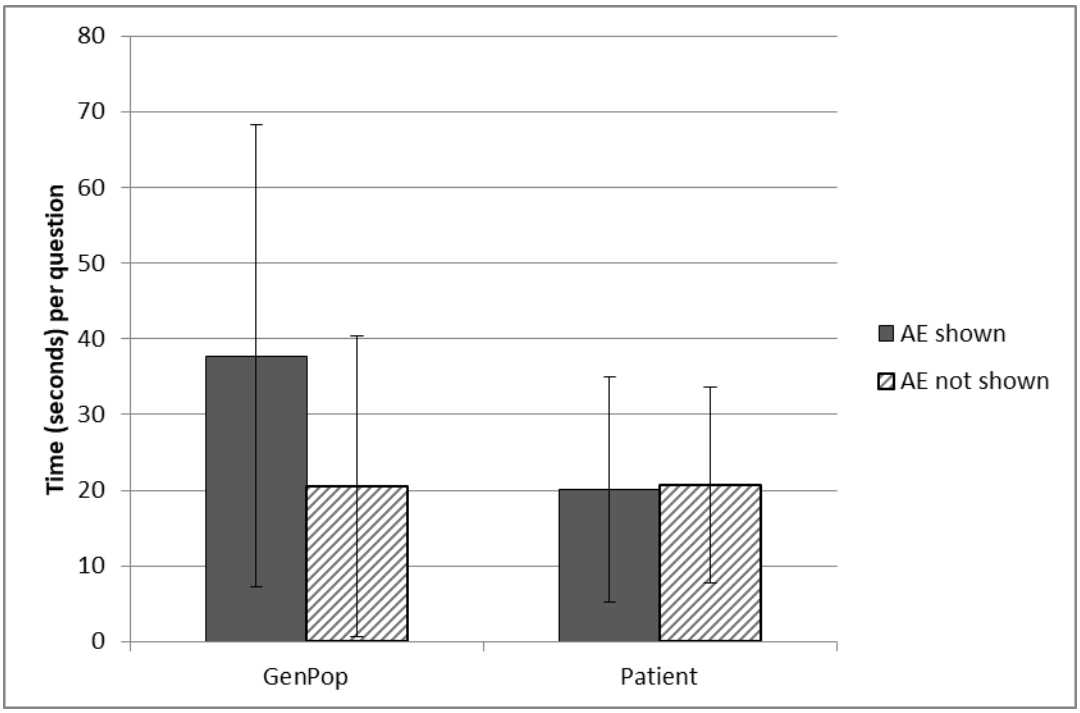


Figure 34. Time to respond by presentation of AE information and participant type

Finally, the interaction of type of visualization and inclusion of AE information was compared within each participant group, and there were no between-subjects effects in either the general population (p-value = 0.99) or the patient group (p-value = 0.8).

Ease of interpretation of risk information: Accuracy of interpretation

Four questions were asked: Direct reporting, probabilistic question "*Out of 100 people who had the surgery, how many had an adverse event by 1 year?*"; Inverse reporting, probabilistic question, *"Out of 100 people who had the surgery, how many died by 1 year?";* Dependent, subtraction question before the survival reveal, "*If someone had this surgery, what is the chance they would be alive and without an adverse event at one year?";* and dependent subtraction question after the survival reveal, "*If you had this surgery, what is the chance you would be alive and without an adverse event at one year?"*

Performance between the patient and general population cohorts was not significantly different for any question type (Figure 35). Overall, survey respondents did not answer the questions requiring interpretation of the two pieces of data together correctly. Respondents correctly answered the direct information from the graph the most often and were slightly worse at correctly answering the inverse question.

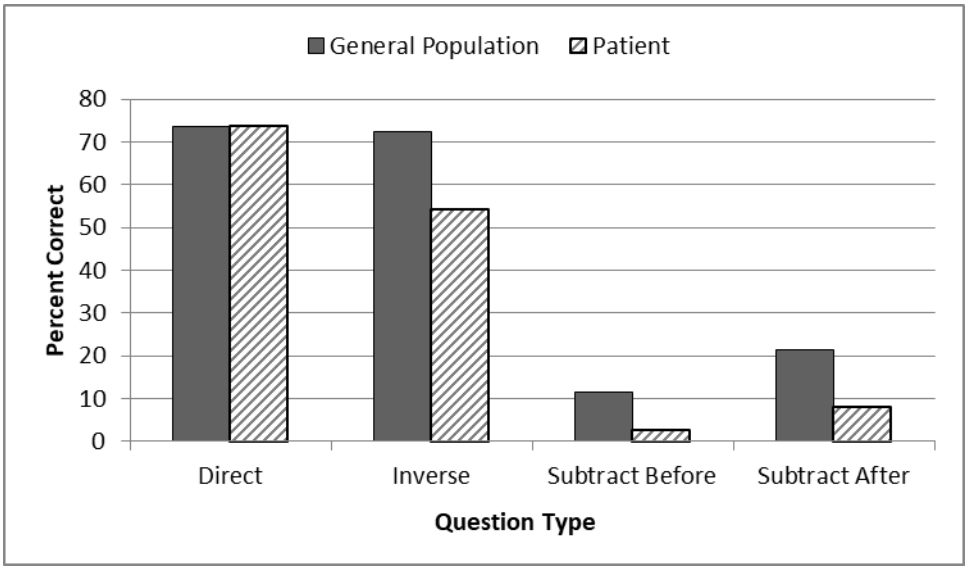


Figure 35. Percent of respondents correctly interpreting each question type

Type of visualization influenced accuracy of interpretation in both respondent populations. In the general population, the bar graph had a significantly higher rate of correct answers for the subtraction question after the personalized survival reveal (p-value = 0.043). Line graphs were correctly interpreted most often (but not to a significant degree) for the direct and inverse questions. In the patient group, pictographs had significantly higher correct interpretation for the direct question (p-value = 0.015), while line graphs had significantly higher correct interpretation for the subtraction question after the personalized survival reveal (p-value

= 0.032). The line graphs were the only visual to be interpreted correctly at all for the subtraction questions in this population. These results are summarized in Figure 36.

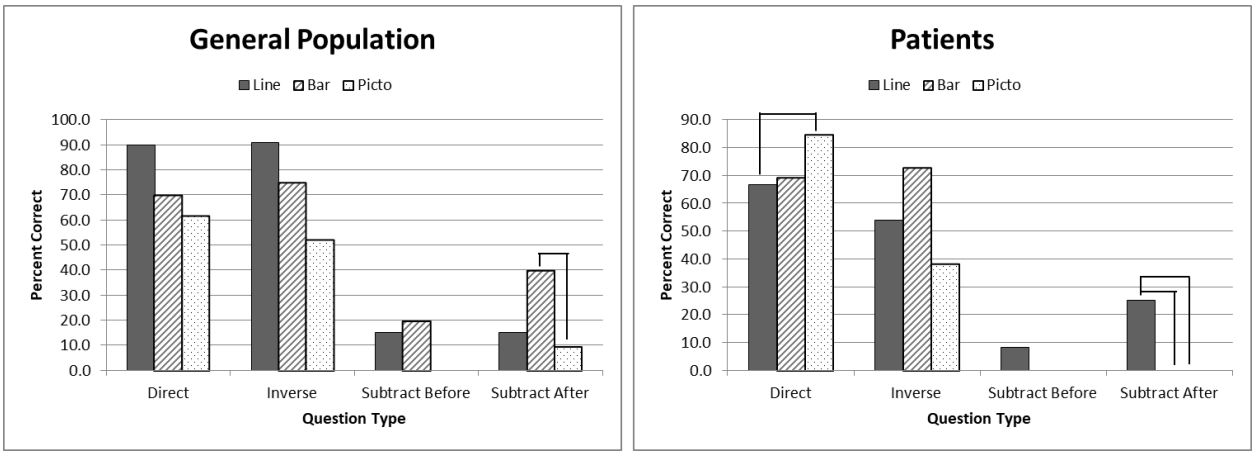


Figure 36. Percent of respondents correctly interpreting each question type by type of visualization

The reveal of personalized survival information that was better or worse than average had no statistically significant effect on the accuracy of interpretation for either respondent group.

In summary, questions asking for interpretation of survival and AE information at the same time were not easy to comprehend. For the direct and inverse reporting questions, line graphs where the easiest to comprehend for the general population and pictographs and bar graphs were easiest to interpret for the patient group.

Effect of risk presentation on risk perception: Risk of dying

After viewing the average survival information with or without AE information presented, the survey participants were asked to rate the size of the risk of dying within one year on a scale from 0 to 10, from least to most risk.

The patient and general populations had significantly different responses. The general population tended to find the risk of death smaller (35.4% over 5 on the scale of 0 to 10) than the patient population (52% over 5) (p-value = 0.039) (Figure 37).

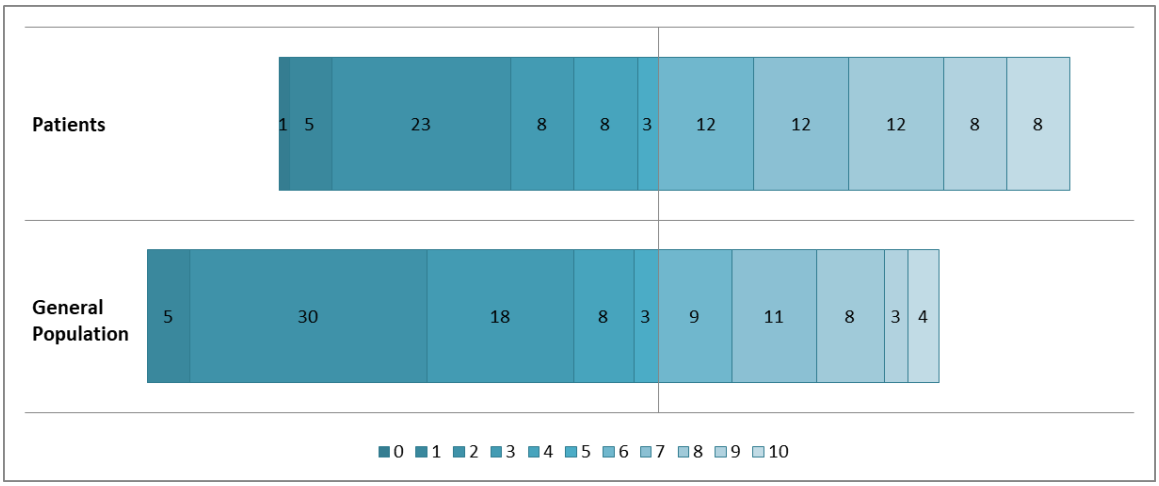


Figure 37. Comparison of size of risk of dying between general and patient populations

When considering the effect of the visualization, the type of graph had no significant impact on the perceived risk of dying within one year for patients or the general population (p-value = 0.556 and 0.546, respectively). There was a trend towards the general population perceiving the lowest risk of dying with bar charts (29.5% over 5) compared to line and pictographs (35.7% and 40.9% over 5, respectively) (Figure 38a). In the patient group, pictographs conveyed the highest risk of dying (62.5% over 5), followed by bar and line graphs (50% and 42.3% over 5, respectively) (Figure 38b).

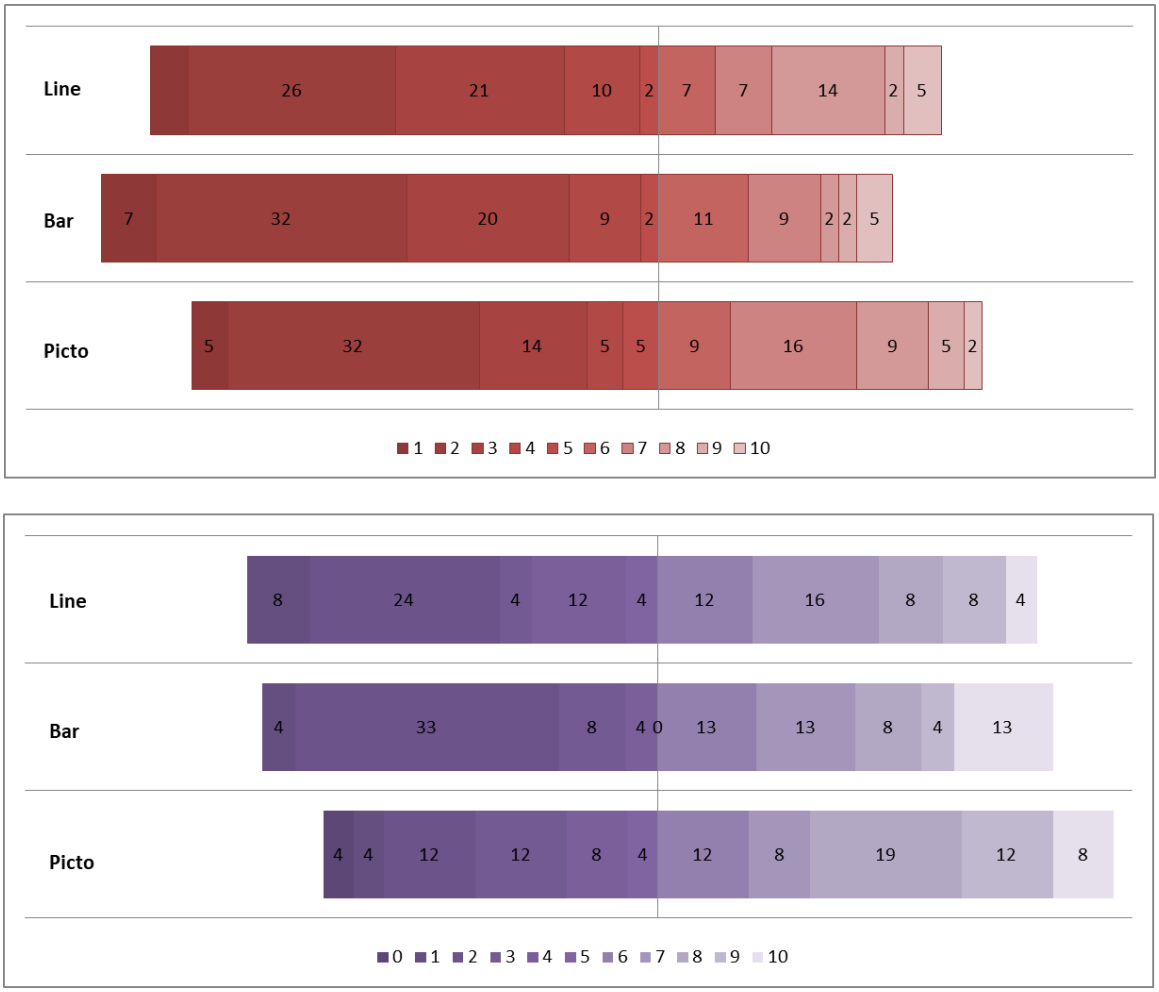


Figure 38. General population (a) and patient population (b) perceived risk of dying by type of visual

There was a statistically significant difference in perception of the risk of dying in the general population when AE information is included (p-value = 0.046), but not in the patient population (p-value = 0.278). In the general population, inclusion of the AE information decreased the perceived risk of dying, with 26.2% of the respondents giving a risk size over 5, compared to 43.5% of the respondents without AE information. This relationship was reversed for the patient population, with 57.9% of the patient population reporting a perceived risk of dying over 5 when AE information is included, compared to 45.9% without. (Figure 39)

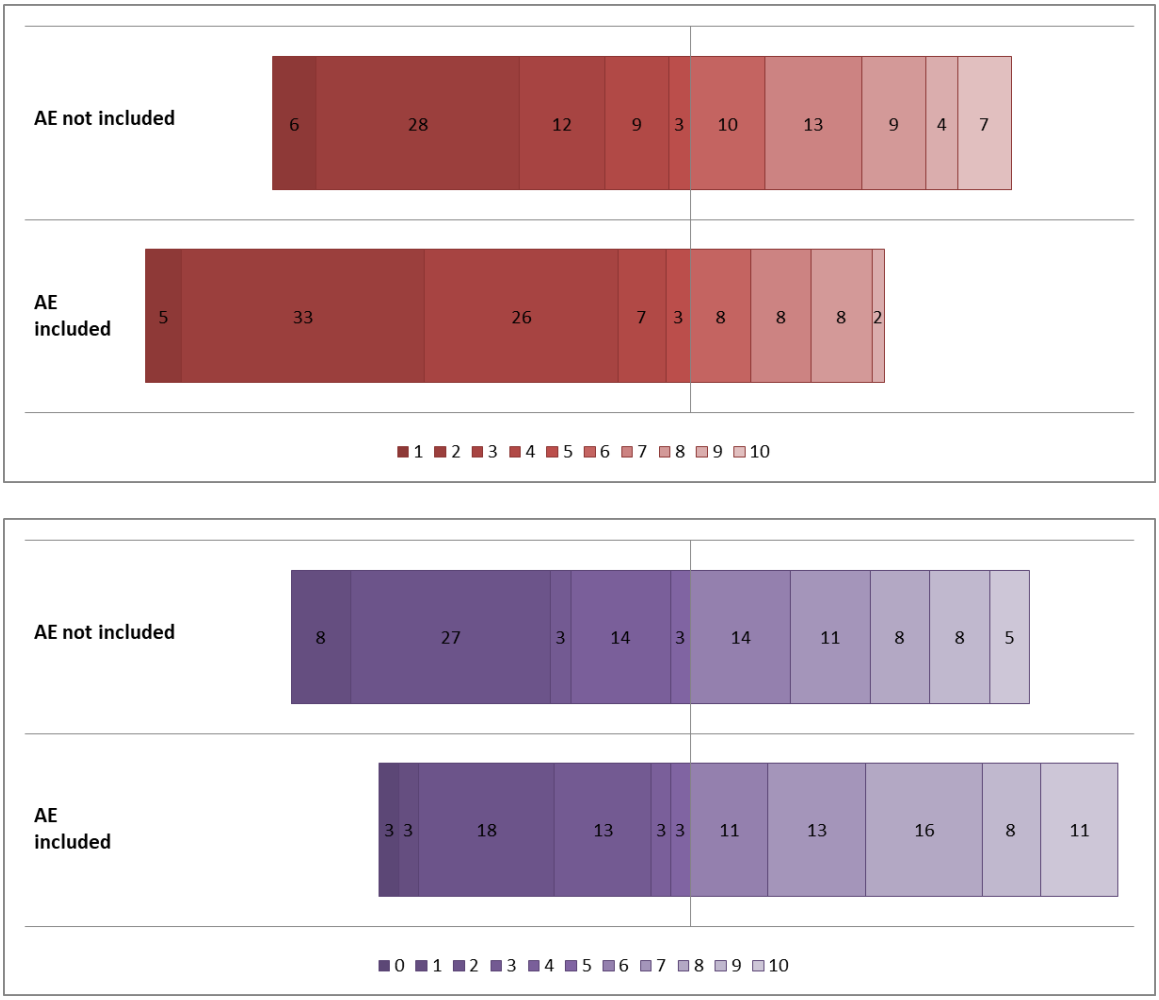


Figure 39. General population (a) and patient population (b) perceived risk of dying with and without inclusion of AE information.

