A Dynamic Bayesian Network approach for time-specific survival probability prediction in patients after Ventricular Assist Device Implantation

Themis P. Exarchos, *Member, IEEE*, George Rigas, Yorgos Goletsis, *Member, IEEE,* Kostas Stefanou, Steven Jacobs, Maria-Giovanna Trivella and Dimitrios I. Fotiadis, *Senior Member, IEEE*

***Abstract*— In this work we present a decision support tool for the calculation of time-dependent survival probability for patients after ventricular assist device implantation. Two different models have been developed, a short term one which predicts survival for the first three months and a long term one that predicts survival for one year after implantation. In order to model the time dependencies between the different time slices of the problem, a dynamic Bayesian network (DBN) approach has been employed. DBNs order to capture the temporal events of the patient disease and the temporal data availability. High accuracy results have been reported for both models. The short and long term DBNs reached an accuracy of 96.97% and 93.55% respectively.**

I. Introduction

Heart failure (HF) is a disease that affects millions of people in the Western societies with high rates of incidence and prevalence. During the last years Ventricular assist devices have become a valuable option for patients with end-stage heart failure, no longer responding to medical therapies. A population of patients that in the past could only be treated with a heart transplantation. However, several complications persist during VAD support (mainly left-VAD/LVAD) due to pre-existing effects of advanced heart failure, the requirement of extensive surgery to implant the device and the effects of VAD in compromised patients [1]. As patient selection and timing are considered as primary determinants of the success of VAD implantation, the use of models and tools that can assess patient status and the risk of adverse events/death can conduce to increased success of LVAD therapy.

For VAD treated patients, several risk scores and assessment tools have been presented in the literature. The Heart Failure Survival Score (HFSS) [2] and the Seattle Heart Failure Model (SHFM) [3] have been proposed for patient selection for LVAD support based on the estimation for expected survival during the next 1 to 3 years. In a similar way, the Randomized Evaluation of Mechanical Assistance in Treatment of Chronic Heart Failure trial (REMATCH) [4] stratifies patient into risk groups. The Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS) [5] has been used for patient classification in risk groups, interval analysis [6], and timing of implant assessment [7]. Also, patient classification regarding the risk of developing other diseases (multi-organ failures) when undergoing LVAD implantation has been addressed with the Model for End-Stage Liver Disease (MELD) [8] without being specific for LVAD patients. Additional analyses of predictors and mortality risk scores can be found in [9-14]. Recently, an adverse event prediction approach based on data mining methods has been presented [15]. Still, the above approaches are based mostly on pre-operative patient data and do not capture the dynamic nature of patient manifestations/clinical history after implantation.

In this work, we present a decision support tool that predicts the survival probabilities of patients after VAD transplantation at specific time intervals. Such a tool can be a valuable assistant in patient selection and the design of treatment plan. The methodology used to develop the tool is based on dynamic Bayesian networks. The proposed approach goes beyond the state of the art since it encompasses temporal information from the different stages of the patient, after VAD implantation.

II. Materials And Methods

A. Formulation of the time dependent problem

In order to appropriately model the target problem of calculating the survival probability at specific time points after LVAD implantation (namely at 1 month, 2 months, 3 months, 6 months, 9 months and 12 months), the methodology of dynamic Bayesian networks (DBNs) was selected. DBNs [16] can model temporal dependencies and therefore are an appropriate choice of modeling survival after the implantation. They can capture the dynamics of patient disease by taking into account the post-operative data and their evolution during the follow-up period after implantation. DBNs have been used in several domains [17-21]. In order to apply the DBNs, the patient features/ measurements are used as input at each time slice. Through DBN modelling we can identify the most important of them

*This research project has been pat funded by the European Union (Project SensorART: A remote controlled Sensorized ARTificial heart enabling patients empowerment and new therapy approaches. ICT 248763.

