Proceedings Paper:
Mbotwa, JL, de Kamps, M orcid.org/0000-0001-7162-4425, Baxter, PD orcid.org/0000-0003-2699-3103 et al. (2 more authors) (2018) Latent Class Regression Modelling: A novel approach to predict survival of patients with Chronic Heart Failure. In: Journal of Epidemiology and Community Health. Society for Social Medicine 62nd Annual Scientific Meeting, 05-07 Sep 2018 BMJ Publishing Group , A67-A68.

https://doi.org/10.1136/jech-2018-SSMabstracts.141

Latent Class Regression Modelling: A novel approach to predict survival of patients with Chronic Heart Failure

JL Mbotwa [1 3],M de Kamps[2], PD Baxter[1] and MS Gilthorpe[1]

[1]Leeds Institute for Data Analytics,Level 11, Worsley Building, Clarendon Way, Leeds, LS2 9NL, UK

[2]Institute of Artificial and Biological Intelligence, School of Computing, University of Leeds,LS2 9JT, Leeds, UK

[3]Department of Applied Studies, Malawi University of Science and Technology, Malawi.

Aim

To compare traditional regression approach with the latent class regression modelling for prediction.

Introduction

Chronic Heart Failure (CHF) is one of the leading cause of hospitalizations and deaths, more especially in old people, and this causes a substantial clinical and economic burden to the government. Using risk prediction models to accurately understand the dynamics of survival patterns amongst patients with CHF condition would provide guidance to health care professionals in decision making on how to improve the delivery of care. However, prediction models used in medical research often fail to accurately predict health outcomes due to methodological limitations. These models particularly perform poorly when predicting narrowly targeted subgroups of patients. We explore the role of latent class regression (LCR) analysis to model the survival of patients with CHF. We seek to show that using LCR improves the modelling of health outcomes as it accounts for unobserved heterogeneity that exists naturally within the patient data.

Methods

LCR generally involves identifying hidden latent classes within data and uses patient's demographic characteristics and other covariates to predict class membership and separate regression models for each class. These latent classes may correspond to subgroups of patients with specific characteristics that affect their survival. The rationale is that one class will be more susceptible to deaths compared to another. The United Kingdom Heart Failure Evaluation and Assessment of Risk Trial (UK-HEART) recruited patients with signs and symptoms of CHF between July 2006 and December 2014. A total of 1802 records were available on patient characteristics as well as medications. We modelled the survival of patients using a Cox proportional hazard model within a latent class framework by estimating a single model across the latent classes. We increased complexity of our model by allowing each class to have a separate survival model.

Latent Class Model

The Latent Class Model (LCM) assumes data is drawn from a family of finite mixture models. The basic LCM has the form:

$$f(\mathbf{y}|\mathbf{x},\mathbf{z},\boldsymbol{\theta}) = \sum_{i=1}^{g} \pi_i(\mathbf{z}|\gamma_i,\boldsymbol{\delta}_i) f_i(\mathbf{y}|\mathbf{x},\boldsymbol{\beta}_i) \quad (1)$$

where $\pi_i(\mathbf{z}|\gamma_i,\boldsymbol{\delta}_i)$'s are class-membership probabilities that are estimated for each class and are dependent on a vector of covariates $\mathbf{z}$ and a vector of coefficients $\boldsymbol{\delta}_i$ such that

$$\sum_{i=1}^{g} \pi_i(\mathbf{z}|\boldsymbol{\gamma}_i,\boldsymbol{\delta}_i) = 1 \quad (2)$$

with $0 \leq \pi_i(\mathbf{z}|\gamma_i,\boldsymbol{\delta}_i) \leq 1$ and $f_i(\mathbf{y}|\mathbf{x},\boldsymbol{\beta}_i)$ is the conditional probability density function for the observed response in the ith class model , $\mathbf{x}$ is the covariate vector.

UNIVERSITY OF LEEDS

Results

We used the area under the receiver operating characteristic (ROC) curve to assess the predictive performance these models. Overall, our novel approach performed better than the traditional one class model approach. Our model gave an area under the ROC curve of 0.84 while the traditional model yielded an AUC of 0.68.

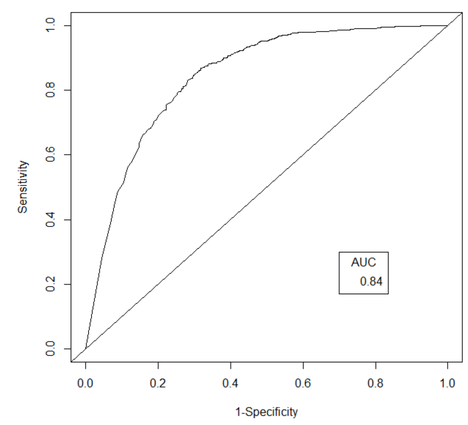

Figure 1: A ROC curve assessing the predictive performance for a Two-Class Latent Class Model

Key finding

Accounting for population heterogeneity in prediction models improves individual predictions.

Conclusion

Ignoring the natural heterogeneity that exists within the patient data affects the accuracy of the estimates in prediction models. Researchers can utilize the available data in order to identify the hidden latent classes within the data. Fitting a regression model to each of the latent classes improves the predictive accuracy of the model.

References

Vermunt JK, and Magidson J. *Latent class analysis.*, 2004;175-198.
Witte KK, Patel PA, Walker AMN, et al. *Socio economic deprivation and mode-specific outcomes in patients with chronic heart failure.* 2018;1-6.

Acknowledgements

The author(s) would like to thank Richard Cubbon for providing data. We would also like to thank all participants who took part in this trial.

Contact Information

- John Lenard Mbotwa, Leeds Institute for Data Analytics,Level 11, Worsley Building, Clarendon Way, Leeds, LS2 9NL, UK
- mt13jlm@leeds.ac.uk

LEEDS Institute for Data Analytics