After having the personalized survival information revealed, the participants were asked again what they perceived to be the size of the risk of dying within one year. Some participants saw a personal risk 6% higher than average (n = 57 for general, n = 40 for patients) and some saw a personal risk 6% lower than average (n = 73 for general, n = 35 for patients). This difference in personal risk had a significant effect on the perception of the risk of dying in both participant populations (p-value = 0.006 and < 0.001 for the general and patient populations, respectively.) In both cases the participants seeing a personal risk below average had a higher perceived risk of dying (42.5%, 51.4% over 5 for general and patient populations, respectively) (Figure 40).

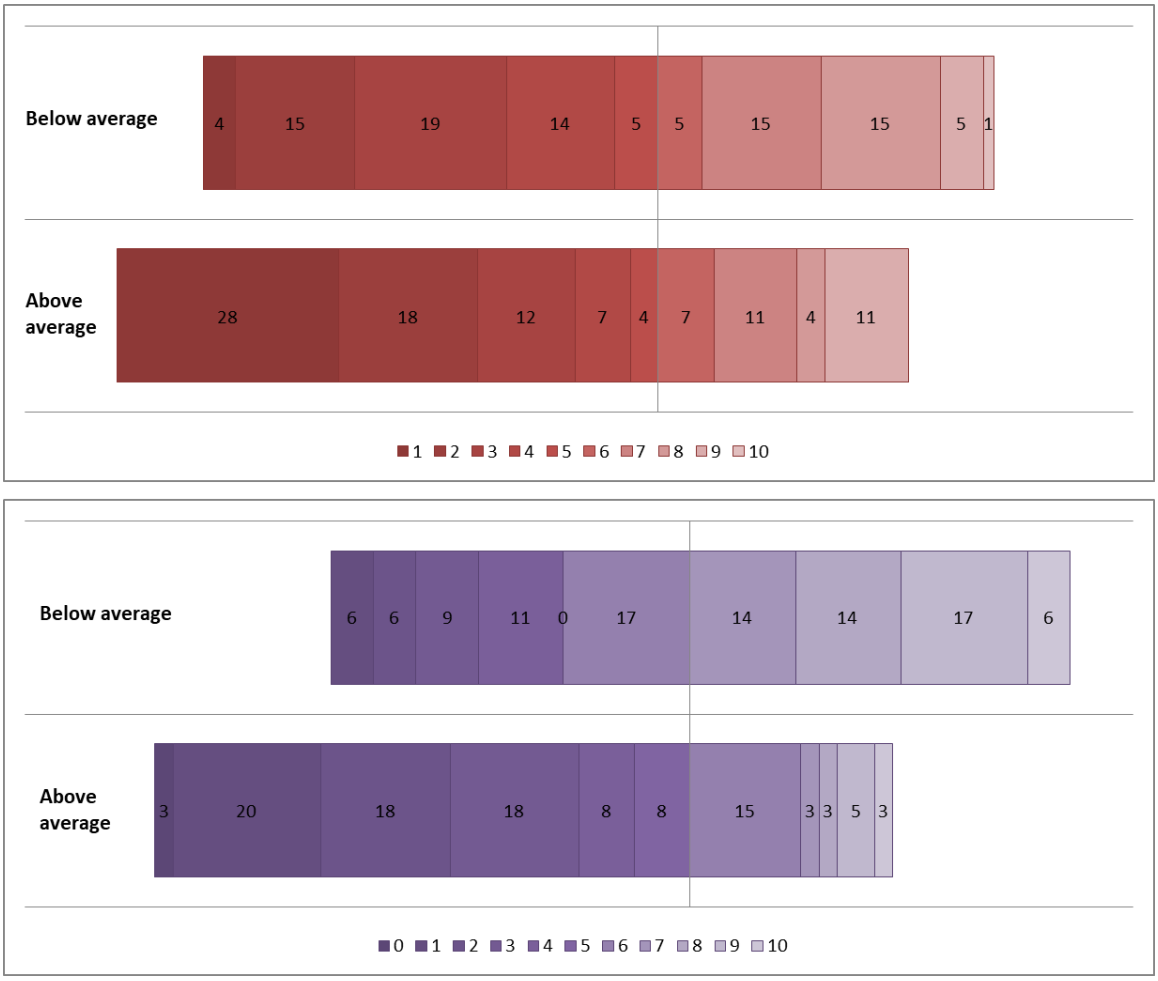


Figure 40. General population (a) and patient population (b) perceived risk of dying after reveal of personalized survival above or below the average.

Effect of risk presentation on risk perception: Likelihood of AE

Only respondents who saw AE information were asked about the likelihood of having an adverse event within one year after surgery.

General population respondents had a significantly different distribution of response from patient respondents, with patient expressing a greater likelihood of having an adverse event (50% versus 27.9% responding over 5, p-value = 0.016) (Figure 41). This matches the pattern in the perceived risk of dying, where patients also indicated a higher risk than the general population.

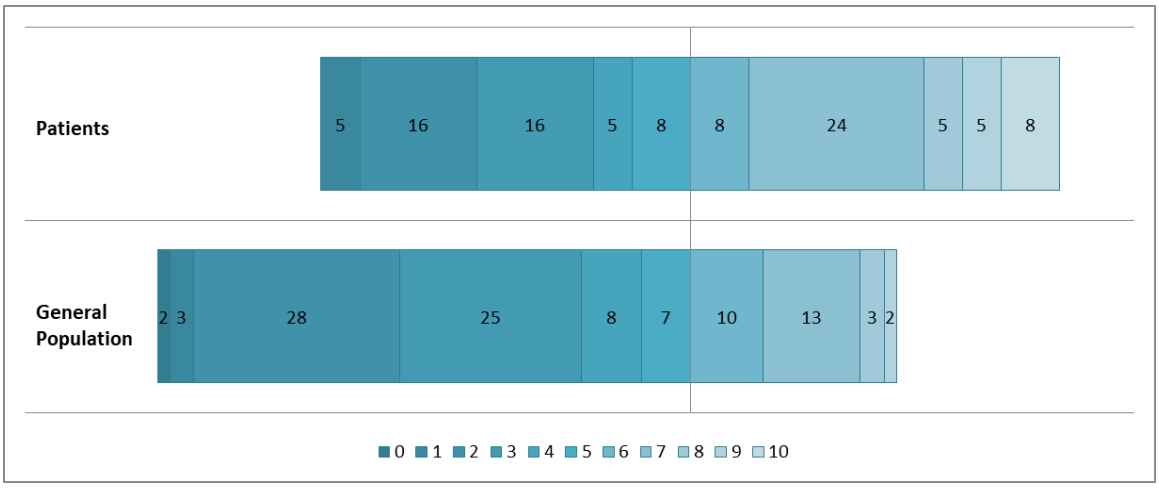


Figure 41. Comparison of size of risk AE occurrence between general and patient populations

The type of visualization did not have a significant effect on the distribution of perceived AE risk in either of the participant populations (p-value = 0.338, 0.926 for general and patient populations, respectively.) Pictographs conveyed the highest likelihood of AE in both the patient and general populations, with line graphs showing the lowest risk for the general population and bar graphs the lowest in the patient population (Figure 44). This matches the response of patients and general populations to pictographs having the highest perceived risk of dying within one year.

Figure 42. General population (a) and patient population (b) perceived risk of AE by type of visual

Willingness to consider surgery

There was no statistically significant difference in the distribution of responses from patients or the general population when asked how willing they would be to consider having the surgery (p-value = 0.840). There was a trend towards patients being less likely to consider surgery, with 78.7% over 5, compared to 82.3% for the general population. This is in line with the patient population expressing a higher perceived risk of death and adverse events in the year after surgery, compared to the general population (Figure 43).

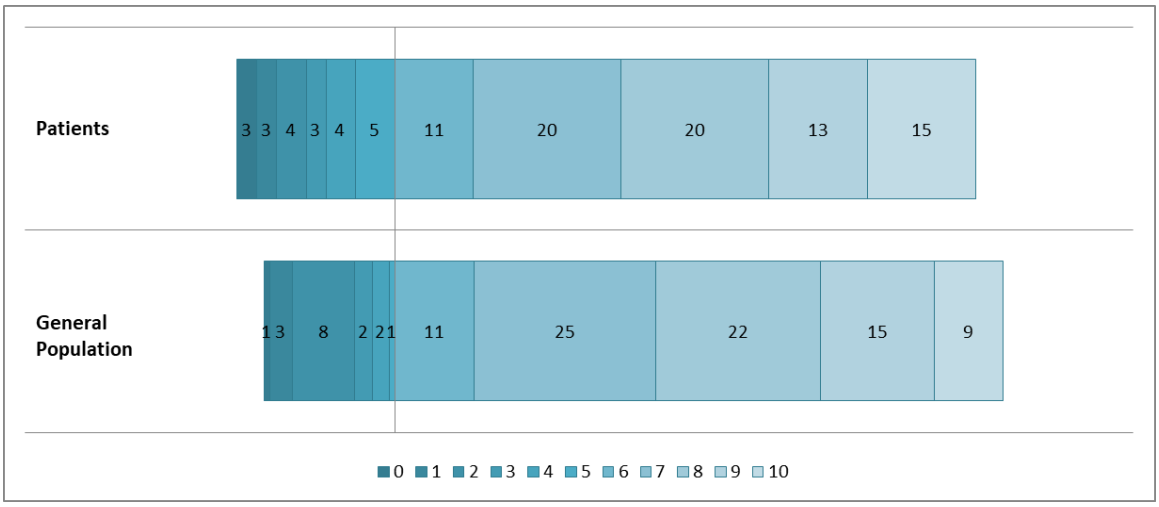


Figure 43. Comparison of size of willingness to consider surgery between general and patient populations

The type of visualization did not have a significant effect on the distribution of perceived willingness to consider surgery in either of the participant populations (p-value = 0.382, 0.251 for general and patient populations, respectively.) Bar graphs conveyed the highest likelihood of considering surgery in both the patient and general populations (83.3% and 86.4% over 5, respectively), with pictographs having the least likelihood for the general population and line graphs for the patient population (76.0% and 75.0% over 5, respectively (Figure 44).

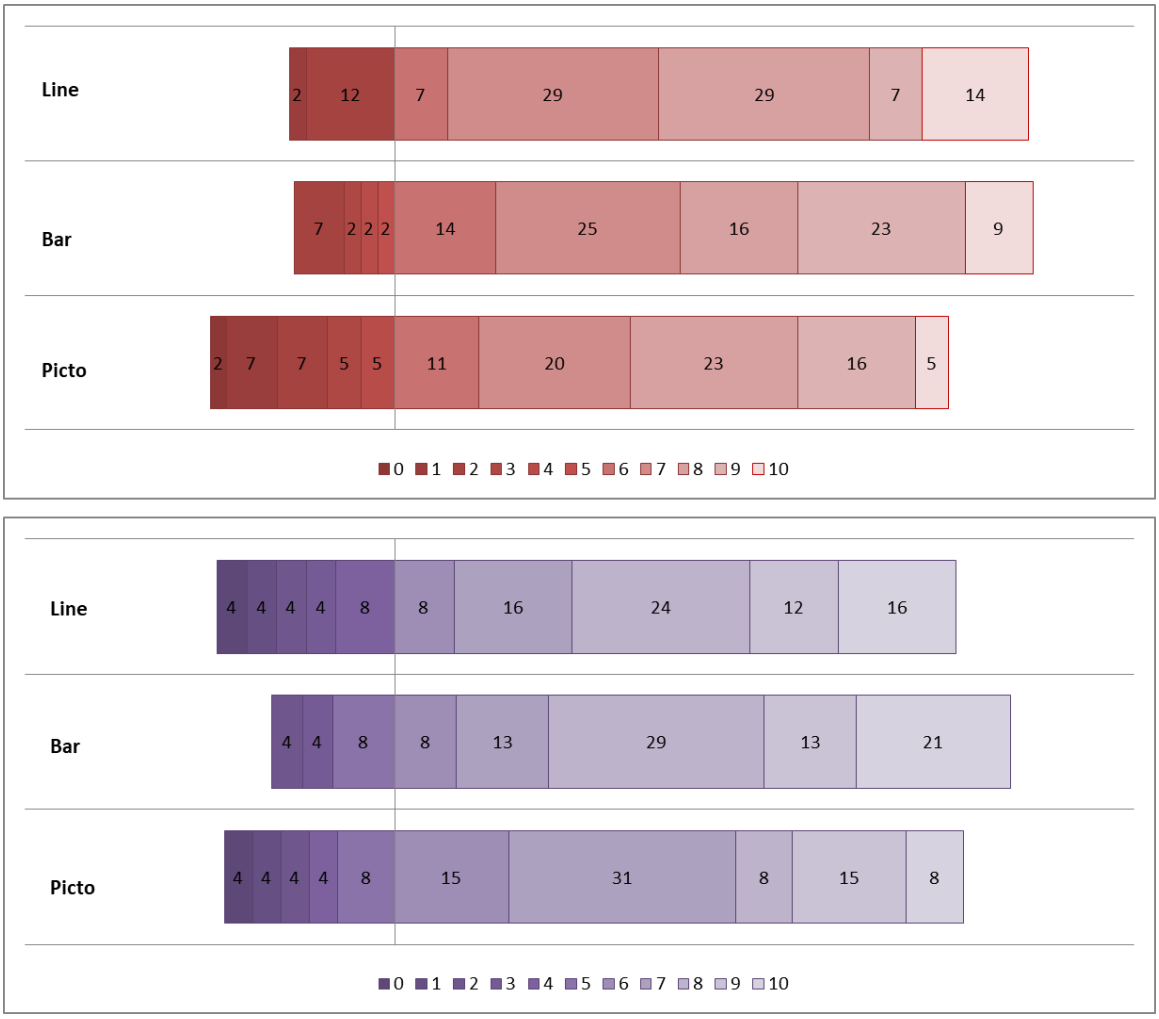


Figure 44. General population (a) and patient population (b) perceived willingness to consider surgery by type of visual

There was no statistically significant difference in the willingness to have surgery with or without the AE information included in the general or patient population (p-value = 0.363, 0.234, respectively). In the both the patient and general population, inclusion of the AE information increased the likelihood of wanting surgery, with 86.8% of the patient respondents giving a likelihood over 5 and 83.6% of the general population respondents (Figure 45.)

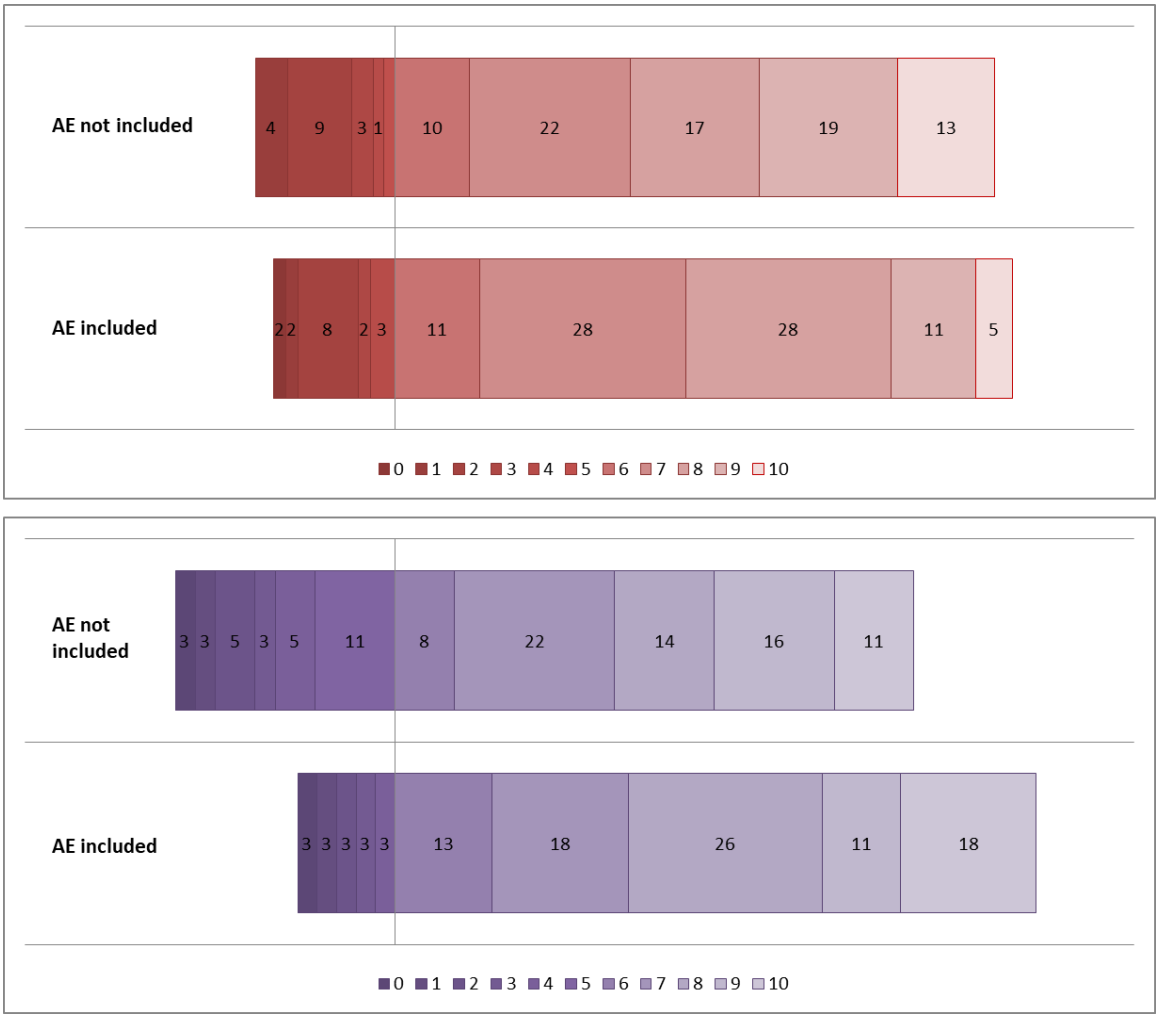


Figure 45. General population (a) and patient population (b) perceived willingness to consider surgery by inclusion of AE information

This contrasts the results from the perceived risk of dying in the patient population, where inclusion of AE information conveyed a higher risk of dying within one year. Despite the higher risk of dying perceived by patients, the willingness to have surgery is also higher.