T.P. Exarchos, G. Rigas, K. Stefanou are with the Dept. of Biomedical Research, Institute of Molecular Biology and Biotechnology, FORTH University of Ioannina, GR 45110 (email: exarchos@cc.uoi.gr, rigas@cs.uoi.gr, kstefan@cc.uoi.gr)

Y. Goletsis is with the Dept of Economics, University of Ioannina, GR 45110 (email: goletsis@cc.uoi.gr)

S. Jacobs is with Department of Cardiac Surgery, University Hospital Leuven, Leuven, Belgium (email: Steven.Jacobs@med.kuleuven.be)

M.G. Trivella is with the Institute of Clinical Physiology, National Research Council, Pisa, Italy (email: trivella@ifc.cnr.it)

D. I. Fotiadis is with the Unit of Medical Technology and Intelligent Information Systems, Dept of Materials Science and Engineering, University of Ioannina and with the Dept. of Biomedical Research, Institute of Molecular Biology and Biotechnology, FORTH (corresponding author phone: +302651008803; fax: +302651008889; e-mail: fotiadis@cc.uoi.gr).

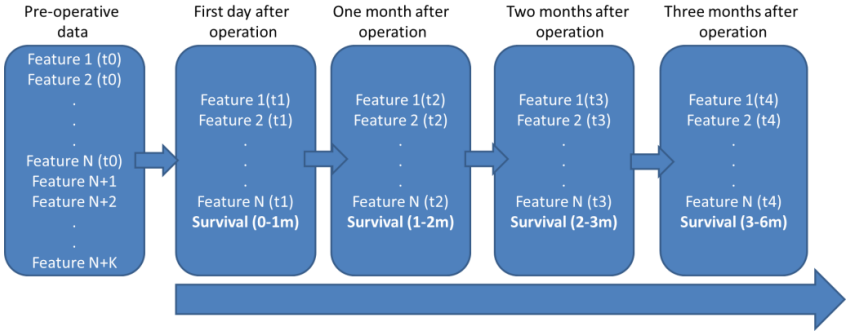

Figure 1: The architecture of the short term Dynamic Bayesian Network used in our application.

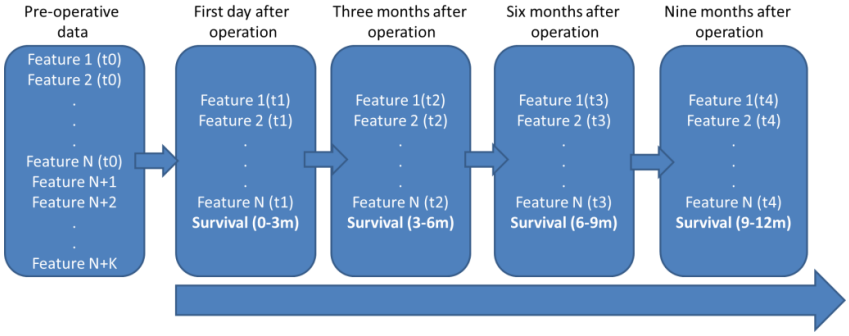

Figure 2: The architecture of the longer term Dynamic Bayesian Network used in our application.

which are related to the survival of the patient. Two different types of DBNs have been developed: (i) a short term one to model the dependencies for the first three months after implantation when patient's situation is less stable (with time points at one day, one month, two months and three months after the implantation) and (ii) and a long term DBN for modeling the dependencies for the period from month 3 to month 12 after the implantation (with time points at months 3, 6, 9 and 12) The concept of the two DBNs is shown in Figs. 1 and 2, respectively. A Bayesian Network (BN) is a directed acyclic graph, where each node is one of the features. For a network described as $B=(G,P)$, where G is a directed acyclic graph, $\mathbf{X} = \{x_1, x_2, \ldots, x_N\}$, is a set of features, and P is the joint probability distribution of features in X, as follows:

