Using Three-Dimensional Cardiac Motion for Predicting Mortality in Pulmonary Hypertension: A Deep Learning Approach

Ghalib A. Bello
MRC London Institute of Medical Sciences
Imperial College London, UK
gbello@ic.ac.uk

Timothy J.W. Dawes
National Heart & Lung Institute
Imperial College London, UK

Jinming Duan
MRC London Institute of Medical Sciences
Imperial College London, UK

Declan P. O'Regan
MRC London Institute of Medical Sciences
Imperial College London, UK

Abstract

Pulmonary hypertension (PH) is a life-threatening and rapidly progressive disease in which functional adaptation of the right ventricle (RV), as quantified by RV ejection fraction (RVEF), is a key prognostic marker. However, RVEF is largely insensitive to regional or early RV dysfunction which may improve prognostication and allow prompt identification of high-risk cases. Cardiac Magnetic Resonance (MR) imaging is a standard modality for quantification of RV function, and can be used to derive anatomically accurate, high-resolution 3D shape models of RV contraction using recently developed computational imaging analysis techniques. These time-resolved 3D models may lend additional insights beyond what is offered by simple conventional measures like RVEF. In this study, we train a neural network to predict mortality in PH patients by learning complex RV contraction patterns from 3D shape models of RV motion. To handle right-censored survival time outcomes, our network utilized a Cox partial likelihood loss function. The network was trained on imaging and mortality data on 148 PH patients. It yielded improved prediction accuracy and superior risk stratification, compared with a multivariable survival model consisting of RVEF and other conventional parameters of RV function.

1 Background

PH often follows a rapidly progressive course with impaired exercise tolerance associated with RV hypertrophy, dilatation and ultimately heart failure. The transition to heart failure in PH is characterised by evolving morphological and functional adaptations of the heart, and yet why some patients rapidly deteriorate while others do not remains unresolved. RVEF, a key prognostic marker of heart failure and mortality in PH, is a simple index representing the volume of blood ejected as a fraction of the chamber's relaxed volume. However, RVEF and other simple volumetric measures of function are insensitive to the complex morphological and contractile patterns of RV dysfunction that characterise progressive impairment. Our group has developed methods to construct time-resolved 3D shape models of the heart which provide a detailed and dynamic representation of the beating heart. Encoded in these representations would be subtle morphological and functional patterns associated with RV dysfunction and thus mortality risk. In this study, we leverage the power of deep learning to develop a prognostic survival model exploiting this rich spatio-temporal information.

1st Conference on Medical Imaging with Deep Learning (MIDL 2018), Amsterdam, The Netherlands.

2 Materials and Methods

2.1 MR Image Processing and Computational Image Analysis

The cine MR imaging protocol has been described in detail elsewhere [1]. After acquisition, the raw cine MR images were processed through a co-registration and segmentation pipeline, producing anatomically smooth 3D shape models of the heart at 20 time frames across the cardiac cycle. These were then used to construct high-resolution 3D mesh representations of heart motion. These highly detailed surface mesh models were constructed from $18,028$ vertices (see Figure 1A).
Additionally, we performed volumetric analysis on the segmented images to derive the following conventional, volumetric indices of cardiac function: RVEF, Right Ventricle End Diastolic Volume (RVEDV) and Right Ventricle End Systolic Volume (RVESV).

2.2 Characterization of Right Ventricular motion

The time-resolved 3D meshes described in the previous section were used to produce a simple representation of RV motion. For this purpose, we utilized a sparser version of the meshes (down-sampled by a factor of 10) with 1803 vertices. To characterize RV motion, we adapted an approach introduced in [2]. Let (x_{vt}, y_{vt}, z_{vt}) represent the Cartesian coordinates of vertex v ($v = 1, .., 1803$) at the t^{th} time frame ($t = 1, .., 20$) of the cardiac cycle. We describe the dynamics and morphology of RV contraction by tracking the motion of each vertex across all 20 time frames (using frame 1 as baseline). Specifically, at each time frame $t = 2, 3, ..., 20$, we compute the coordinate-wise displacement of each vertex from its position at time frame 1. This yields the following one-dimensional input vector:

$$\boldsymbol{X} = \begin{pmatrix} x_{vt} - x_{v1}, & y_{vt} - y_{v1}, & z_{vt} - z_{v1} \end{pmatrix}_{2 \leq t \leq 20}^{1 \leq v \leq 1803}$$

Vector $\boldsymbol{X}$ has length 102,771 ($3 \times 19 \times 1803$), we used it as the input feature for our network.