After having the personalized survival information revealed, the participants were asked again whether they would consider having the surgery. The addition of personalized risk had a significant effect on the willingness to consider surgery in the general population (p-value = 0.020) (Figure 46a) but not the patient population (p-value = 0.719) (Figure 46b). In the general population, the participants seeing a personal survival probability below average were less likely to consider surgery than those seeing a higher survival (74.0% versus 78.9% over 5, respectively.) The patient population had the opposite trend, with a slightly higher likelihood of considering surgery when the survival probability was below the average (77.1% versus 75% over 5, respectively).

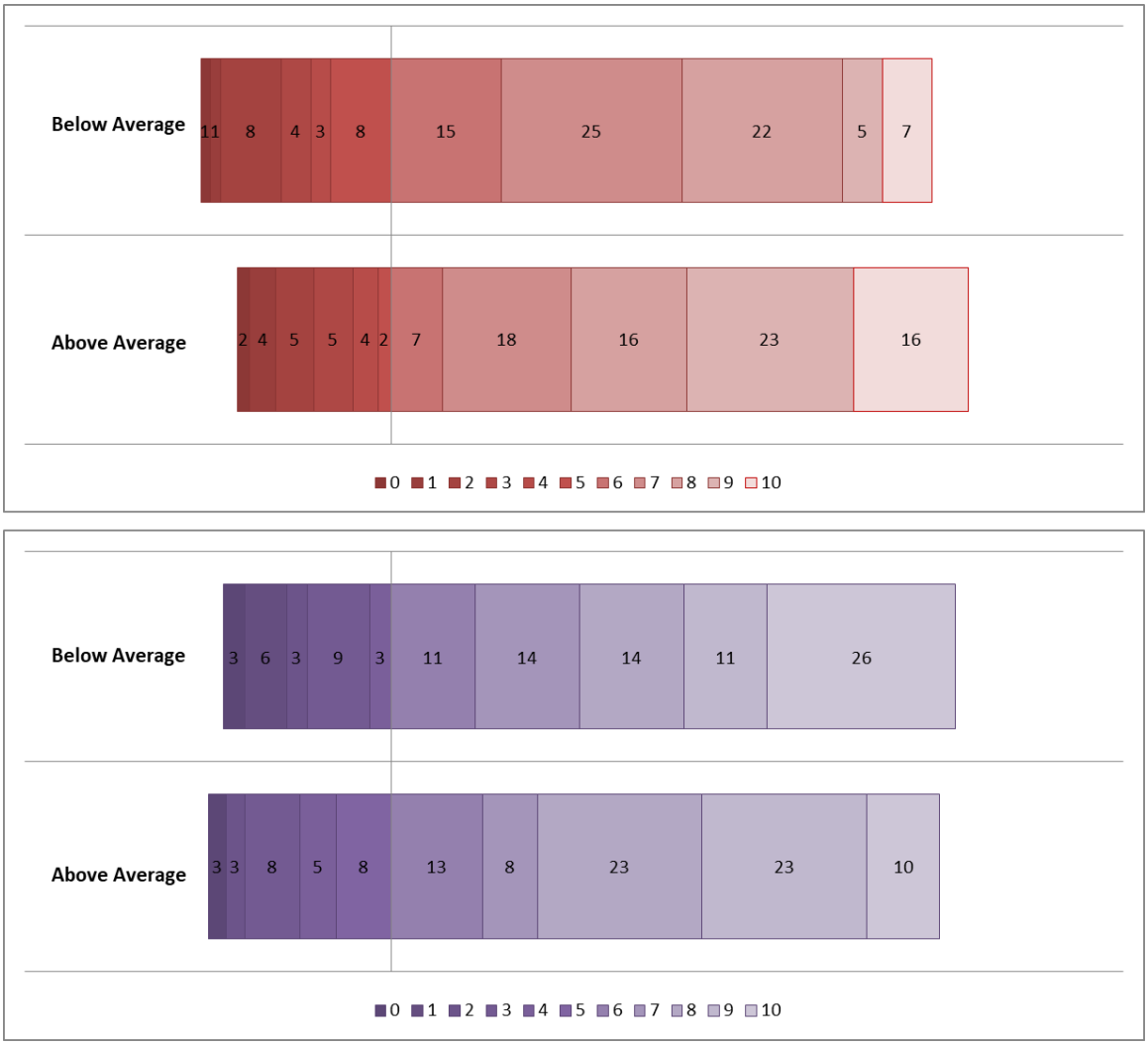


Figure 46. General population (a) and patient population (b) likelihood of wanting surgery after survival probability is shown

Effect of personal vs average risk information on risk perception: Change in perceived risk of dying within one year

There was not a statistically significant difference in the total delta in the perceived risk between patients and the general population before and after the personalized information reveal. There was a slight trend for patients to have an increase in their risk perception compared to the general population (Figure 47).

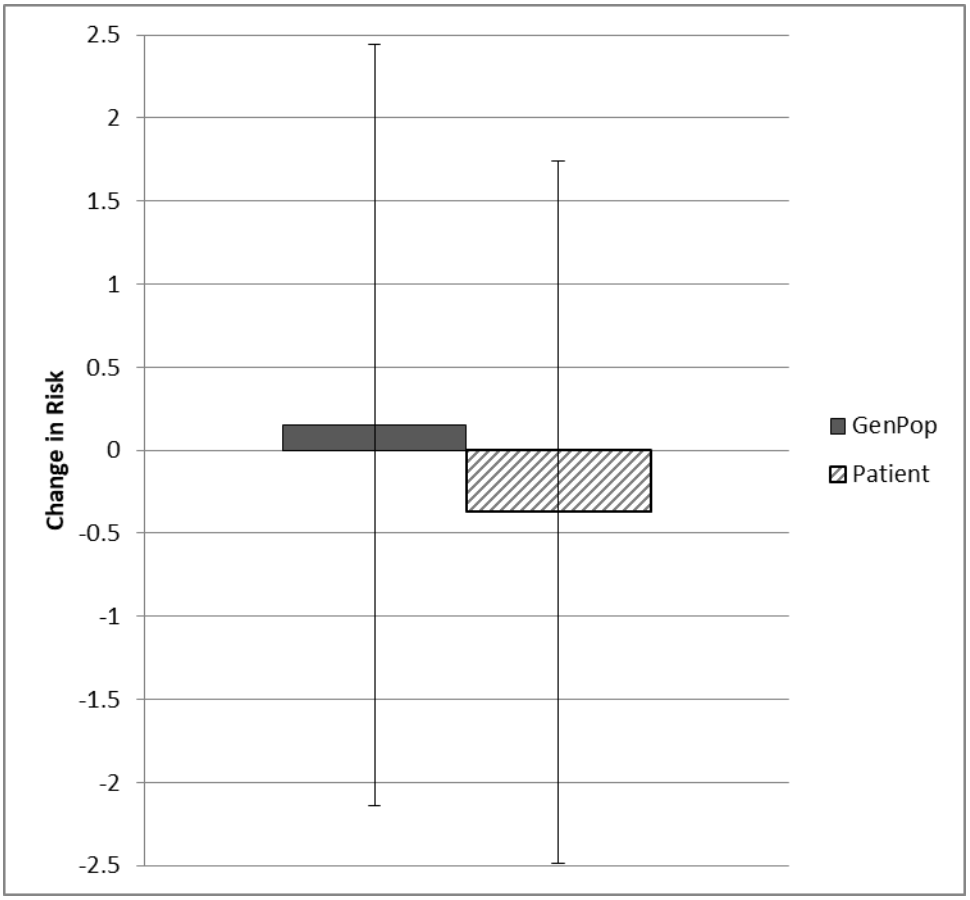


Figure 47. Comparison of change in risk of dying perception between general and patient populations

When considering whether the participants saw either a personalized survival probability that was above or below the average survival, both populations had a significant difference between groups. In the general population, the perceived risk dropped by a mean of 0.86 in the above average survival group, compared with a rise of 0.95 when the survival was worse than average (p-value <0.001). Similarly, in the patient population, the perceived risk dropped by a mean of 0.98 in the above average survival group, compared with a rise of 0.31 when the survival was worse than average (p-value = 0.008) (Figure 48).

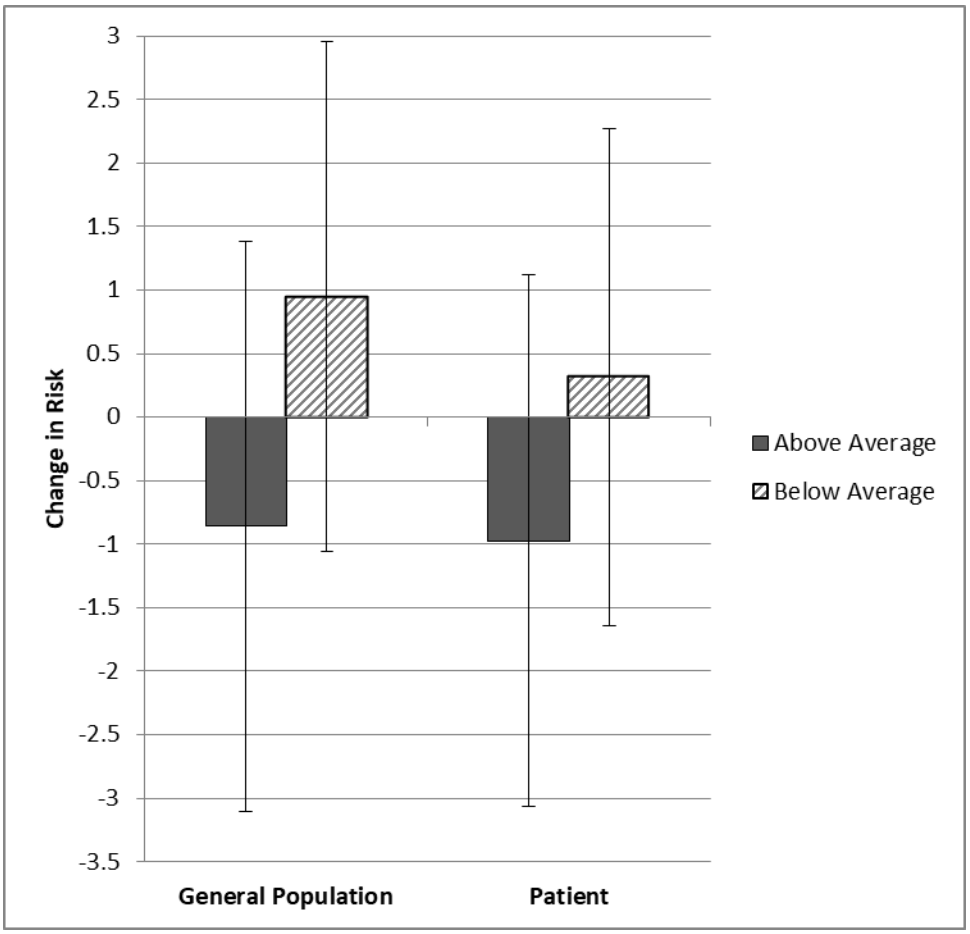


Figure 48. Comparison of change in risk of dying perception between general and patient populations, with survival probability above and below average

The effect of the type of visual on the change in perceived risk of dying within one year was only significant in the below average survival general population group (p-value = 0.037). All other groups had no significant differences.

The trend for the general population is that with an above average survival risk, all graph types have a decreased risk perception, with line graphs showing the biggest decrease in risk (1.37) and bar graphs the least (0.44). For the below average, general population group line graphs also had the biggest impact on risk perception change (1.70) and pictographs had the least (0.21).

For the patient population, the below average survival group had the largest change in risk perception with bar graphs, with an increased risk of 0.70. Interestingly, in this group the participants seeing a pictograph reported a *decreased* risk perception of 0.60. In the above average survival probability patient group, line graphs had the biggest decrease in risk perception (1.6). The pictographs had a decrease in risk of 0.56, very similar to the response to pictographs in the below average group (Figure 49).

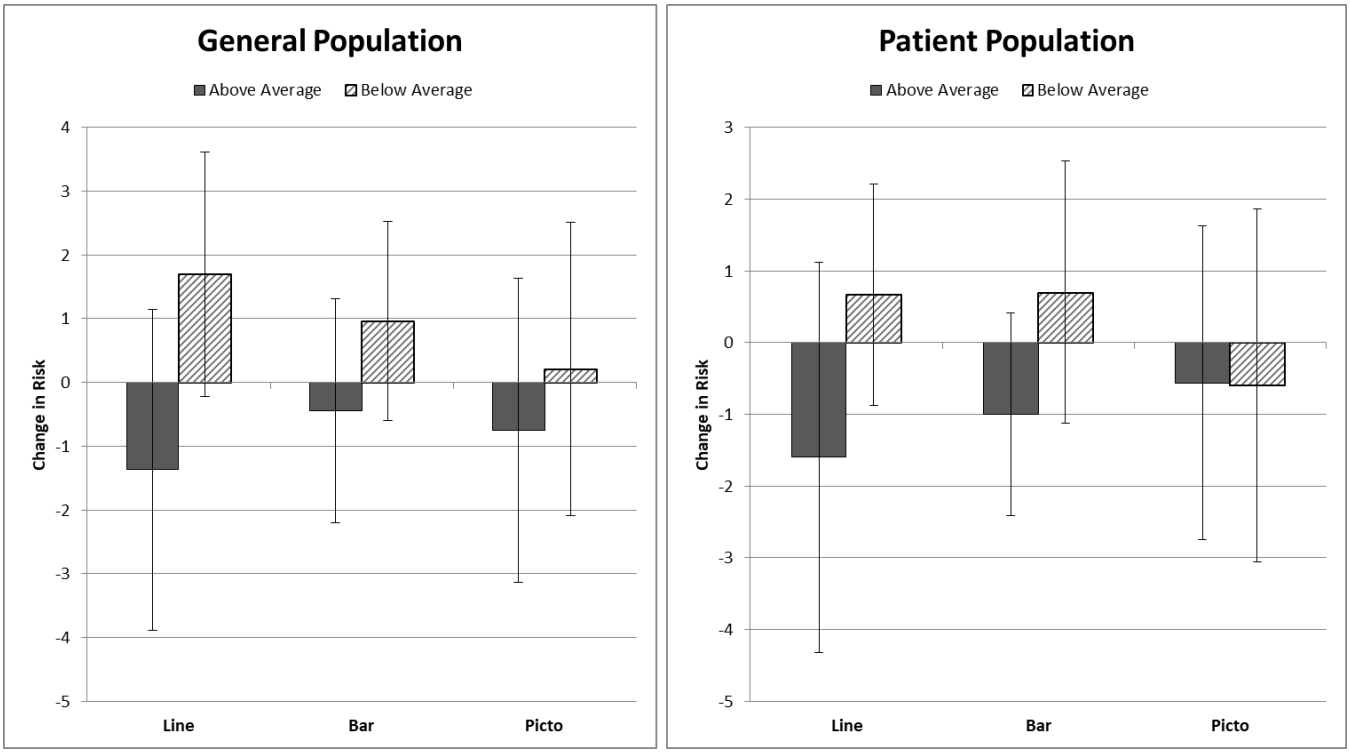


Figure 49. Comparison of change in risk of dying perception between visualizations, with survival probability above and below average in general (a) and patient (b) populations

Presentation or exclusion of AE information had no significant effect on the change in perceived risk in any respondent group. In the general and patient populations with above average survival probability there is a larger decrease in risk with the AE shown versus not shown (-0.90 with AE versus -0.81 without AE in the general population, -1.2 with AE versus -0.75 without in patient). In the general population group with below average survival, there is a larger increase in risk when AE is not shown (0.68 with AE versus 1.14 without). In the patient population with below average survival this pattern is reversed, including AE information created the larger increase in risk (0.50 with AE versus 0.12 without) (Figure 50).

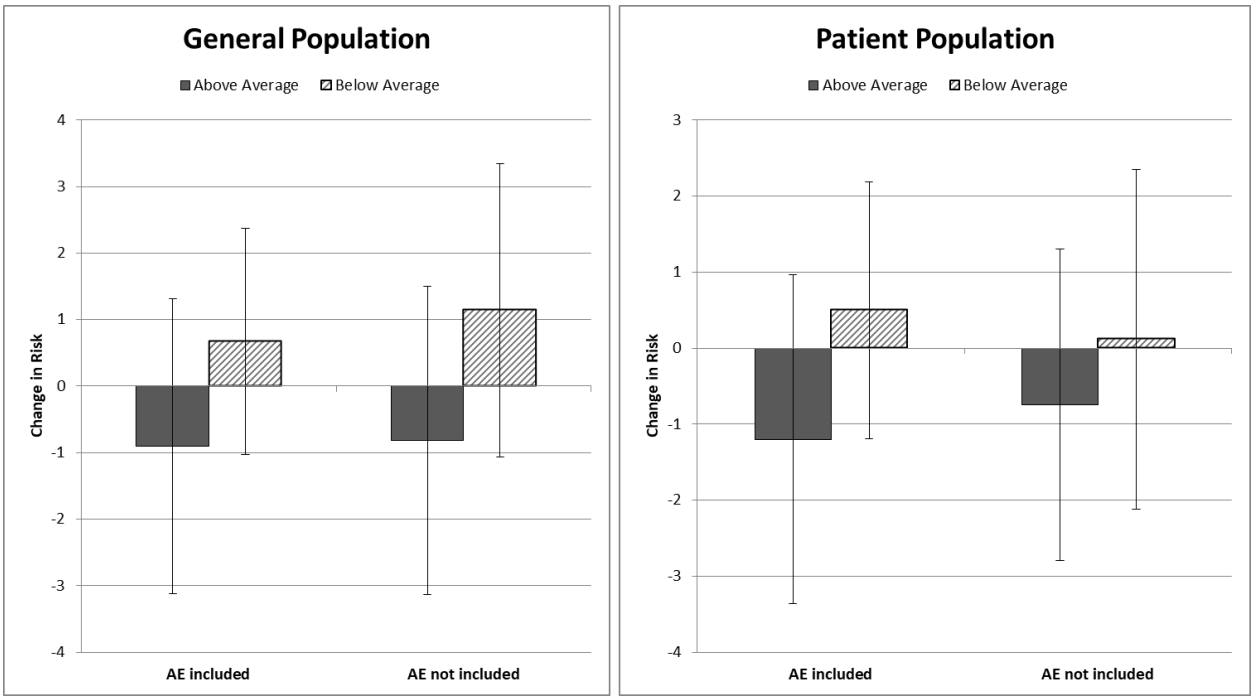


Figure 50. Comparison of change in risk of dying perception with AE information included or not included, with survival probability above and below average in general (a) and patient (b) populations

There was no effect from the interaction of graph type and inclusion of AE information on the change in perceived risk of dying within one year in any of the populations, above or below average survival.