$$P(\mathbf{X}) = \prod_{i=1}^{N} P(x_i \mid \pi_G(x_i)), \quad (1)$$

where $\pi_G(x)$ denotes the parents of x in G. A DBN is defined as a pair $DB=(B_0,B_{trans})$, where B_0 is a BN, defining the prior $P(X_0)$ and B_{trans} is a two-slice temporal BN which defines $P(X_t|X_{t-1})$. The semantics of a DBN can be defined using the 2 slice temporal BN, in all time-slices. The resulting joint distribution is given by:

$$P(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_T) = \prod_{t=1}^{T} \prod_{i=1}^{N} P(x_i^t \mid \pi_G(x_i^t)). \quad (2)$$

Our aim is to use the training data in order to define the architecture of the network, which means to identify the dependencies between the features within each time slice (inter-dependencies) as well as across the different time slices (intra-dependencies). Two different algorithms were employed, the Bayesian Search and the PC algorithm [22], for searching across the feature space and identification of the optimal network architecture that provides the highest accuracy. The first task is to identify all dependencies among the features of the network and after that, to provide evidence to the trained network and conjecture about the value each feature in the network in the next post operation phase time slices. Due to the transparent architecture of the DBNs, we are able to identify new information regarding the correlation of the features with the patient survival probability and thus the underlying processes that take place after the VAD implantation. From the set of resulting DBN architectures, we are also able to identify the optimal set of features, both from the pre-operation features, as well as, in each time slice after the implantation/operation.

B. Dataset

The dataset for training and evaluating the DBN methodology has been provided by Katholieke Universiteit of Leuven, Belgium. The dataset contains 71 patients, with 41 pre-operative features and 9 features in every post-operation time slice. Patients were followed for an up to 12 month period after implantation unless they died. Patients that had a heart transplant, or at their latest measurement were alive or had a VAD explantation were considered as patients that survived at the specific time slice of the measurements of one of the above mentioned events. Out of the 71 patients, 53 died within the 12 month period. The 41 pre-operative features are shown in Table I, with their mean (or median) and their standard deviation. These features take into consideration medical literature on pre-operative risk factors for VAD patients (see e.g [23]). The post-operative features, are shown in Table II.

III. Results

In order to evaluate our methodology, due to the limited sample size, 10-fold cross validation method was employed. In addition, the evaluation is also extended to predict the

TABLE I: PRE-OPERATIVE FEATURES

Feature	Units	mean ± std
Age (yrs)	years	48.07 ± 14.82
Gender	-	14F, 57 M
BSA body surface area	m^2	1.89±0.21
BMI body mass index	Kg/ m^2	24.53±3.52
In Hospital Prior to Implant?	Inpatient/ outpatient	11 Oupatinet 60 inpatient
HF heart failure Etiology	Ischemic/ dilated cardiomyopathy	35 Isch 36 DCM
Inotropes	Yes/no	51 yes, 16 No
PCWP pulmonary capillary wedge pressure	mmHg	28.11±7.29
PAP (S) pulmonary artery pressure (systolic)	mmHg	49.66±13.59
PAP (D) pulmonary artery pressure (diastolic)	mmHg	29.32±7.22
PAP (M) pulmonary artery pressure (mean)	mmHg	36.04±8.71
RVP right ventricular pressure	mmHg	50.85±13.22
RAP right atrial pressure	mmHg	16.47±7.32
MAP mean arterial pressure	mmHg	71.16±11.10
CO cardiac output	L/min	3.51±1.12
CI cardiac index	L/min/ m^2	1.82±0.52
SVR systemic vascular resistance	dyn·s/cm^5	1389.92±463. 19
PVR pulmonary vascular resistance	dyn·s/cm^5	2.95±1.72
CP Cardiac Power	W	0.56±0.21
CPI Cardiac Power Index	W/ m^2	0.29±0.10
HR heart rate	Beats per minute	87.81±21.18
BP (S) blood pressure (systolic)	mmHg	92.10±18.98
BP (D) blood pressure (diastolic)	mmHg	59.98±10.46
LVEF left ventricular ejection fraction	%	17.93±7.00
LVEDD left ventricular end-diastolic diameter	mm	62.12±8.38
BNP NT-proBNP level	ng/L	74.35±36.64
Hb Hemoglobin	g/dL	11.23±2.54
WBC white blood cell count	10^9 /L	9.84±4.79
PC Platelet Count	10^9 /L	193.10±89.15
AST	IU/L	240.2±570.7
ALT	IU/L	202.8±461. 4
LDH	IU/L	1130.92±142 8.62
T. Bilirubin Total bilirubin level	mg/L	1.63±1.48
Na Sodium	mmol/L	135.06±8.30
BUN blood urea nitrogen	mg/dL	76.51±42.59
Creatinine	mg/L	1.52±0.86
Creatinine Clearance	ml/min	62.08±32.21
CRP C-reactive protein	mg/L	74.54±97.31
POD Transfer to Ward post operative day	days	19.80±17.00
POD Discharge (Days)	days	35.96±26.27
INTERMACS Profile		13 (1), 23 (2), 20 (3), 13 (4), 2 (5)