2.3 Network architecture and training parameters

Our network utilized a simple multilayer feedforward architecture. The (102771×1) feature vector representing vertex-wise displacement was used as the input layer for the network, followed by two hidden, fully connected layers (consisting of 256 and 128 units respectively). The final/output layer had a single unit which represents predicted mortality risk. The outcome in this study was survival/follow-up time, i.e. the time elapsed from MRI acquisition until death (all-cause mortality) or for those still alive, the last date of follow-up. This type of outcome is called a *right-censored time-to-event* outcome, and is typically handled using survival analysis techniques, the most popular of which is Cox's proportional hazards regression model [4]. We used as our network loss function the negative Cox proportional hazards partial likelihood (given below), thereby extending the standard Cox regression model to accommodate a neural network architecture:

$$\mathcal{L}(\xi, \boldsymbol{X}) = -\sum_{i=1}^{n} \delta_i \left\{ F(\boldsymbol{X_i}, \xi) - log \sum_{j \in R(t_i)} e^{F(\boldsymbol{X_j}, \xi)} \right\}$$

where $\boldsymbol{X_i}$ is the input vector for subject i, ξ represents the neural network parameters to be trained, and $F(\boldsymbol{X_i}, \xi)$ is the network's output/prediction (mortality risk) for subject i. δ_i is an indicator of subject i's status (0=Alive, 1=Dead) and $R(t_i)$ represents subject i's risk set, i.e. subjects still alive (i.e. at risk) at the time subject i died or became censored (t_i). This loss function was minimized (with respect to network parameters) via backpropagation. To avoid overfitting, we applied l_1/l_2 regularization, dropout, and the max-norm constraint. The rectified linear unit (ReLU) activation function was used for the hidden layers, and linear activation for the output layer. Batch normalization was applied after each hidden layer. Using the Adaptive Moment Estimation (*Adam*) algorithm, the network was trained for 100 epochs with a learning rate of 10^{-4}.

2.4 Conventional Parameter model

As a benchmark comparison to our RV motion model, we trained a Cox proportional hazards model utilizing conventional volumetric indices of RV function (RVEF, RVEDV and RVESV) as mortality

predictors. To account for collinearity among these predictor variables, an l_2-norm regularization term was added to the Cox partial likelihood function.

2.5 Model validation

Prediction accuracies of the RV motion model and the conventional parameter model were evaluated using Harrell's Concordance Index [3], an extension of AUC to survival outcomes. This index quantifies the extent of agreement between predictions and censored time-to-event outcomes, and has the same range and interpretation as AUC. We estimated (and report below) the mean concordance index across three-fold cross-validation analyses.

3 Results

3.1 Cohort Characteristics and Predictive Performance

Imaging and all-cause mortality data were collected from a sample of 148 patients diagnosed with chronic thromboembolic PH, who had not undergone pulmonary endarterectomy by the end of the follow-up period. The patients underwent routine diagnostic assessment and cardiac MR imaging from 2004-2017, and were followed up until November 2017. By end of follow-up, 45% had died. The RV motion model yielded a higher concordance index (0.67) than the conventional parameter model (0.58), the improvement being statistically significant (p-value $= 0.01$). Additionally, the RV motion model demonstrated greater risk differentiation, as shown in Figure 1B.

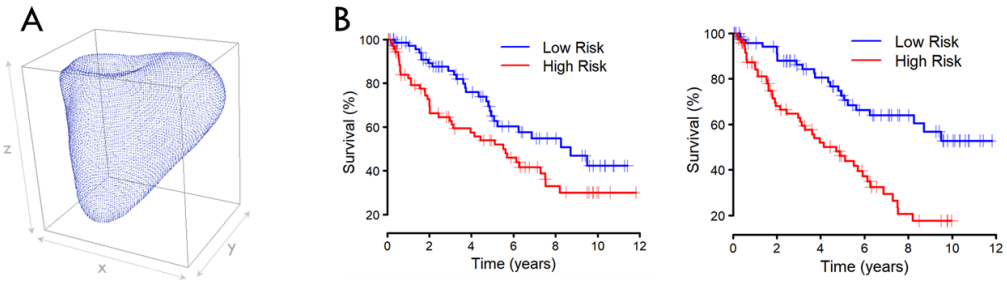

Figure 1: **A**: Exemplar RV 3D mesh model with 18028 vertices, **B**: Kaplan-Meier curves showing risk stratification quality of RV motion network model (right) vs. conventional parameter model (left). For each risk model, patients were divided into low- and high-risk groups according to the risk score medians. The curves compare survival rates over time between risk groups

4 Conclusion and Future Work

This study demonstrates the utility of deep learning for identification of prognostic spatio-temporal patterns in 3D models of RV motion and the results suggest the potential of more specialized/spatially-aware network architectures to achieve further improvement in prognostic accuracy.

5 References

[1] Dawes, T.J.W. et al (2016) Machine Learning of Three-dimensional Right Ventricular Motion Enables Outcome Prediction in Pulmonary Hypertension: A Cardiac MR Imaging Study. *Radiology* **283**(2):381.

[2] Bai, W. et al (2015) A Bi-ventricular Cardiac Atlas Built from 1000+ High-Resolution MR Images of Healthy Subjects and an Analysis of Shape and Motion. *Med Image Anal*, **26**(1):133.

[3] Harrell, F.E. et al (1982) Evaluating the yield of medical tests. *JAMA* **247**(18):2543.

[4] Cox, D.R. (1972) Regression Models and Life-Tables. *J Roy Stat Soc B* **34**(2):187.