Effect of personal vs average risk information on risk perception: Change in willingness to have surgery

There was no statistical difference in the change in willingness to have surgery between patients and the general population before and after the personalized information reveal (Figure 51).

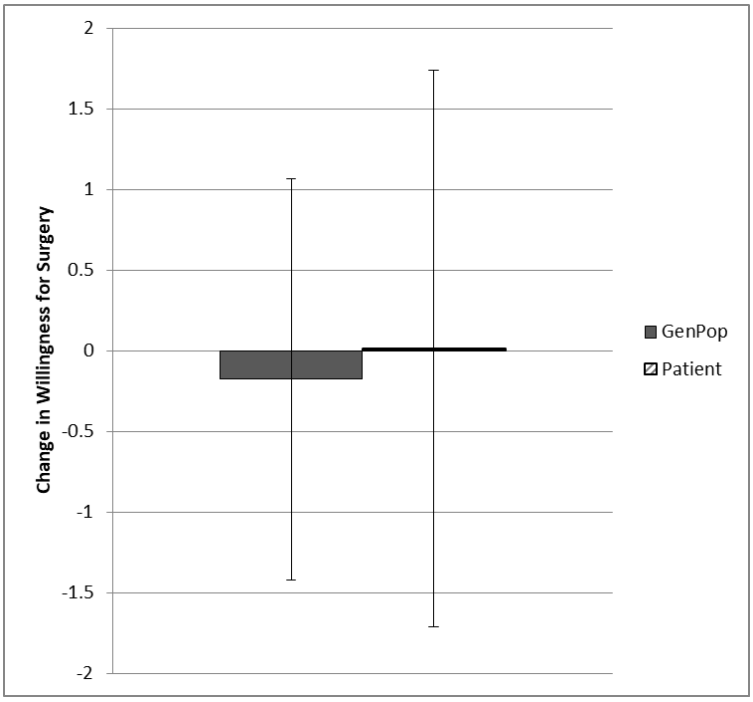


Figure 51. Comparison of change in willingness to have surgery between general and patient populations

Whether the participants saw a personalized survival probability that was above or below average had a significant effect on the willingness to have surgery in the general population. The willingness increased by a mean of 0.21 in the above average survival group, compared with a decrease of 0.48 when the survival was worse than average (p-value = 0.001).

The change in the patient population group was not significant (p- value = 0.162) but followed the same pattern as in the general population. The willingness to have surgery increased by a mean of 0.28 in the above average survival group, compared with a decrease of 0.29 when the survival was worse than average (Figure 52).

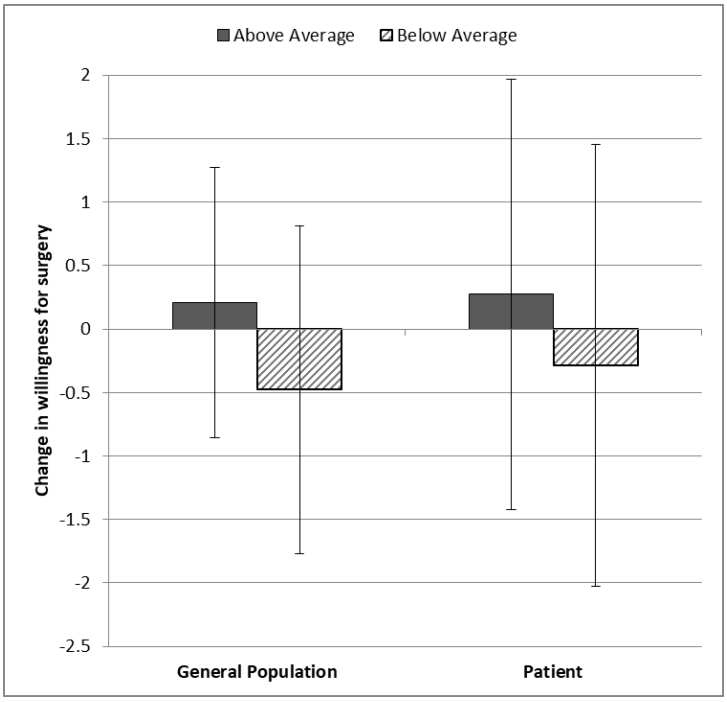


Figure 52. Comparison of change in willingness to have surgery between general and patient populations with above and below average survival probabilities

The effect of visualization was not significant in any of the respondent groups (Figure 53). Of the above average survival general population, respondents viewing bar graphs had the largest increase in willingness to have surgery (0.33), followed by line graphs (0.21) and pictographs (0.10). In the below average survival general population, bar graphs decreased the willingness the least (-0.27) followed by pictographs (-0.54) and line graphs (-0.65). This conveys that the bar graph makes the general population feel the most optimistic about considering surgery.

In the patient population, the above average survival group had a very different response than the above average general population group. These participants had a decrease in willingness to have surgery when data was viewed on a bar graph, despite having a better than average survival outcome (-0.29). The line graph had the largest increase in willingness to

consider surgery (0.60) followed by pictographs (0.56). In the below average survival patient population the pictograph surprisingly increased the willingness to consider surgery (0.60), while the line graph decreased the willingness the most (-0.80) followed by the bar graph (-0.40). Taken all together, patients viewing the data on pictographs have the most optimistic change in willingness to consider surgery, while those patients viewing bar graphs have the greatest decline in willingness.

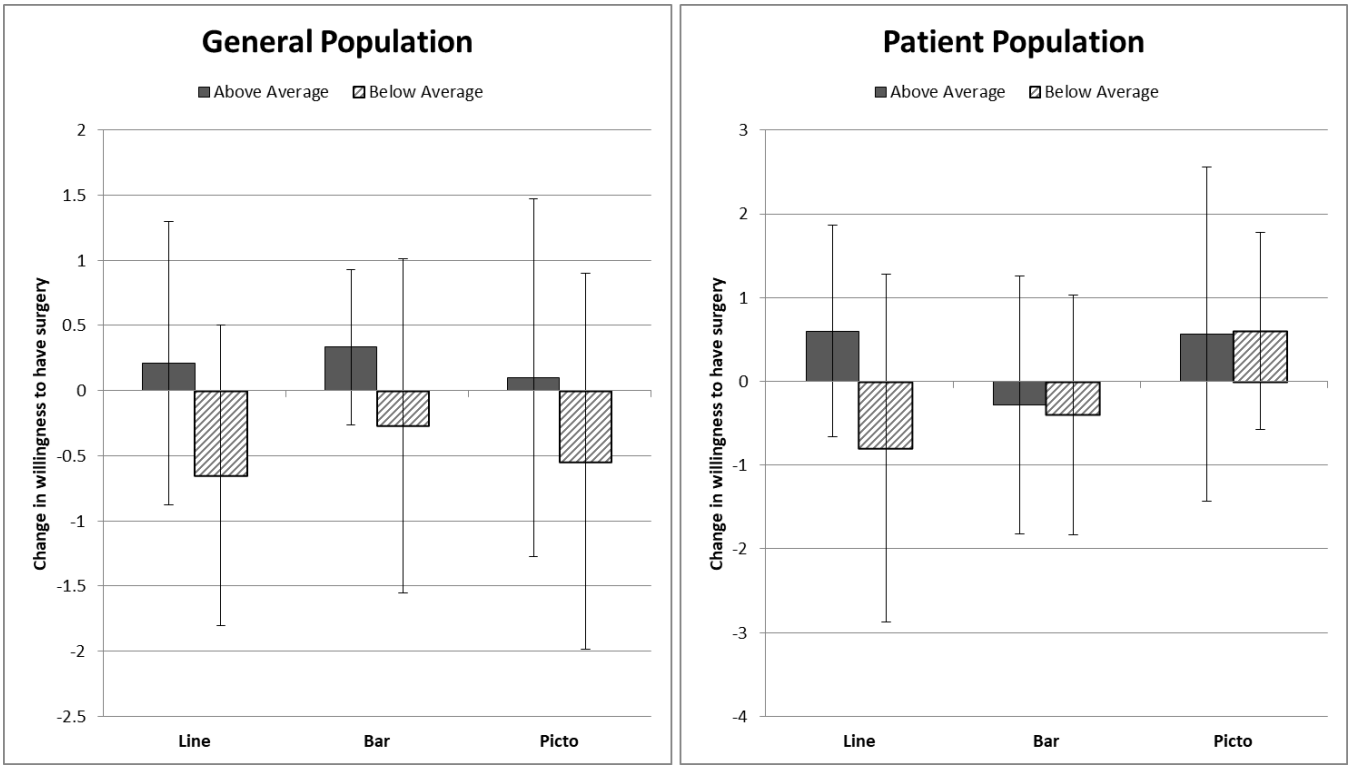


Figure 53. Comparison of change in willingness to have surgery between general and patient populations with above and below average survival probabilities by type of visualization

Inclusion or exclusion of AE information had no significant effect on willingness to consider surgery in any respondent group (Figure 54). In the general populations with above average survival probability there is a larger increase in willingness with the AE shown (0.37)

versus not shown (0.04). General population respondents seeing below average survival, had a greater decreased in willingness to consider surgery when AE information was not included (-0.71) versus when it was included (-0.16). Overall, the inclusion of AE information seems to make the general population respondents more willing to consider surgery.

In the patient population there is even less difference in effect with and without AE information. Patients with above average survival had a slight increase in willingness with AE included (0.20) and with AE excluded (0.35). When patients saw a below average survival probability, they had a slight decrease in willingness to have surgery with AE (-0.22) and without AE (-0.35).

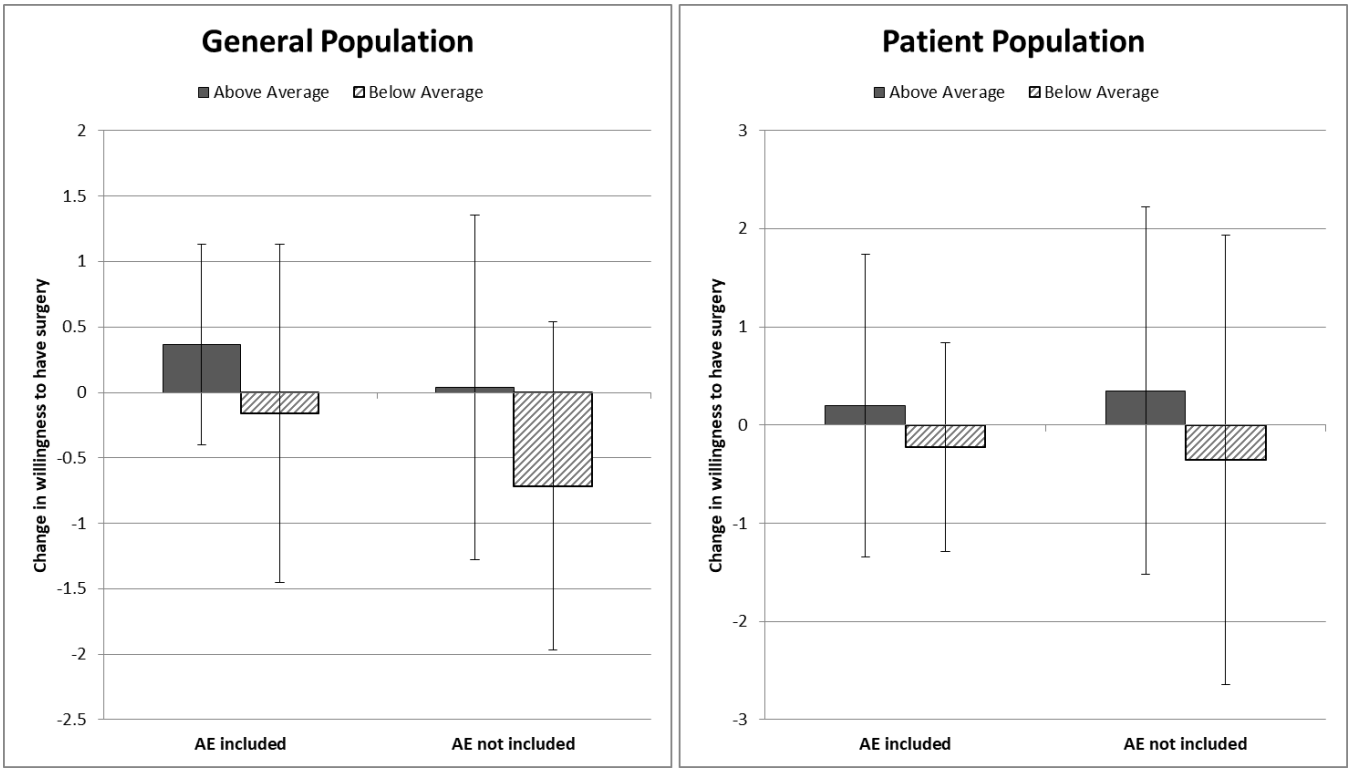


Figure 54. Comparison of change in willingness to have surgery between general and patient populations with above and below average survival probabilities by inclusion or exclusion of AE information

There is no effect from the interaction of graph type and AE information included or not on the change in willingness to have surgery in any of the populations, above or below average survival.

5. Perception of personalization

There is no statisically significant difference in the personalization perceived by the survey respondents by population type, data visualization, inclusion of AE, or reveal of personal survival probability above or below the average.

Overall patients rated the information as being more personalized than the general population (80% over 5 versus 70% over 5, respectively) (Figure 55).

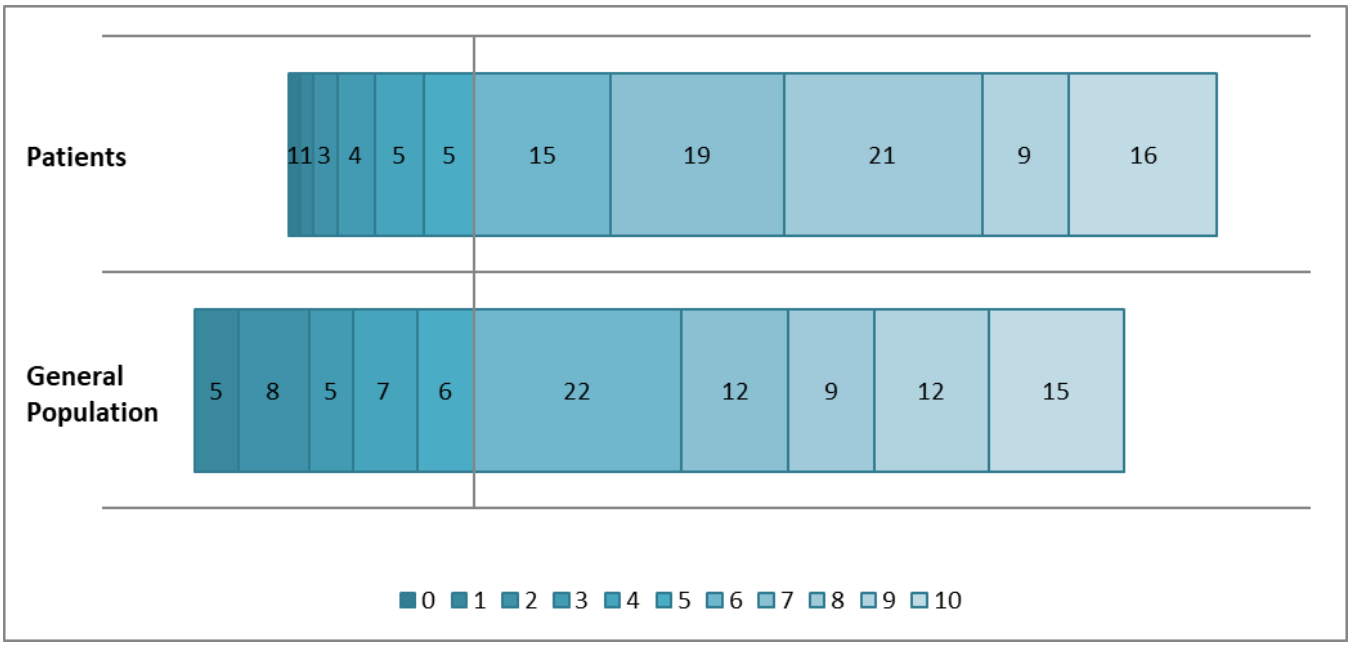
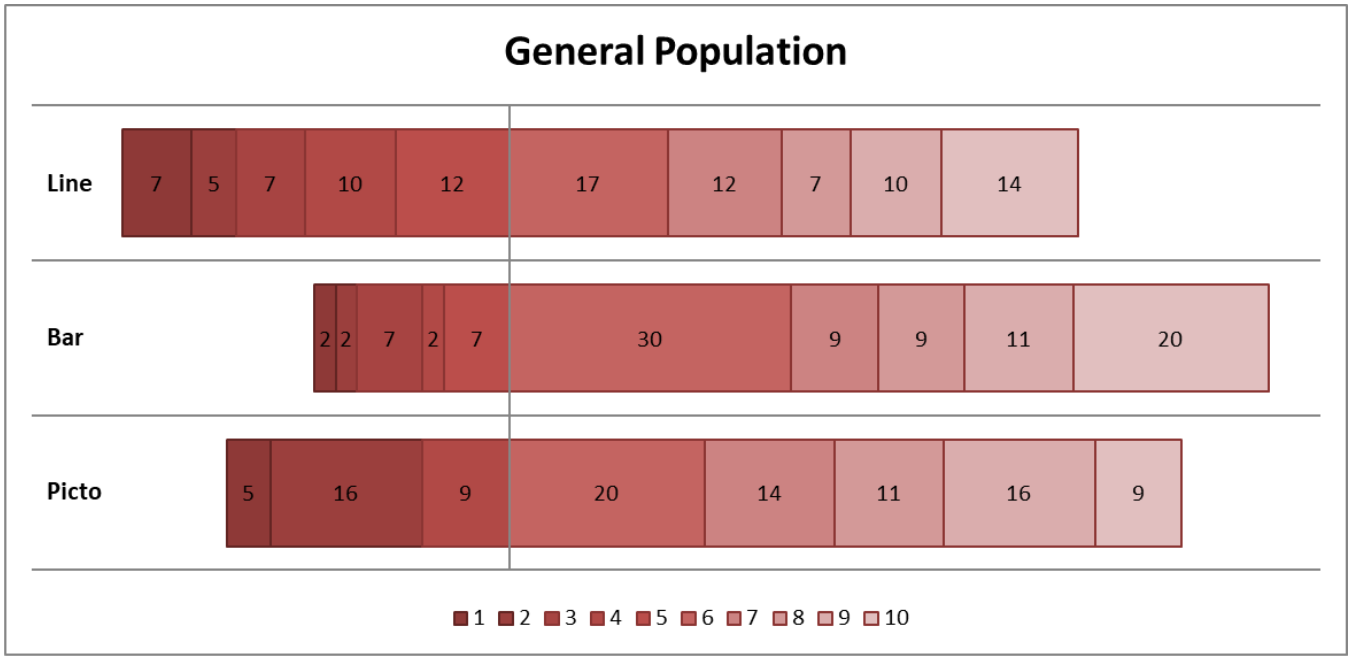


Figure 55. Rank of perceived personalization of information

In the general respondent population, the bar graph has the highest rating of personalization (79.5% over 5) compared to the pictograph (70%) and the line graph (59.5%). This relationship is the same in the patient population, with bar graphs having the highest personalization rating (91.7%) compared to pictographs (80.8%) and line graphs (56%) (Figure 56).