TABLE II: POST-OPERATIVE FEATURES

Feature	Units
NT-proBNP	Ng/L
C-reactive protein	g/dL
Leukocytes	10^9 /L
Creatinine	mg/dL
Urea	mg/dL
Sodium	mmol/L
Bilirubine total	mg/L
AST	IU/L
ALT	IU/L

next time slice condition, based on the previous conditions (either *t-1*, or *t-1*+*t-2* etc until *t-1*+*t-2*+...+*t-n*). In this case we provide evidence to the model for the respective time slice(s) and evaluate the model in the prediction of the next time slices.

Below we present the DBN developed for the survival after VAD implantation problem. The short and long term DBN (their structure is the same) that take into consideration also the pre-operative data is shown in Fig. 3. Although the two DBNs share the same structure, estimated parameters (probabilities) differ in each case. The corresponding accuracy results for the short and long term DBNs are shown in the Tables III and IV. Accuracy is computed as number of correctly identified events (predictions) divided to the total number of predictions. Having available data for previous time slices the table presents the prediction accuracy for the following time slice (in terms of survival).

TABLE III: ACCURACY RESULTS FOR THE SHORT TERM DBN

	1T	2T	3T	4T
Data 0 T	74.65%	67.65%	68.18%	50.77%
Data 1 T	-	91.18%	89.39%	69.23%
Data 2 T	-	-	96.97%	70.77%
Data 3 T	-	-	-	70.77%

Where 0T denotes pre-op data, 1T denotes the time slice in the first day after the operation, the 2T the time slice after 30 days of the operation, the 3T the time slice after 60 days of the operation and the 4T the time slice after 90 days of the operation. Each result presents the respective accuracy in predicting the next time slice condition, based on the previous ones.

TABLE IV: ACCURACY RESULTS FOR THE LONG TERM DBN

	1T	2T	3T	4T
Data 0 T	74.65%	64.71%	34.78%	93.55%
Data 1 T	-	72.55%	56.52%	93.55%
Data 2 T	-	-	78.26%	93.55%
Data 3 T	-	-	-	74.19%

In Table IV, 0T denotes pre-op data, 1T denotes the time slice in the first day after the operation, the 2T the time slice after 3 months of the operation, the 3T the time slice after 6 months of the operation and the 4T the time slice after 9 months of the operation.