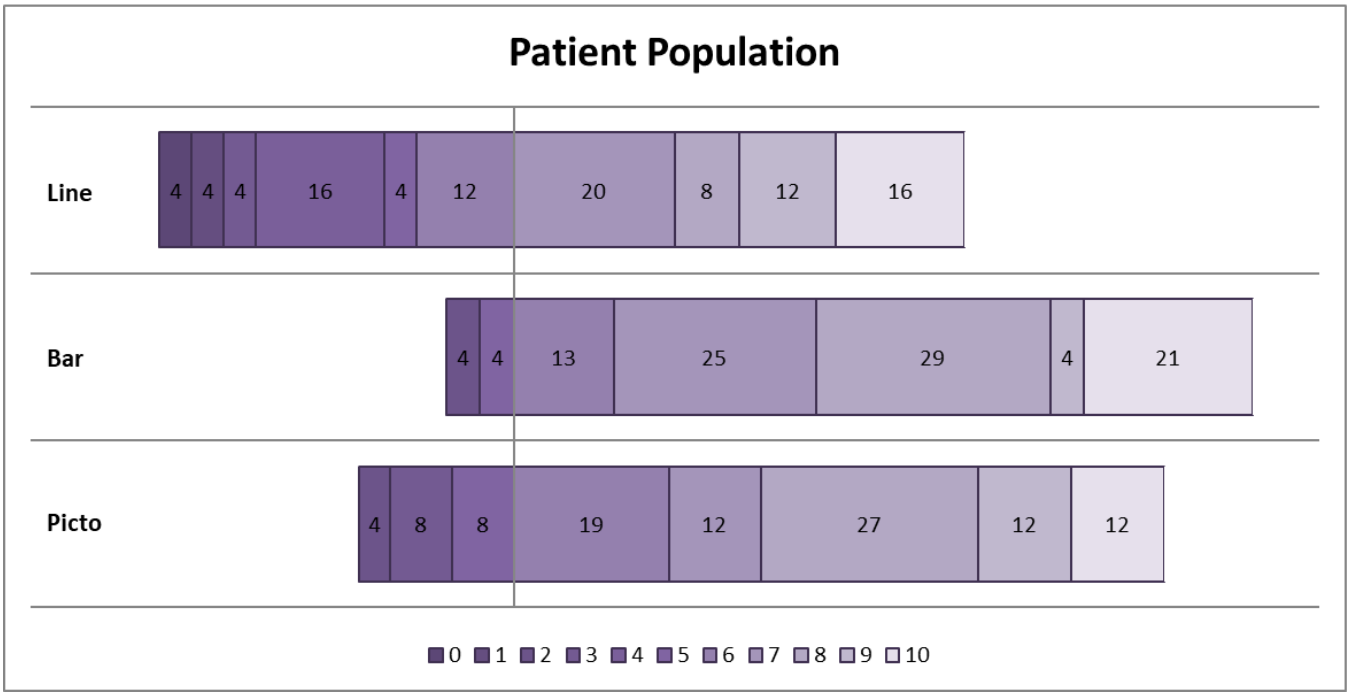


Figure 56. General population (a) and patient population (b) perceived level of personalization by type of visual

When AE information is included in the visualization, general population respondents report that the visual seems less personalized (63.9% over 5) than when AE information is excluded (75.4%). The response is reversed with patient respondents; inclusion of AE information has a higher perception of personalization (81.6% over 5) than without (78.4%) (Figure 57).

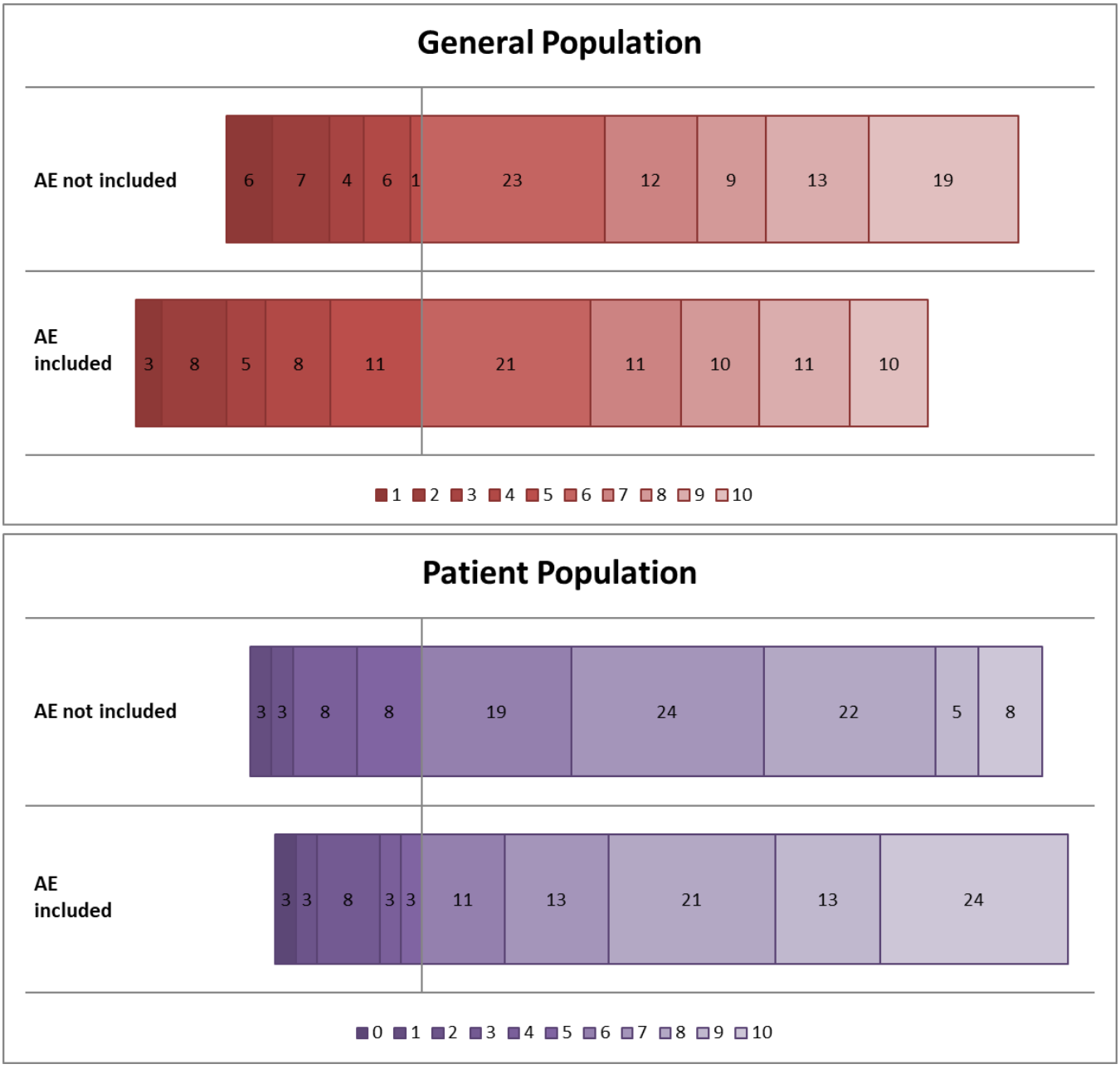


Figure 57. General population (a) and patient population (b) perceived level of personalization by inclusion or exclusion of AE information

Finally, both the general and patient populations who were given a below average survival probability found the visualization to be more personalized (42.5% and 51.4% over 5, respectively.) When their personal survival was above average they rated personalization of the visual at (31.6% and 27.5% over 5, respectively) (Figure 58).

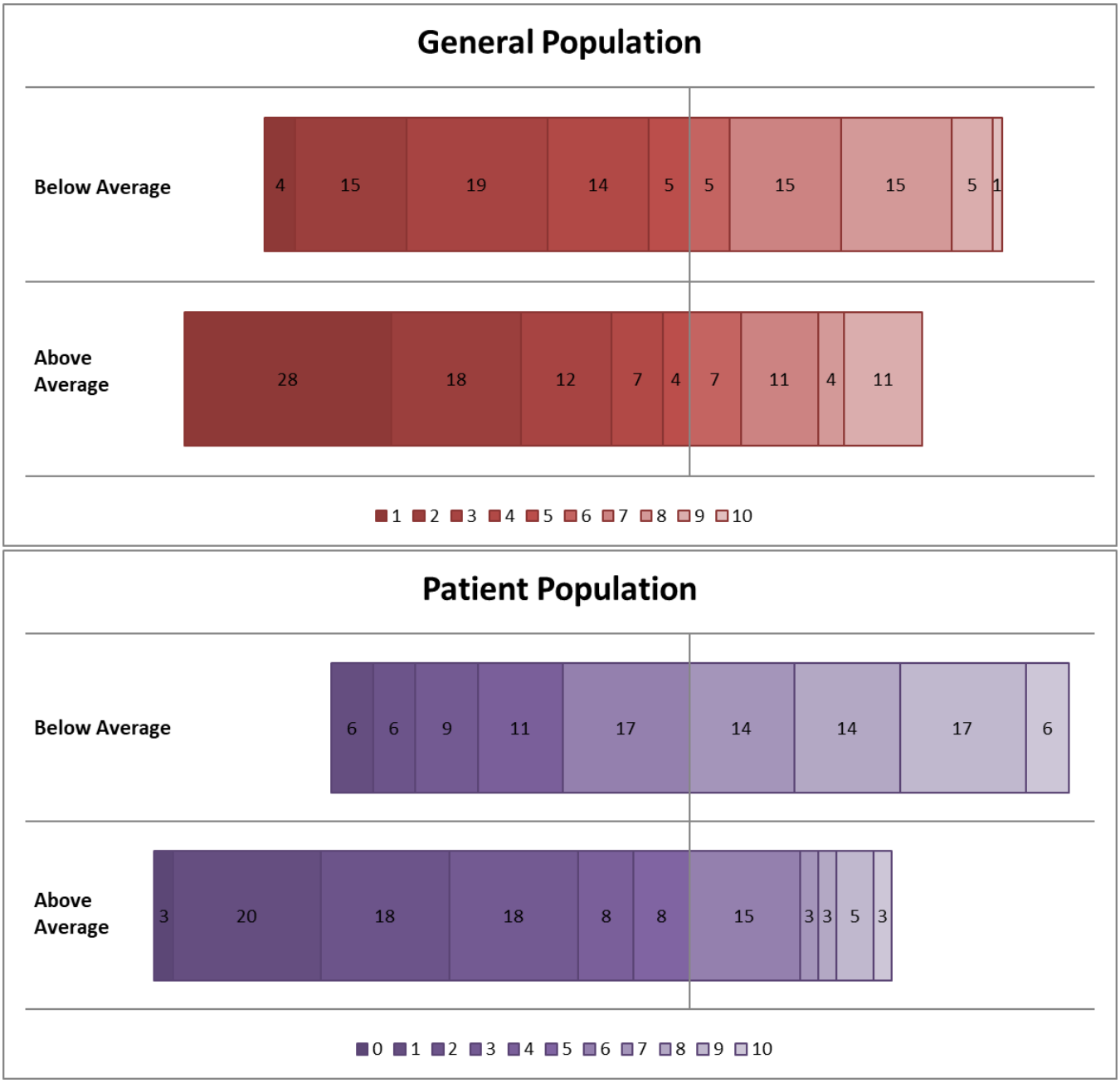


Figure 58. General population (a) and patient population (b) perceived level of personalization by survival information above or below average

4.2.4 Discussion

Two metrics were used to measure the ease of interpretation of graphs: amount of time to answer questions and accuracy of interpretation. Patients were significantly faster at answering questions that the general population. This may be due to the younger overall age of the patient cohort than the general population. Assessment of numeracy has shown that older people (over 55 years of age) had significantly decreased numeracy compared to people 24-54 years of age [130]. Comparing the type of visualizations for speed of interpretation, line graphs were interpreted fastest by the general population and pictographs were interpreted fastest by patients. In terms of accuracy, line graphs were interpreted more correctly in all cases where there was a significant difference. Inclusion of adverse event information did not affect accuracy of interpretation. Taken all together, line graphs had the greatest ease of interpretation in both population types.

Questions asking for the AE and survival to be interpreted at the same time were mostly answered incorrectly, with no correct answers from patients looking at bar graphs or pictographs. Risks that should be interpreted together may need a different style of visualization, such as part-to-whole area graphs to indicate the dependence of the two risks [129].

Effect of visualization of perception of risk was measured by three different areas: risk of dying, risk of AE, and willingness to consider surgery. For the risk of dying and risk of AE, patient and general populations were significantly different, with the patient group expressing a greater perception of risk in both cases. Having experienced varying levels of the heart failure disease progression, this population may be more sensitive to medical risks. Type of visualization and inclusion of AE had no effect on the risk perception.

The use of comparative data after the personalized survival probability reveal influenced risk perception for willingness to consider surgery in the general population. This is in line with a study looking at comparisons between hypothetical and average risk data in a non-patient population showed that people seeing a higher-than-average risk were more likely to want an intervention [131]. Interestingly, this effect was not seen in the patient population. The lack of change in the patient group willingness to consider surgery, despite a difference in perceived risk, suggests that the desire to pursue treatment options is not prognostic-sensitive.

Analysis from the University of Colorado on patient attitudes when considering LVAD implant showed that patient decision making can be characterized as either automatic, deciding without much consideration of the risk data, or reflective, considering the risks, benefits, and burden[17]. Most patients were automatic and automatic deciders all opted for receiving an LVAD, while reflective patients were split: some received the pump and while others declined. In this study, patient response mirrored the automatic deciders, where perception of risk and personalized survival information did not change interest in receiving an intervention.

Risk information's lack of impact on interest in interventions indicates that other types of educational materials may be needed to engage patients in decision making. This could be the use of patient testimonials [132], statements from physicians [133], or values clarification exercises [134]. Future work on development and evaluation of the myCORA patient counselor will explore the effect of these elements on patient risk perceptions.

When determining which visualization was perceived as most personalized to the user, there was no significant different between populations, graphs types, or inclusion of AE information. Directionally, bar graphs were rated as being most personalized and line graphs the least.

4.2.5 Conclusion

The goal of this study was to determine which method of visualization was most easily interpretable for patients and their caregivers and how different visualizations effected risk perception. Line graphs were the best interpreted overall and did not bias risk perception. Patients were more sensitive to risk data in terms of perceiving higher risks but were less sensitive to risk when indicating how willing they would be to have surgery. Future presentation of prognostic data can use line graphs to show single risk types but may need different methods of presentation to show dependent risks. Other types of information than risks, such as patient testimonials or physician recommendations, may be needed to affect patient preferences for treatments.

5.0 CONCLUSIONS

5.1 SUMMARY

Development of a decision support tool requires the right information, presented to the right user, in the right format, at the right time. Determining what constitutes the 'right' approach for each element requires extensive work in the field of not only data analysis, but also behavioral science and decision making, human-computer interaction, and clinical practice. In this thesis, I incorporate research from all these fields to develop decision support tools for both physicians and patients considering LVAD implantation.

New models to predict patient survival after LVAD implant were developed with a current mix of patients, including BTT and DT patients on both axial and centrifugal flow pumps, and elucidate the factors that drive early and late mortality risk. The predictions perform better than current risk scores and provide users with information on which patient features contribute most to their survival predictions.

Models to predict recurrent GI bleeding and ischemic stroke expand the utility of the physician decision support tool, supplying information on the both the most frequent and the deadliest patient adverse event, respectively. Both predictive models identify features of high-risk patients, allowing physicians to consider additional evaluation for high-risk patients.

Presentation of model information to physicians requires an intuitive user interface that fits into their regular workflow. Initial pilot testing of our user interface uncovered issues with layout and interactive elements that I need to address. I also need to validate the model outputs in a site-dependent fashion to prove the model accuracy and utility for physician buy-in.

Patient numeracy influenced patient interest in participating in decision support, indicating that the presentation of prognosis information is a key component to support tool design. Analysis of patient and general population responses to different visual tools found that line graphs were most universally well-interpreted for different prognostic information and did not bias perception of risk. This visual will be further explored in the future for the ability to present multiple risks concurrently and in low-numeracy patient populations.

5.2 FUTURE WORK

There are more outcomes to model for the LVAD patient population. To improve throughput for model building, I have developed an automated machine learning workflow with my colleague, Carmen Khoo. We plan to use the new workflow to model right heart failure (using the latest definitions decided by INTERMACS and ISHLT), hemorrhagic stroke, renal failure, infection, thrombosis and late mortality (3yrs+). I will also continue to maintain and update the current models with new INTERMACS data as it is made available.

In addition to predicting outcomes for patients who received an LVAD, I will create models for patients who *did not* receive mechanical support and instead were on optimal medical management. This is a critical aspect in providing a balanced decision tool for patients and

physicians, showing what the alternative treatments are. The main dataset I intend to use for this model development is from the ROADMAP trial[8].

Content development for the patient counselor will continue, with the patient surveys presented in this thesis work informing the design, along with myLVAD website discussion text analysis, and a future patient interview study. The patient counselor will use "lite" versions of the predictive algorithms from Aim 1 alongside educational information sourced from the literature and physician interviews.

An observational study of heart failure patients receiving treatment information from doctors and thinking through options for their treatment (e.g., whether to receive an LVAD) will be used to identify patient concerns and values during decision making. Analysis for this study will employ the latest in language technologies, including topic discovery and sequence analysis. The design goal is to create a resource that patients can access with their caregivers from home to follow up after discussions with their doctors.

I will evaluate the resulting patient decision support tool for usability through a think-aloud study with LVAD patients and their caregivers. In this study, I will observe how patients and their caregivers navigate through the interface and ask questions about their perceived utility of the tool. Following this pilot test, I plan to apply for PCORI funding to evaluate the tool with a prospective patient population.

APPENDIX

CORA™: CARDIAC OUTCOMES RISK ASSESSMENT

A Personalized Cardiac Counselor for Optimal VAD Therapy

Interview Questions for Patients

Identifier # _____

Date:

This survey consists of four Parts. The first part asks a few questions about you and your health; the second part relates to your interaction with your medical team; the third part asks about your familiarity with technology; and the fourth part relates to your familiarity with decision aids. This entire survey should take about 30 minutes.

PART-1: About You

1. Is this the first time you are taking this survey? ☐ Yes ☐ No
   a. if not the first time, approximately how long ago did you last take the survey?

2. What is the purpose of your visit today?

3. Do you know the diagnosis of your heart disease? How would you describe it?

4. In your opinion, how severe your medical condition? (circle one)

a. My health is about as good as people I know my age.
b. I have heart failure that limits the things I can do, but it's not a major problem.
c. I have heart failure that prevents me from doing some of the things I like to do.
d. I have severe heart failure that might kill me eventually (a year from now.)
e. I have severe heart failure that might prevent me from ever leaving the hospital.

5. ----------D
o you think your condition is so severe that you need a heart transplant? Yes / No

6. ----------I
f you were offered the option of a heart transplant, would you accept it? Yes / No

7. How familiar are you with ventricular assist devices? (circle one)

a. Never heard of them
b. Somewhat familiar: I've heard of them, but not sure what they are, how they work, or what good they could do for me.
c. Very familiar: I've read up on them and/or my doctor has told me about them and I have a good understanding about how they could (or could not) help me.

Part 2: Interaction with your medical team.

1. Do you know who your heart failure cardiologist is? (circle one)
   a. Yes, his / her name is _____
   b. Yes, I met him / her, but I cannot remember his name.
   c. I really don't know.
2. Do you feel comfortable discussing your physical and emotional state with your physicians?
   a. I am comfortable discussing both my physical and emotional state
   b. I am comfortable discussing my physical condition, but not my feelings or emotions.
   c. I am generally uncomfortable asking questions about my physical and emotional state.
3. Which of the following methods of communication makes you feel most comfortable asking questions from you doctor? (circle one)

   a) In person b) Over the phone c) By email
4. If you had a safe and secure way of communicating with your medical team using either an email or messaging system, would you consider using it? (circle one)

   a) Yes b) No c) Maybe
5. How interested are you in understanding your condition? (circle one)
   a. I am very interested in learning everything I can about my condition.
   b. I am somewhat interested.
   c. I rely on the experts who know what they are doing.
6. Is there anything in particular you wish you knew more about? (i.e. more information to help interpret your test results or to better understand your treatment options)

7. About how much <u>total</u> time have you spent speaking with your doctor about your condition prior to taking this survey?

   a. less than 15 minutes b. 15-30 minutes c. 30-60 minutes d. over an hour

8. Do you feel you spent adequate time, or wish you could spent more time with you doctor?
   a. I am satisfied with the time spent with my doctor.
   b. I was not able to ask all the questions of my doctor, but the staff (nurses, coordinators, etc.) were able to fill in my missing questions.
   c. I wish I had more time to ask questions of my doctor.
   d. I was satisfied at first, but later remembered questions I wish I had asked.

9. If you had access to your electronic health records, would you look at them and try to understand it?
   a. Yes, I am eager to look at my records
   b. No, I am not really interested in my records.
   c. No, I don't think I would understand my records.
   d. No, for another reason: _____

PART-3: Your familiarity with technology.

8. Have you done any internet research in the past regarding your heart failure?
   a. No: Proceed to next question.
   b. Yes: we would like to know what sites you visited, and your impressions of their helpfulness

	Did not visit	Visited: not useful.	Visited: useful
WebMD	☐	☐	☐
HeartHope	☐	☐	☐
Other:_____*	☐	☐	☐
Other:_____*	☐	☐	☐

   * if you cannot remember the name, you can leave this blank.

9. If you visited one of these sites, was there any information that you were unable to find?

10. Have you ever requested access to your medical records? (circle one)
    a. No.
    b. Yes, but I really could not understand the information.
    c. Yes, I found it to be informative.

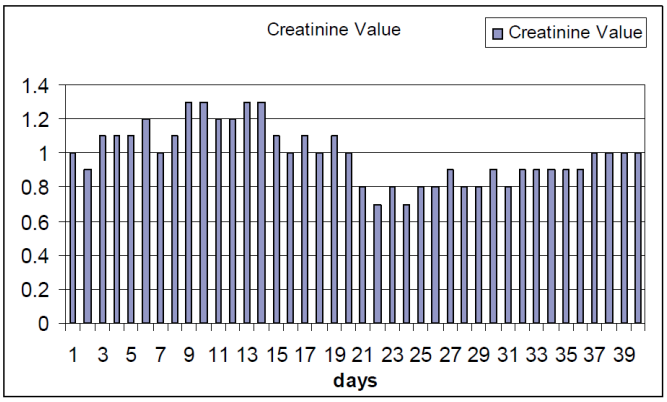
11. How frequently do you use the following electronic devices?

	Every day	Occasionally	Never
Smart phone (like iPhone)	☐	☐	☐
Computer (laptop or desktop)	☐	☐	☐
Computer tablet	☐	☐	☐

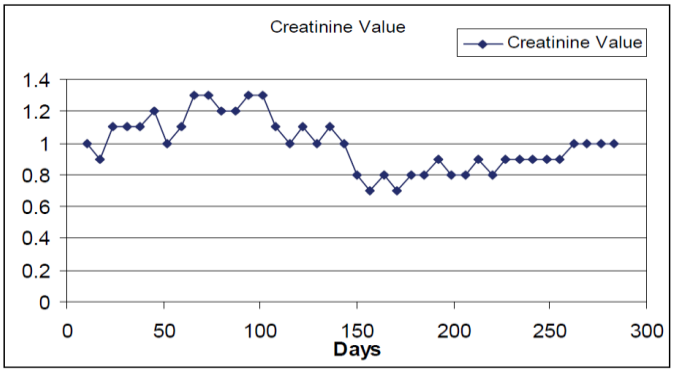
12. How comfortable are you understanding and interpreting the following types of graphs?

	Do not understand	Understand somewhat	Understand well
Bar Graph	☐	☐	☐
Line Graph	☐	☐	☐
Pie Chart	☐	☐	☐
Survival Chart	☐	☐	☐

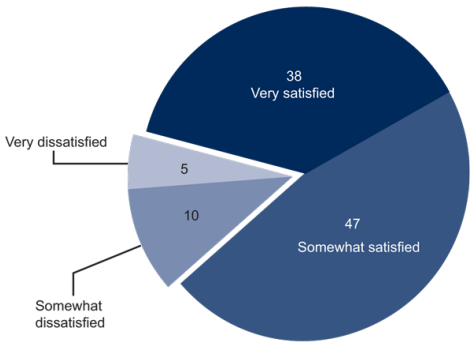
Bar Graph



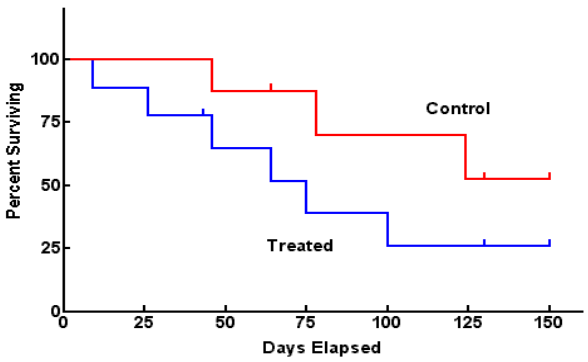
Line Plot



Pie Chart

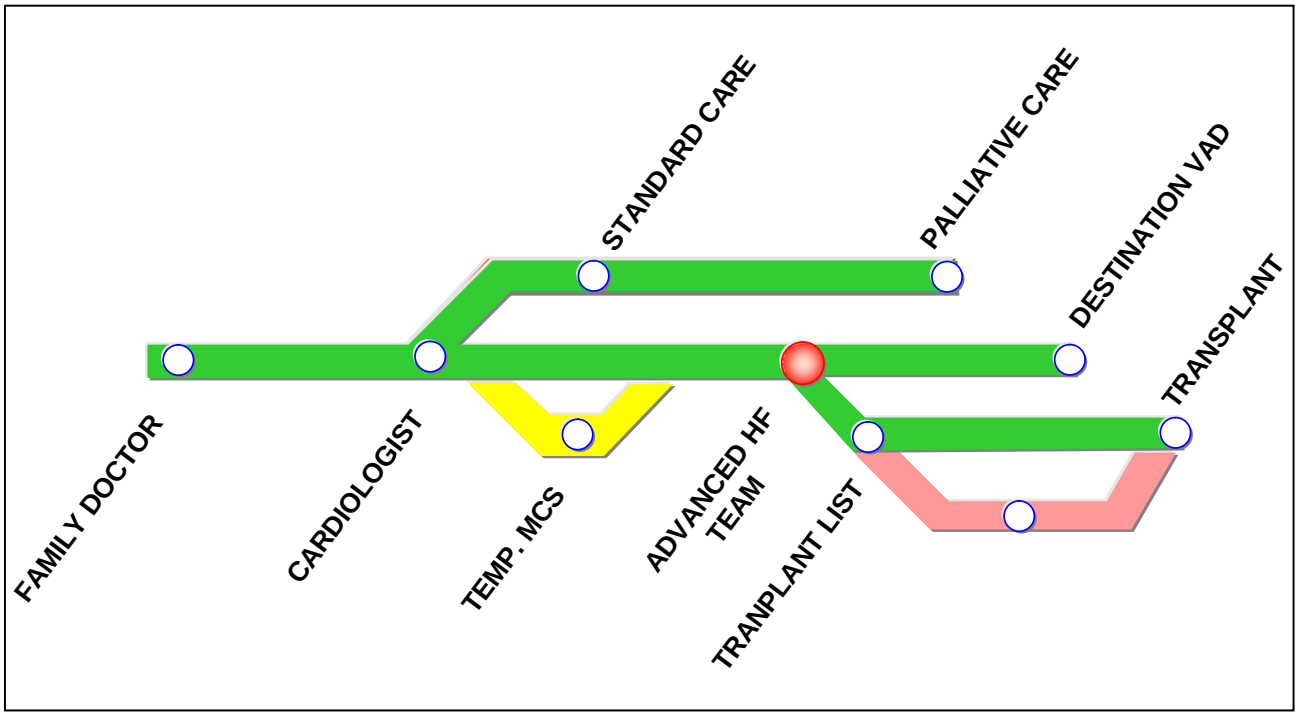


Survival Curve

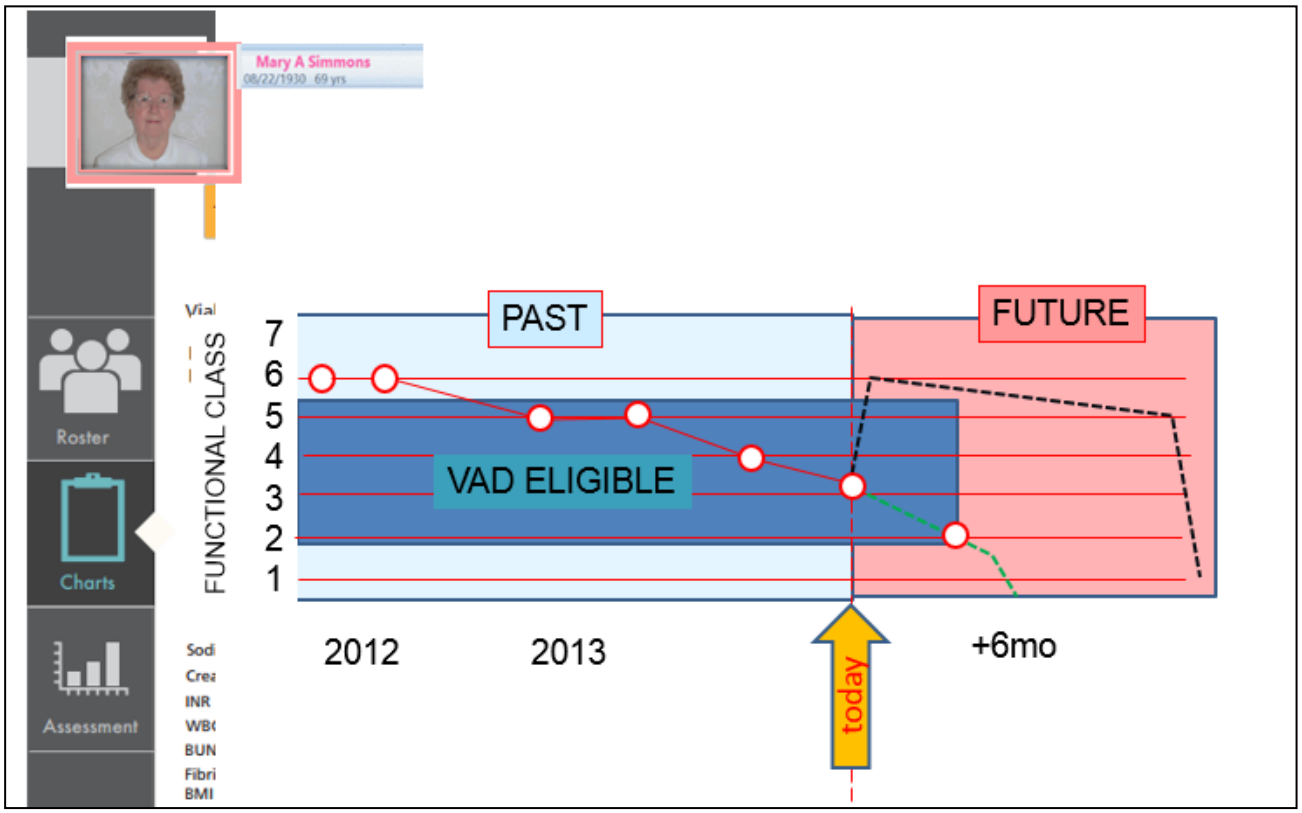


Part 4: Decision Aids and Shared Decision Making

13. Have you ever used a decision tool… like for buying a car or choosing a college?
    a) yes (explain) _____ b) no
14. Which of the following statements best describes how you feel about your involvement in your treatment?
    a) I feel like I have control over what treatments I receive and when.
    b) I have no say whatsoever, the doctors just do what they want and never ask me.
    c) I feel like I'm *too involved* … the doctors can't make a decision on their own, without asking me.
    d) none of the above
15. If you were given a "roadmap" that shows the progression of your health, and the decision points in your care, would you find that useful?
    a. Yes, I think it would be very useful.
    b. I think it would be somewhat useful.
    c. I don't think it would be useful for me.



16. If there was a website or computer program that would show your prognosis (risk of death, becoming more sick, or side-effects of treatment) would that be useful?
    a. Yes, I think it would be very useful.
    b. I think it would be somewhat useful.
    c. I don't think it would be useful for me.



17. If there was a website where you could watch short videos of other patients like you telling stories of their experiences, would that interest you?
    yes maybe no

18. If you can imagine a computer "wizard" that could answer all your questions about your health, or your treatment choices, what would you ask?

This concludes the survey. Thanks for your participation!

BIBLIOGRAPHY

1. (2015). What is Heart Failure? In Health Topics, Volume 2017. (National Heart, Lung, and Blood Institute: U.S. Department of Health & Human Services).

2. (2017). Classes of Heart Failure. In About Heart Failure, Volume 2017. (American Heart Association).

3. Inamdar, A.A., and Inamdar, A.C. (2016). Heart Failure: Diagnosis, Management and Utilization. Journal of Clinical Medicine *5*, 62.

4. Stevenson, L.W., Pagani, F.D., Young, J.B., Jessup, M., Miller, L., Kormos, R.L., Naftel, D.C., Ulisney, K., Desvigne-Nickens, P., and Kirklin, J.K. (2009). INTERMACS profiles of advanced heart failure: the current picture. J Heart Lung Transplant *28*, 535-541.

5. Patel, C. (2015). Update in Heart Failure Management. (Health & Medicine Slide Share: Duke Heart Center), p. 51.

6. Allen, L.A., Stevenson, L.W., Grady, K.L., Goldstein, N.E., Matlock, D.D., Arnold, R.M., Cook, N.R., Felker, G.M., Francis, G.S., Hauptman, P.J., et al. (2012). Decision making in advanced heart failure: a scientific statement from the American Heart Association. Circulation *125*, 1928-1952.