IV. DISCUSSION

We have presented a novel approach for the prediction of survival probability of patients at specific time points after ventricular assist device implantation. Our approach is based on Dynamic Bayesian Networks. DBNs model the

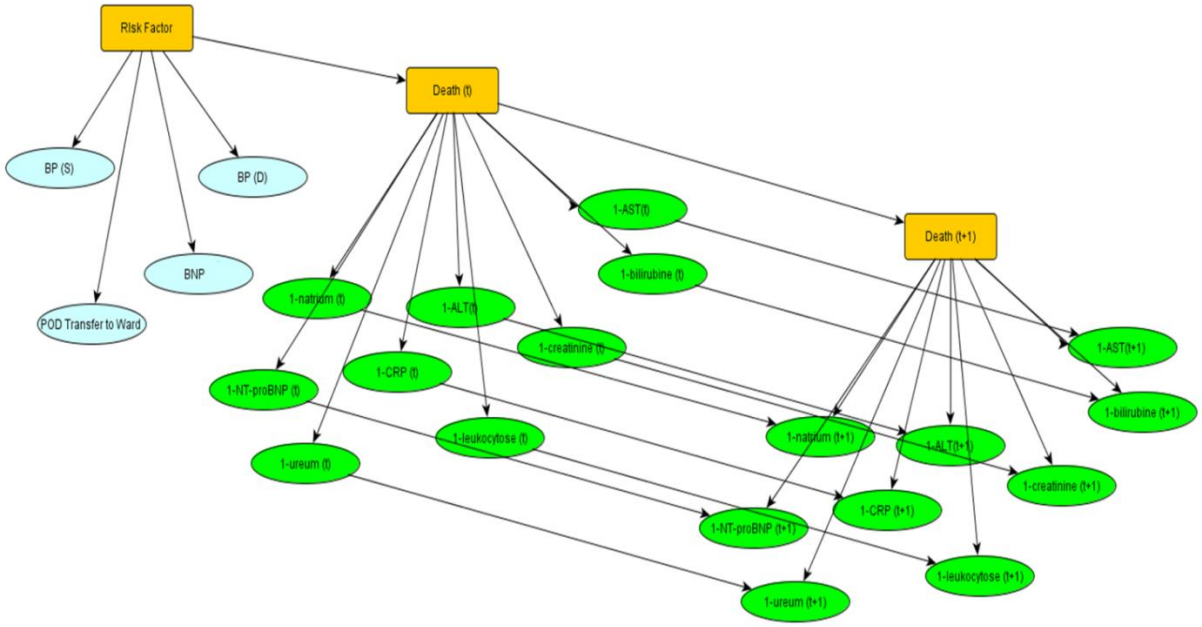

Figure 3: The Dynamic Bayesian Networks developed for the short and long term survival prediction

dependencies between the data and provide an insight to the survival process. The accuracy results shown in Tables III and IV present some differences in accuracy, in relation to the available input. When more information is used as input the accuracy for the next time slice is improved.

Future work will focus on two different directions. The first one refers to the extension of the available dataset, so as to train and test the developed DBNs with higher number of patient cases. The second direction will focus extended discussion with cardiologists and cardiosurgeons so as to incorporate the developed models in the clinical practice. Moreover, after having confirmed the proof of concept of this approach, the next step is to perform a large scale study so as to provide results comparing the efficacy of our approach to existing prognostic scores such as HFSS and other time dependent methodologies, i.e. Cox regression.

REFERENCES

[1] M.S. Slaughter et al., "Clinical management of continuous-flow left ventricular assist devices in advanced heart failure", *J. Heart Lung Transplant*, vol. 29, pp S1-S39, 2010.

[2] K.D. Aaronson et al., "Development and prospective validation of a clinical index to predict survival in ambulatory patients referred for cardiac transplant evaluation", *Circulation*, vol. 95, pp. 2660-7, 1997.

[3] W.C. Levy et al., "The Seattle heart failure model: prediction of survival in heart failure", *Circulation*, vol. 113, pp. 1424-33, 2006.

[4] W.C. Levy et al., "Can the Seattle heart failure model be used to risk-stratify heart failure patients for potential left ventricular assist device therapy?", *J Heart Lung Transplant,* vol. 28, pp.231-6, 2009.