7. Kirklin, J.K., Naftel, D.C., Pagani, F.D., Kormos, R.L., Stevenson, L.W., Blume, E.D., Myers, S.L., Miller, M.A., Baldwin, J.T., and Young, J.B. (2015). Seventh INTERMACS annual report: 15,000 patients and counting. J Heart Lung Transplant *34*, 1495-1504.

8. Estep, J.D., Starling, R.C., Horstmanshof, D.A., Milano, C.A., Selzman, C.H., Shah, K.B., Loebe, M., Moazami, N., Long, J.W., Stehlik, J., et al. (2015). Risk Assessment and Comparative Effectiveness of Left Ventricular Assist Device and Medical Management in Ambulatory Heart Failure Patients: Results From the ROADMAP Study. Journal of the American College of Cardiology *66*, 1747-1761.

9. Magid, M., Jones, J., Allen, L.A., McIlvennan, C.K., Magid, K., Thompson, J.S., and Matlock, D.D. (2016). The Perceptions of Important Elements of Caregiving for a Left Ventricular Assist Device Patient: A Qualitative Meta-Synthesis. The Journal of cardiovascular nursing *31*, 215-225.

10. Birati, E.Y.J., Mariell (2015). Left Ventricular Assist Devices in the Management of Heart Failure1. Cardiac Failure Review *1*, 25-30.

11. Russell, S.B. (2011). HeartMate II Left Ventricular Assist Device Fact Sheet. (Thoratec Corporation), p. 3.

12. GlobeNewswire (2017). Medtronic HeartWare(TM) HVAD(TM) System Approved for Destination Therapy. In Patients with End-Stage Heart Failure Now Have New Options for Care. (NASDAQ Corporate Solutions: Medtronic plc).

13. Schumer, E.M., Black, M.C., Monreal, G., and Slaughter, M.S. (2016). Left ventricular assist devices: current controversies and future directions. Eur Heart J *37*, 3434-3439.

14. Najjar, S.S., Slaughter, M.S., Pagani, F.D., Starling, R.C., McGee, E.C., Eckman, P., Tatooles, A.J., Moazami, N., Kormos, R.L., Hathaway, D.R., et al. (2014). An analysis of pump thrombus events in patients in the HeartWare ADVANCE bridge to transplant and continued access protocol trial. The Journal of Heart and Lung Transplantation *33*, 23-34.

15. Smith, E.M., and Franzwa, J. (2015). Chronic outpatient management of patients with a left ventricular assist device. J Thorac Dis *7*, 2112-2124.

16. Kirklin, J.K., Pagani, F.D., Kormos, R.L., Stevenson, L.W., Blume, E.D., Myers, S.L., Miller, M.A., Baldwin, J.T., Young, J.B., and Naftel, D.C. (2017). Eighth annual INTERMACS report: Special focus on framing the impact of adverse events. J Heart Lung Transplant *36*, 1080-1086.

17. McIlvennan, C.K., Jones, J., Allen, L.A., Lindenfeld, J., Swetz, K.M., Nowels, C., and Matlock, D.D. (2015). Decision Making for Destination Therapy Left Ventricular Assist Devices: Implications for Caregivers. Circulation. Cardiovascular quality and outcomes *8*, 172-178.

18. Slaughter, M.S., Pagani, F.D., Rogers, J.G., Miller, L.W., Sun, B., Russell, S.D., Starling, R.C., Chen, L., Boyle, A.J., Chillcott, S., et al. (2010). Clinical management of continuous-flow left ventricular assist devices in advanced heart failure. The Journal of Heart and Lung Transplantation *29*, S1-S39.

19. Bruce, C.R. (2013). A Review of Ethical Considerations for Ventricular Assist Device Placement in Older Adults. Aging and Disease *4*, 100-112.

20. Pagani, F.D., Aaronson, K.D., Kormos, R., Mann, D.L., Spino, C., Jeffries, N., Taddei-Peters, W.C., Mancini, D.M., McNamara, D.M., Grady, K.L., et al. The NHLBI REVIVE-IT study: Understanding its discontinuation in the context of current left ventricular assist device therapy. The Journal of Heart and Lung Transplantation *35*, 1277-1283.

21. Feldman, D., Pamboukian, S.V., Teuteberg, J.J., Birks, E., Lietz, K., Moore, S.A., Morgan, J.A., Arabia, F., Bauman, M.E., Buchholz, H.W., et al. (2013). The 2013

International Society for Heart and Lung Transplantation Guidelines for mechanical circulatory support: executive summary. J Heart Lung Transplant *32*, 157-187.

22. Cook, J.L., Colvin, M., Francis, G.S., Grady, K.L., Hoffman, T.M., Jessup, M., John, R., Kiernan, M.S., Mitchell, J.E., Pagani, F.D., et al. (2017). Recommendations for the Use of Mechanical Circulatory Support: Ambulatory and Community Patient Care: A Scientific Statement From the American Heart Association. Circulation.

23. Worzala, C. (2009). Policy Update: Federal Incentives for the Adoption of Electronic Health Records. Journal of Oncology Practice *5*, 262-263.

24. Campbell, R.J. (2013). The Five Rights of Clinical Decision Support: CDS Tools Helpful for Meeting Meaningful Use. Volume 84. (Journal of AHIMA), pp. 42-47.

25. Elwyn, G., O'Connor, A.M., Bennett, C., Newcombe, R.G., Politi, M., Durand, M.-A., Drake, E., Joseph-Williams, N., Khangura, S., Saarimaki, A., et al. (2009). Assessing the Quality of Decision Support Technologies Using the International Patient Decision Aid Standards instrument (IPDASi). PLoS ONE *4*, e4705.

26. Teuteberg, J.J., Ewald, G.A., Adamson, R.M., Lietz, K., Miller, L.W., Tatooles, A.J., Kormos, R.L., Sundareswaran, K.S., Farrar, D.J., and Rogers, J.G. (2012). Risk assessment for continuous flow left ventricular assist devices: does the destination therapy risk score work? An analysis of over 1,000 patients. Journal of the American College of Cardiology *60*, 44-51.

27. Cowger, J., Sundareswaran, K., Rogers, J.G., Park, S.J., Pagani, F.D., Bhat, G., Jaski, B., Farrar, D.J., and Slaughter, M.S. (2013). Predicting survival in patients receiving continuous flow left ventricular assist devices: the HeartMate II risk score. J Am Coll Cardiol *61*, 313-321.

28. Lanfear, D.E., Levy, W.C., Stehlik, J., Estep, J.D., Rogers, J.G., Shah, K.B., Boyle, A.J., Chuang, J., Farrar, D.J., and Starling, R.C. (2017). Accuracy of Seattle Heart Failure Model and HeartMate II Risk Score in Non-Inotrope-Dependent Advanced Heart Failure Patients: Insights From the ROADMAP Study (Risk Assessment and Comparative Effectiveness of Left Ventricular Assist Device and Medical Management in Ambulatory Heart Failure Patients). Circulation. Heart failure *10*.

29. Ravichandran, A.K., and Cowger, J. (2015). Left ventricular assist device patient selection: do risk scores help? Journal of Thoracic Disease *7*, 2080-2087.

30. Levy, W.C., Mozaffarian, D., Linker, D.T., Sutradhar, S.C., Anker, S.D., Cropp, A.B., Anand, I., Maggioni, A., Burton, P., Sullivan, M.D., et al. (2006). The Seattle Heart Failure Model. Prediction of Survival in Heart Failure *113*, 1424-1433.

31. Peura, J.L., Colvin-Adams, M., Francis, G.S., Grady, K.L., Hoffman, T.M., Jessup, M., John, R., Kiernan, M.S., Mitchell, J.E., O'Connell, J.B., et al. (2012). Recommendations for the Use of Mechanical Circulatory Support: Device Strategies and Patient Selection. A Scientific Statement From the American Heart Association *126*, 2648-2667.

32. Matthews, J.C., Pagani, F.D., Haft, J.W., Koelling, T.M., Naftel, D.C., and Aaronson, K.D. (2010). Model for end-stage liver disease score predicts left ventricular assist device operative transfusion requirements, morbidity, and mortality. Circulation *121*, 214-220.

33. McIlvennan, C.K., Thompson, J.S., Matlock, D.D., Cleveland, J.C., Jr., Dunlay, S.M., LaRue, S.J., Lewis, E.F., Patel, C.B., Walsh, M.N., and Allen, L.A. (2016). A Multicenter Trial of a Shared Decision Support Intervention for Patients and Their Caregivers Offered Destination Therapy for Advanced Heart Failure: DECIDE-LVAD: Rationale, Design, and Pilot Data. The Journal of cardiovascular nursing *31*, E8-e20.

34. Kostick, K.D., Estevan D.; Wilhelms, Lidija A.; Bruce, Courtenay R.; Estep, Jerry D.; Loebe, Matthias; Minard, Charles; and Blumenthal-Barby, Jennifer S. (2016). Development and Pilot-Testing of a Patient Decision Aid for Left Ventricular Assist Device Placement. the VAD Journal *2*.

35. Yang, Q., Zimmerman, J., Steinfeld, A., Carey, L., and Antaki, J.F. (2016). Investigating the Heart Pump Implant Decision Process: Opportunities for Decision Support Tools to Help. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. (San Jose, California, USA: ACM), pp. 4477-4488.

36. Loghmanpour, N.A., Kanwar, M.K., Druzdzel, M.J., Benza, R.L., Murali, S., and Antaki, J.F. (2015). A new Bayesian network-based risk stratification model for prediction of short-term and long-term LVAD mortality. ASAIO J *61*, 313-323.

37. Santelices, L.C., Wang, Y., Severyn, D., Druzdzel, M.J., Kormos, R.L., and Antaki, J.F. (2010). Developing a Hybrid Decision Support Model for Optimal Ventricular Assist Device Weaning. The Annals of Thoracic Surgery *90*, 713-720.

38. Lohmueller, L.C., Alexander, W.A., and Antaki, J.F. (2017). Predicting Recurrent GI Bleeding in Patients with CF-LVADs Using Pre-Implant Data. The Journal of Heart and Lung Transplantation *36*, S123-S124.

39. Loghmanpour, N.A., Kormos, R.L., Kanwar, M.K., Teuteberg, J.J., Murali, S., and Antaki, J.F. (2016). A Bayesian Model to Predict Right Ventricular Failure Following Left Ventricular Assist Device Therapy. JACC Heart Fail *4*, 711-721.

40. Thomas, S.S., Nahumi, N., Han, J., Lippel, M., Colombo, P., Yuzefpolskaya, M., Takayama, H., Naka, Y., Uriel, N., and Jorde, U.P. (2014). Pre-operative mortality risk assessment in patients with continuous-flow left ventricular assist devices: application of the HeartMate II risk score. J Heart Lung Transplant *33*, 675-681.

41. Kanwar, M.K., Lohmueller, L.C., Kormos, R.L., Loghmanpour, N.A., Benza, R.L., Mentz, R.J., Bailey, S.H., Murali, S., and Antaki, J.F. (2016). Low Accuracy of the HeartMate Risk Score for Predicting Mortality using the INTERMACS Registry Data. ASAIO J.

42. Sabashnikov, A., Mohite, P.N., Zych, B., Garcia, D., Popov, A.F., Weymann, A., Patil, N.P., Hards, R., Capoccia, M., Wahlers, T., et al. (2014). Outcomes and predictors of

early mortality after continuous-flow left ventricular assist device implantation as a bridge to transplantation. ASAIO J *60*, 162-169.

43. Ravichandran, A.K., and Cowger, J. (2015). Left ventricular assist device patient selection: do risk scores help? J Thorac Dis *7*, 2080-2087.

44. Cowger, J.A., Castle, L., Aaronson, K.D., Slaughter, M.S., Moainie, S., Walsh, M., and Salerno, C. (2016). The HeartMate II Risk Score: An Adjusted Score for Evaluation of All Continuous-Flow Left Ventricular Assist Devices. Asaio j *62*, 281-285.

45. Cowger, J.A., Stulak, J.M., Shah, P., Dardas, T.F., Pagani, F.D., Dunlay, S.M., Maltais, S., Aaronson, K.D., Singh, R., Mokadam, N.A., et al. (2017). Impact of Center Left Ventricular Assist Device Volume on Outcomes After Implantation: An INTERMACS Analysis. JACC Heart Fail *5*, 691-699.

46. Kanwar, M.K., Lohmueller, L.C., Kormos, R.L., Loghmanpour, N.A., Benza, R.L., Mentz, R.J., Bailey, S.H., Murali, S., and Antaki, J.F. (2016). Low Accuracy of the HeartMate Risk Score for Predicting Mortality using the INTERMACS Registry Data. ASAIO Journal *Publish Ahead of Print*.

47. Tromp, T.R., de Jonge, N., and Joles, J.A. (2015). Left ventricular assist devices: a kidney's perspective. Heart Failure Reviews *20*, 519-532.

48. Aggarwal, A., Pant, R., Kumar, S., Sharma, P., Gallagher, C., Tatooles, A.J., Pappas, P.S., and Bhat, G. (2012). Incidence and Management of Gastrointestinal Bleeding With Continuous Flow Assist Devices. The Annals of Thoracic Surgery *93*, 1534-1540.

49. Morgan, J.A., Paone, G., Nemeh, H.W., Henry, S.E., Patel, R., Vavra, J., Williams, C.T., Lanfear, D.E., Tita, C., and Brewer, R.J. (2012). Gastrointestinal bleeding with the HeartMate II left ventricular assist device. J Heart Lung Transplant *31*, 715-718.

50. Li, F., Hinton, A., Chen, A., Mehta, N.K., Eldika, S., Zhang, C., Hussan, H., Conwell, D.L., and Krishna, S.G. (2017). Left Ventricular Assist Devices Impact Hospital Resource Utilization Without Affecting Patient Mortality in Gastrointestinal Bleeding. Dig Dis Sci *62*, 150-160.

51. Stulak, J.M., Lee, D., Haft, J.W., Romano, M.A., Cowger, J.A., Park, S.J., Aaronson, K.D., and Pagani, F.D. (2014). Gastrointestinal bleeding and subsequent risk of thromboembolic events during support with a left ventricular assist device. J Heart Lung Transplant *33*, 60-64.

52. Sparrow, C.T., Nassif, M.E., Raymer, D.S., Novak, E., LaRue, S.J., and Schilling, J.D. (2015). Pre-Operative Right Ventricular Dysfunction Is Associated With Gastrointestinal Bleeding in Patients Supported With Continuous-Flow Left Ventricular Assist Devices. JACC Heart Fail *3*, 956-964.

53. Velez, M., and Johnson, M.R. (2009). Management of allosensitized cardiac transplant candidates. Transplant Rev (Orlando) *23*, 235-247.

54. Holley, C.T., Harvey, L., Roy, S.S., Cogswell, R., Eckman, P., Liao, K., and John, R. (2015). Gastrointestinal Bleeding during Continuous-Flow Left Ventricular Assist Device Support is Associated with Lower Rates of Cardiac Transplantation. ASAIO journal (American Society for Artificial Internal Organs : 1992) *61*, 635-639.

55. Wever-Pinzon, O., Selzman, C.H., Drakos, S.G., Saidi, A., Stoddard, G.J., Gilbert, E.M., Labedi, M., Reid, B.B., Davis, E.S., Kfoury, A.G., et al. (2013). Pulsatility and the risk of nonsurgical bleeding in patients supported with the continuous-flow left ventricular assist device HeartMate II. Circ Heart Fail *6*, 517-526.

56. Guha, A., Eshelbrenner, C.L., Richards, D.M., and Monsour, H.P., Jr. (2015). Gastrointestinal bleeding after continuous-flow left ventricular device implantation: review of pathophysiology and management. Methodist Debakey Cardiovasc J *11*, 24-27.

57. Draper, K.V., Huang, R.J., and Gerson, L.B. (2014). GI bleeding in patients with continuous-flow left ventricular assist devices: a systematic review and meta-analysis. Gastrointest Endosc *80*, 435-446 e431.

58. Sami, S.S., Al-Araji, S.A., and Ragunath, K. (2014). Review article: gastrointestinal angiodysplasia – pathogenesis, diagnosis and management. Alimentary Pharmacology & Therapeutics *39*, 15-34.

59. Meyer, A.L., Malehsa, D., Budde, U., Bara, C., Haverich, A., and Strueber, M. (2014). Acquired von Willebrand syndrome in patients with a centrifugal or axial continuous flow left ventricular assist device. JACC Heart Fail *2*, 141-145.

60. Tabit, C.E., Chen, P., Kim, G.H., Fedson, S.E., Sayer, G., Coplan, M.J., Jeevanandam, V., Uriel, N., and Liao, J.K. (2016). Elevated Angiopoietin-2 Level in Patients With Continuous-Flow Left Ventricular Assist Devices Leads to Altered Angiogenesis and Is Associated With Higher Nonsurgical Bleeding. Circulation *134*, 141-152.

61. Baumann Kreuziger, L.M., Kim, B., and Wieselthaler, G.M. (2015). Antithrombotic therapy for left ventricular assist devices in adults: a systematic review. Journal of Thrombosis and Haemostasis *13*, 946-955.

62. Kurgan, L.A., and Cios, K.J. (2004). CAIM Discretization Algorithm. IEEE Trans. on Knowl. and Data Eng. *16*, 145-153.

63. Gurvits, G.E., and Fradkov, E. (2017). Bleeding with the artificial heart: Gastrointestinal hemorrhage in CF-LVAD patients. World Journal of Gastroenterology *23*, 3945-3953.

64. Cochrane, J., Jackson, C., Schlepp, G., and Strong, R. (2016). Gastrointestinal angiodysplasia is associated with significant gastrointestinal bleeding in patients with continuous left ventricular assist devices. Endoscopy International Open *4*, E371-E377.

65. Joy, P.S., Kumar, G., Guddati, A.K., Bhama, J.K., and Cadaret, L.M. (2016). Risk Factors and Outcomes of Gastrointestinal Bleeding in Left Ventricular Assist Device Recipients. The American Journal of Cardiology *117*, 240-244.

66. Chait, M.M. (2010). Lower gastrointestinal bleeding in the elderly. World Journal of Gastrointestinal Endoscopy *2*, 147-154.

67. (2015). Appendix M. Intermacs(r) Site Users' Guide. In Manual of Operatins and Procedures Version 4.0. (INTERMACS), pp. 1-105.