[5] J.K. Kirklin et al., "Second INTERMACS annual report: More than 1,000 primary left ventricular assist device implants", *J Heart Lung Transplant*, vol. 29, pp. 1-10, 2010.

[6] W.L. Holman et al., "INTERMACS: interval analysis of registry data", *J Am Coll Surg*, vol. 208, pp. 755-61, 2009.

[7] L.W. Stevenson et al., "INTERMACS profiles of advanced heart failure: the current picture", *J Heart Lung Transplant* , vol. 28, pp. 535-41, 2009.

[8] J.C. Matthews et al., "Model for End-Stage Liver Disease Score Predicts Left Ventricular Assist Device Operative Transfusion Requirements, Morbidity, and Mortality", *Circulation*, vol. 121, pp. 214-220, 2010.

[9] D.J. Farrar, "Preoperative predictors of survival in patients with Thoratec ventricular assist devices as a bridge to heart transplantation. Thoratec Ventricular Assist Device Principal Investigators" *J Heart Lung Transplant,* vol. 13, pp. 93-100, 1994.

[10] M.C. Oz et al., "Screening scale predicts patients successfully receiving long-term implantable left ventricular assist devices" *Circulation*, vol. 92, pp. 169-73, 1995.

[11] V. Rao et al., "Revised screening scale to predict survival after insertion of a left ventricular assist device" *J Thorac Cardiovasc Surg*, vol. 125, pp. 855-62, 2003.

[12] C.Y. Lin et al., "Evaluation of outcome scoring systems for patients on extracorporeal membrane oxygenation" *Ann Thorac Surg*, vol. 84, pp. 1256-62, 2007.

[13] K. Lietz et al., "Outcomes of left ventricular assist device implantation as destination therapy in the post-REMATCH era: implications for patient selection" *Circulation,* pp. 116, pp. 497-505, 2007.

[14] Y. Wang et al., "A Classification Approach for Risk Prognosis of Patients on Mechanical Ventricular Assistance", *Proc Int Conf Mach Learn Appl.12 December*, pp. 293–298, 2010.

[15] E. Karvounis et al., "An Intelligent Decision Support System for the treatment of patients receiving Ventricular Assist Device Support", *Methods of Information in Medicine*, 2014 to appear.

[16] Z.R Xiang, et al., "miniTUBA: medical inference by network integration of temporal data using Bayesian analysis." *Bioinformatics*, vol. 23, pp. 2423-2432, 2007.

[17] M.A. van Gerven, et al., "Dynamic Bayesian networks as prognostic models for clinical patient management." *J Biomed Inform*, vol. 41, pp. 515-529, 2008.

[18] A.H. Marshall, et al., "Continuous Dynamic Bayesian networks for predicting survival of ischaemic heart disease patients", IEEE 23rd International Symposium on Computer-Based Medical Systems (CBMS), pp. 178-183, 2010.

[19] R. Chen, S.M. Resnick, C. Davatzikos, E.H. Herskovits, "Dynamic Bayesian network modeling for longitudinal brain morphometry." Neuroimage. vol. 59, pp. 2330-8, 2012

[20] L.N. Peelen, et al., "Using hierarchical dynamic Bayesian networks to investigate dynamics of organ failure in patients in the Intensive Care Unit." *J Biomed Inform*, vol. 43, pp. 273-286, 2010.

[21] K.P. Exarchos, Y. Goletsis, D.I. Fotiadis, "A multiscale and multiparametric approach for modelling the progression of oral cancer", *BMC Med. Inform. and Decision Making,* vol. 12, 136, 2012.

[22] P. Spirtes, C. Glymour, R. Sheines, *Causation, Prediction and Search*,. Cambridge, Massachusetts. U.S.A.: MIT Press, 2000.

[23] A.J. Boyle at al., "Preoperative Risk Factors of Bleeding and Stroke during LVAD Support: An Analysis of Over 900 HeartMate II Outpatients", *J. Am. Col. of Card.*, vol. 13, pp. S0735-1097, 2013.