68. Ravi, Y., Lella, S.K., Copeland, L.A., Zolfaghari, K., Grady, K., Emani, S., and Sai-Sudhakar, C.B. (2018). Does recipient work status pre-transplant affect post-heart transplant survival? A United Network for Organ Sharing database review. J Heart Lung Transplant.

69. Heilmann, C., Geisen, U., Benk, C., Berchtold-Herz, M., Trummer, G., Schlensak, C., Zieger, B., and Beyersdorf, F. (2009). Haemolysis in patients with ventricular assist devices: major differences between systems☆. European Journal of Cardio-Thoracic Surgery *36*, 580-584.

70. Tomizawa, M., Shinozaki, F., Hasegawa, R., Shirai, Y., Motoyoshi, Y., Sugiyama, T., Yamamoto, S., and Ishige, N. (2016). Low hemoglobin levels are associated with upper gastrointestinal bleeding. Biomedical Reports *5*, 349-352.

71. Tchantchaleishvili, V., Sagebin, F., Ross, R.E., Hallinan, W., Schwarz, K.Q., and Massey, H.T. (2014). Evaluation and treatment of pump thrombosis and hemolysis. Annals of Cardiothoracic Surgery *3*, 490-495.

72. Hayes, H.M., Dembo, L.G., Larbalestier, R., and O'Driscoll, G. (2010). Management options to treat gastrointestinal bleeding in patients supported on rotary left ventricular assist devices: a single-center experience. Artif Organs *34*, 703-706.

73. Kirklin, J.K., Pagani, F.D., Kormos, R.L., Stevenson, L.W., Blume, E.D., Myers, S.L., Miller, M.A., Baldwin, J.T., Young, J.B., and Naftel, D.C. Eighth annual INTERMACS report: Special focus on framing the impact of adverse events. The Journal of Heart and Lung Transplantation *36*, 1080-1086.

74. Levy, W.C., Mozaffarian, D., Linker, D.T., Sutradhar, S.C., Anker, S.D., Cropp, A.B., Anand, I., Maggioni, A., Burton, P., Sullivan, M.D., et al. (2006). The Seattle Heart Failure Model: prediction of survival in heart failure. Circulation *113*, 1424-1433.

75. Kanwar, M., Lohmueller, L., Kormos, R., Teuteberg, J., Rogers, J., J, L., Bailey, S., McIlvennan, C., Benza, R., Murali, S., et al. (2018). A Bayesian Model to Predict Mortality Following Left Ventricular Assist Device Therapy. JACC HF.

76. Acharya, D., Loyaga-Rendon, R., Morgan, C.J., Sands, K.A., Pamboukian, S.V., Rajapreyar, I., Holman, W.L., Kirklin, J.K., and Tallaj, J.A. (2017). INTERMACS Analysis of Stroke During Support With Continuous-Flow Left Ventricular Assist Devices: Risk Factors and Outcomes. JACC Heart Fail *5*, 703-711.

77. Willey, J.Z., Gavalas, M.V., Trinh, P.N., Yuzefpolskaya, M., Reshad Garan, A., Levin, A.P., Takeda, K., Takayama, H., Fried, J., Naka, Y., et al. (2016). Outcomes after stroke complicating left ventricular assist device. J Heart Lung Transplant *35*, 1003-1009.

78. Harvey, L., Holley, C., Roy, S.S., Eckman, P., Cogswell, R., Liao, K., and John, R. (2015). Stroke After Left Ventricular Assist Device Implantation: Outcomes in the Continuous-Flow Era. Ann Thorac Surg *100*, 535-541.

79. Pagani, F.D., Aaronson, K.D., Kormos, R., Mann, D.L., Spino, C., Jeffries, N., Taddei-Peters, W.C., Mancini, D.M., McNamara, D.M., Grady, K.L., et al. (2016). The NHLBI REVIVE-IT study: Understanding its discontinuation in the context of current left ventricular assist device therapy. J Heart Lung Transplant *35*, 1277-1283.

80. Cho, S.M., Moazami, N., and Frontera, J.A. (2017). Stroke and Intracranial Hemorrhage in HeartMate II and HeartWare Left Ventricular Assist Devices: A Systematic Review. Neurocritical Care *27*, 17-25.

81. Willey, J.Z., Demmer, R.T., Takayama, H., Colombo, P.C., and Lazar, R.M. (2014). Cerebrovascular disease in the era of left ventricular assist devices with continuous flow: Risk factors, diagnosis, and treatment. The Journal of Heart and Lung Transplantation *33*, 878-887.

82. Pinney, S.P., Anyanwu, A.C., Lala, A., Teuteberg, J.J., Uriel, N., and Mehra, M.R. (2017). Left Ventricular Assist Devices for Lifelong Support. Journal of the American College of Cardiology *69*, 2845-2861.

83. Rogers, J.G., Pagani, F.D., Tatooles, A.J., Bhat, G., Slaughter, M.S., Birks, E.J., Boyce, S.W., Najjar, S.S., Jeevanandam, V., Anderson, A.S., et al. (2017). Intrapericardial Left Ventricular Assist Device for Advanced Heart Failure. N Engl J Med *376*, 451-460.

84. DeVore, A.D., and Stewart, G.C. (2017). The Risk of Stroke on Left Ventricular Assist Device Support: Steady Gains or Stalled Progress? JACC Heart Fail *5*, 712-714.

85. Coffin, S.T., Haglund, N.A., Davis, M.E., Xu, M., Dunlay, S.M., Cowger, J.A., Shah, P., Aaronson, K.D., Pagani, F.D., Stulak, J.M., et al. (2015). Adverse neurologic events in patients bridged with long-term mechanical circulatory support: A device-specific comparative analysis. J Heart Lung Transplant *34*, 1578-1585.

86. Nassif, M.E., Tibrewala, A., Raymer, D.S., Andruska, A., Novak, E., Vader, J.M., Itoh, A., Silvestry, S.C., Ewald, G.A., and LaRue, S.J. (2015). Systolic blood pressure on discharge after left ventricular assist device insertion is associated with subsequent stroke. The Journal of Heart and Lung Transplantation *34*, 503-508.

87. Lampert, B.C., Eckert, C., Weaver, S., Scanlon, A., Lockard, K., Allen, C., Kunz, N., Bermudez, C., Bhama, J.K., Shullo, M.A., et al. (2014). Blood pressure control in continuous flow left ventricular assist devices: efficacy and impact on adverse events. Ann Thorac Surg *97*, 139-146.

88. Mooe, T., Eriksson, P., and Stegmayr, B. (1997). Ischemic stroke after acute myocardial infarction. A population-based study. Stroke *28*, 762-767.

89. Suarez, J.I. (2006). Acute myocardial infarction, ischemic stroke, sympathetic stress, and inflammation: birds of a feather. Stroke *37*, 2449-2450.

90. Van de Graaff, E., Dutta, M., Das, P., Shry, E.A., Frederick, P.D., Blaney, M., Pasta, D.J., and Steinhubl, S.R. (2006). Early coronary revascularization diminishes the risk of ischemic stroke with acute myocardial infarction. Stroke *37*, 2546-2551.

91. Mehrpour, M., Khuzan, M., Najimi, N., Motamed, M.R., and Fereshtehnejad, S.-M. (2012). Serum uric acid level in acute stroke patients. Medical Journal of the Islamic Republic of Iran *26*, 66-72.

92. Yang, X.L., Kim, Y., Kim, T.J., Jung, S., Kim, C.K., and Lee, S.H. (2016). Association of serum uric acid and cardioembolic stroke in patients with acute ischemic stroke. Journal of the neurological sciences *370*, 57-62.

93. Wang, L., Hu, W., Miao, D., Zhang, Q., Wang, C., Pan, E., and Wu, M. (2017). Relationship between serum uric acid and ischemic stroke in a large type 2 diabetes population in China: A cross-sectional study. Journal of the neurological sciences *376*, 176-180.

94. Chamorro, A., Obach, V., Cervera, A., Revilla, M., Deulofeu, R., and Aponte, J.H. (2002). Prognostic significance of uric acid serum concentration in patients with acute ischemic stroke. Stroke *33*, 1048-1052.

95. Di Napoli, M., Papa, F., and Bocola, V. (2001). C-reactive protein in ischemic stroke: an independent prognostic factor. Stroke *32*, 917-924.

96. Stulak, J.M., Deo, S., Schirger, J., Aaronson, K.D., Park, S.J., Joyce, L.D., Daly, R.C., and Pagani, F.D. (2013). Preoperative atrial fibrillation increases risk of thromboembolic events after left ventricular assist device implantation. Ann Thorac Surg *96*, 2161-2167.

97. Frontera, J.A., Starling, R., Cho, S.-M., Nowacki, A.S., Uchino, K., Hussain, M.S., Mountis, M., and Moazami, N. (2017). Risk factors, mortality, and timing of ischemic and hemorrhagic stroke with left ventricular assist devices. The Journal of Heart and Lung Transplantation *36*, 673-683.

98. Boyle, A.J., Jorde, U.P., Sun, B., Park, S.J., Milano, C.A., Frazier, O.H., Sundareswaran, K.S., Farrar, D.J., Russell, S.D., and HeartMate, I.I.C.I. (2014). Pre-operative risk factors of bleeding and stroke during left ventricular assist device support: an analysis of more than 900 HeartMate II outpatients. J Am Coll Cardiol *63*, 880-888.

99. Lietz, K., Long, J.W., Kfoury, A.G., Slaughter, M.S., Silver, M.A., Milano, C.A., Rogers, J.G., Miller, L.W., Deng, M., Naka, Y., et al. (2009). Impact of center volume on outcomes of left ventricular assist device implantation as destination therapy: analysis of the Thoratec HeartMate Registry, 1998 to 2005. Circulation. Heart failure *2*, 3-10.

100. Haglund, N.A., Feurer, I.D., Ahmad, R.M., DiSalvo, T.G., Lenihan, D.J., Keebler, M.E., Schlendorf, K.H., Stulak, J.M., Wigger, M.A., and Maltais, S. (2014). Institutional volume of heart transplantation with left ventricular assist device explantation influences graft survival. J Heart Lung Transplant *33*, 931-936.

101. Krim, S.R., Vivo, R.P., Campbell, P., Estep, J., Fonarow, G.C., Naftel, D.C., and Ventura, H.O. (2015). REGIONAL DIFFERENCES IN UTILIZATION AND OUTCOMES OF LEFT VENTRICULAR ASSIST DEVICES: INSIGHTS FROM THE INTERMACS REGISTRY. The Journal of heart and lung transplantation : the official publication of the International Society for Heart Transplantation *34*, 912-920.

102. Osheroff, J.A., Teich, J., Information, H., Society, M.S., Levick, D., Saldana, L., Velasco, F., Sittig, D., Rogers, K., and Jenders, R. (2012). Improving Outcomes with Clinical Decision Support: An Implementer's Guide, (HIMSS).

103. Congress, t.U.S. (2010). Patient Protection and Affordable Care Act. (Public Law), pp. 111-148.

104. Shared Decision Making. (NHS England).

105. Hoffmann, T.C., Legare, F., Simmons, M.B., McNamara, K., McCaffery, K., Trevena, L.J., Hudson, B., Glasziou, P.P., and Del Mar, C.B. (2014). Shared decision making: what do clinicians need to know and why should they bother? Med J Aust *201*, 35-39.

106. Sheridan, S.L., Shadle, J., Simpson, R.J., and Pignone, M.P. (2006). The impact of a decision aid about heart disease prevention on patients' discussions with their doctor and their plans for prevention: a pilot randomized trial. BMC Health Services Research *6*, 121.

107. Hess, E.P., Knoedler, M.A., Shah, N.D., Kline, J.A., Breslin, M., Branda, M.E., Pencille, L.J., Asplin, B.R., Nestler, D.M., Sadosty, A.T., et al. (2012). The chest pain choice decision aid: a randomized trial. Circ Cardiovasc Qual Outcomes *5*, 251-259.

108. Jacqueline Jones, C.N., B. Karen Mellis, Amy Jenkins, Heather Nuanes, Paul Varosy, Richard Thomson, Glyn Elwyn, David J. Magid, Angela Brega, Travis Vermilye, Fred Masoudi, Daniel Matlock (2015). Implementation of decision aids for implantable cardioverter-defibrillators: Lessons learned and patient perspectives In 37th Annual Meeting of the Society for Medical Decision Making. (St. Louis, MO).

109. Thompson, J.S., Matlock, D.D., McIlvennan, C.K., Jenkins, A.R., and Allen, L.A. (2015). Development of a Decision Aid for Patients With Advanced Heart Failure Considering a Destination Therapy Left Ventricular Assist Device. JACC Heart Fail *3*, 965-976.

110. Morgan, M.W., Deber, R.B., Llewellyn-Thomas, H.A., Gladstone, P., Cusimano, R.J., O'Rourke, K., Tomlinson, G., and Detsky, A.S. (2000). Randomized, Controlled Trial of an Interactive Videodisc Decision Aid for Patients with Ischemic Heart Disease. Journal of General Internal Medicine *15*, 685-693.

111. Trevena, L.J., Zikmund-Fisher, B.J., Edwards, A., Gaissmaier, W., Galesic, M., Han, P.K.J., King, J., Lawson, M.L., Linder, S.K., Lipkus, I., et al. (2013). Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. Bmc Med Inform Decis *13*.

112. Gigerenzer, G., and Edwards, A. (2003). Simple tools for understanding risks: from innumeracy to insight. BMJ *327*, 741-744.

113. McIlvennan, C.K., Matlock, D.D., Narayan, M.P., Nowels, C., Thompson, J.S., Cannon, A., Bradley, W.J., and Allen, L.A. (2015). Perspectives from mechanical circulatory support coordinators on the pre-implantation decision process for destination therapy left ventricular assist devices. Heart Lung *44*, 219-224.

114. Eyler, R.F., Cordes, S., Szymanski, B.R., and Fraenkel, L. (2017). Utilization of Continuous "Spinners" to Communicate Risk. Medical Decision Making *37*, 725-729.

115. Faisal, S., Blandford, A., and Potts, H.W.W. (2013). Making sense of personal health information: Challenges for information visualization. Health Informatics Journal *19*, 198-217.

116. Le, T., Chi, N.-C., Chaudhuri, S., Thompson, H.J., and Demiris, G. (2016). Understanding Older Adult Use of Data Visualizations as a Resource for Maintaining Health and Wellness. Journal of Applied Gerontology, 0733464816658751.

117. Weng, H.C. (2009). A multisource and repeated measure approach to assessing patient-physician relationship and patient satisfaction. Eval Health Prof *32*, 128-143.

118. Magnezi, R., Bergman, L.C., and Urowitz, S. (2014). Would Your Patient Prefer to Be Considered Your Friend? Patient Preferences in Physician Relationships. Health Education & Behavior *42*, 210-219.

119. Leto, L., and Feola, M. (2014). Cognitive impairment in heart failure patients. Journal of Geriatric Cardiology : JGC *11*, 316-328.

120. Blei, D.M.N., Andrew Y; Jordan, Michael I (2003). Latent dirichlet allocation. Journal of Machine Learning research *3*, 99-1022.

121. Daly, R., Shen, Q., and Aitken, S. (2011). Learning Bayesian networks: approaches and issues. Knowl Eng Rev *26*, 99-157.

122. Schölkopf, B., Smola, A., and Müller, K.-R. (1998). Nonlinear component analysis as a kernel eigenvalue problem. Neural computation *10*, 1299-1319.

123. Willems, S., De Maesschalck, S., Deveugele, M., Derese, A., and De Maeseneer, J. (2005). Socio-economic status of the patient and doctor–patient communication: does it make a difference? Patient Education and Counseling *56*, 139-146.

124. Mantwill, S., Monestel-Umana, S., and Schulz, P.J. (2015). The Relationship between Health Literacy and Health Disparities: A Systematic Review. PLoS One *10*, e0145455.

125. Smith, S.G., Wolf, M.S., and von Wagner, C. (2010). Socioeconomic status, statistical confidence, and patient-provider communication: an analysis of the Health Information National Trends Survey (HINTS 2007). J Health Commun *15 Suppl 3*, 169-185.

126. Ciampa, P.J., Osborn, C.Y., Peterson, N.B., and Rothman, R.L. (2010). Patient Numeracy, Perceptions of Provider Communication and Colorectal Cancer Screening Utilization. Journal of health communication *15*, 157-168.

127. Gratch, J., Lucas, G.M., King, A.A., and Morency, L.-P. (2014). It's only a computer: the impact of human-agent interaction in clinical interviews. In Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems. (Paris, France: International Foundation for Autonomous Agents and Multiagent Systems), pp. 85-92.

128. Fitzpatrick, K.K., Darcy, A., and Vierhile, M. (2017). Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. JMIR Ment Health *4*, e19.

129. Ancker, J.S., Senathirajah, Y., Kukafka, R., and Starren, J.B. (2006). Design Features of Graphs in Health Risk Communication: A Systematic Review. Journal of the American Medical Informatics Association : JAMIA *13*, 608-618.

130. Prins, E., and Monnat, S. (2015). Examining Associations between Self-Rated Health and Proficiency in Literacy and Numeracy among Immigrants and U.S.-Born Adults: Evidence from the Program for the International Assessment of Adult Competencies (PIAAC). PLoS ONE *10*, e0130257.

131. Fagerlin, A., Zikmund-Fisher, B.J., and Ubel, P.A. (2007). “If I’m better than average, then I’m ok?”: Comparative information influences beliefs about risk and benefits. Patient Education and Counseling *69*, 140-144.

132. Trevena, L.J., Zikmund-Fisher, B.J., Edwards, A., Gaissmaier, W., Galesic, M., Han, P.K.J., King, J., Lawson, M.L., Linder, S.K., Lipkus, I., et al. (2013). Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. Bmc Med Inform Decis *13*, S7.

133. Bruine de Bruin, W., Wallin, A., Parker, A.M., Strough, J., and Hanmer, J. (2017). Effects of Anti- Versus Pro-Vaccine Narratives on Responses by Recipients Varying in Numeracy: A Cross-sectional Survey-Based Experiment. Medical Decision Making *37*, 860-870.

134. Feldman-Stewart, D., Tong, C., Siemens, R., Alibhai, S., Pickles, T., Robinson, J., and Brundage, M.D. (2012). The impact of explicit values clarification exercises in a patient decision aid emerges after the decision is actually made: evidence from a randomized controlled trial. Med Decis Making *32*, 616-626.