Master Thesis Econometrics and Management Science

Business Analytics and Quantitative Marketing

June 21, 2021

Bayesian multilevel modelling of patient data in South Australia and the Fully Bayesian approach

---

Martijn Hofstra (387276)

Supervisor EUR: P.H.B.F. Franses

Co-reader EUR: Martina Zaharieva

Supervisors SAHMRI: Lukah Dykes & Johan Verjans

1 Introduction

1.1 Problem description

ED overcrowding/'Access block' means that patients who require emergency hospital admission spend an unreasonable amount of time in an emergency department (ED) because they cannot gain access to appropriate hospital beds (Richardson, 2002). Access block is regarded as the major issue currently facing emergency medicine in Australia (Richardson et al., 2009) (Geelhoed and de Klerk, 2012) (Cameron et al., 2009). It is linked with longer hospital stay afterwards. Patients that remain in the ED for 8-12 hours are 20% more likely to have an excess stay (i.e. longer than the state average for the relevant admission problem) in the hospital afterwards. For patients that remain in the ED for more than 12 hours this can be as high as 50%. Research in Australia has shown than increased ED occupancy is associated with significantly higher short-term patient mortality (Liew et al., 2003).

There was a 6.8% total growth in ED presentations per 1000 persons in South Australia in the period 2000-2001 to 2009-2010. The annual growth in this period was 0.8% (FitzGerald et al., 2012). In a selection of major metropolitan hospitals[1] the total number of ED presentations increased from 375,703 in 2012/2013 to 412,935 in 2017/2018, corresponding to an increase of approximately 6% when adjusting for the change in population (SAH, 2019). In 2017-2018, 61% of the patients admitted to the ED in South Australia were seen on time (i.e. < 4 hours). In 2014/2015, the return rate to the ED in South Australia was about 0.89% (AIH, 2018). These number indicate that the problem of 'access block' is becoming bigger.

In Australia there are almost 1 million ED presentations every year of a suspected cardiac chest patient. In South Australia this represents approximately 30,000 presentations per year. However, about 85% of these patients do not actually have an acute coronary syndrome. In 2015, the mean cost per chest pain patient in admitted to the ED in Australia was $5,272 (Cullen et al., 2015). In the Australian Capital Territory 16.2% of the ED patients had cardiac complaints (Richardson, 2002) and in New South Wales this was even 22.7% (Chan et al., 2008). This strongly indicates that a relatively high proportion of ED patients in Australia has cardiac complaints and that chest pain presentations are a major cause of access block. Hence, it can be beneficial to target chest pain patient in order to reduce ED waiting times and to improve patient flow.

In metropolitan South Australia (in 2011-2012) refugees (Refugee and Asylum Seeker Countries (RASC)), aboriginals and those aged 75 and older had the highest risk of presenting to the emergency department (ED). Moreover, aboriginals and those aged 75 and older had the highest risk of multiple ED presentations (re-admission). The excess costs of ED presentations associated with vulnerable

[1]The Queen Elizabeth Hospital, Royal Adelaide Hospital, Lyell McEwin Hospital, Modbury Hospital, Women's and Children's Hospital – Paediatrics, Flinders Medical Centre and Noarlunga Hospital

groups (RASC, aboriginals and patients aged 75 years and older) were $A 22 million per year (Banham et al., 2019). Access block is more likely to lead to longer hospital stays for people older than 65 years, females and patients presented to the ED outside working hours (Liew et al., 2003). Higher ED occupancy was associated with a higher proportion of elderly, female and was more likely during weekdays and during winter (Sprivulis et al., 2006).

1.2 Social relevance

In order to explain the social relevance of this thesis, it is important to note that this thesis is written in collaboration and in assignment of SAHMRI in South Australia, specifically, in collaboration with the researchers in the RAPIDx AI project. Hence, the context of the application of this thesis is the RAPIDx AI project. The goal of the RAPIDx AI project is to assist with the medical management of patients presenting to the emergency department with potential acute coronary syndrome (ACS). This will be done by integrating clinical care with validated real-time data and modern analytical methods to better support clinical decision-making and help establish the South Australian health system as an effective learning health system. The RAPIDx AI project will deploy an AI-based diagnostic algorithm for identifying patients with potential Type I or Type II myocardial infarction (MI)/myocardial injury within the emergency departments (EDs) of six South Australian hospitals, and will provide protocolised recommendations for medical management of these patients. By supporting cardiologists in correctly identifying Type I MI patients, the RAPIDx AI will hopefully improve patient flow and reduce access block. The HeartAI system provides the digital platform to enable real-time data and analytical methods. In a supporting partnership with the RAPIDx AI project, Siemens will deploy the AI-Pathway Companion to provide a robust interface at the clinical point-of-care (Dykes, 2021). Hopefully, developing Bayesian causal models will support the RAPIDx AI predictive models.

An important component in detecting a MI is the troponin level. Recent advancements in troponin testing have led to an improvement in detection of myocardial injuries. However, this has not necessarily made decision-making for clinicians in the ED easier. Namely, chest pain patients with a T2MI cardiac outcome are falsely labelled as T1MI patients, leading to unnecessary admittance to the ED (Dykes, 2021). Troponin are structural proteins unique to the heart. For the diagnosis of a myocardial infarction, the troponin levels are superior to all other available clinical biomarkers (e.g. Creatine-kinase or the white blood cell count) (Reichlin et al., 2009). Troponin is known to be highly correlated with the chance of having a myocardial infection (Fathil et al., 2015).

Correct identification of Type I MI's by predictive models can be very beneficial in supporting the cardiologists' decision in whether or not to discharge a cardiac patient presented to the ED. Discharging a cardiac patient with a Type I MI can have severe consequences. However, unnecessary admittance of cardiac patients to the ED leads to excess costs and increase of access block. In gen-

eral, cardiologists with less field experience (i.e. junior cardiologists) tend to be over-conservative in their clinical decision-making and thus be hesitant to discharge patients. Whereas predictive models are generally focused on attaining the highest prediction accuracy, causal models are focused on describing causal structures. In order to support the RAPIDx AI predictive models, this paper attempts to describe causalities of data of cardiac patients admitted to the ED. These causal structures can be useful for the RAPIDx AI predictive models because they can be used for the likelihood specification.

In the Australian Capital Territory (ACT) 16.2% of the ED patients had cardiac complaints (Richardson, 2002) and in New South Wales (NSW) this was even 22.7% (Chan et al., 2008). As to my knowledge, no exact numbers exist for the proportion of cardiac complaints at the ED. However, a report by AIH (2018) contains data on ED admissions specified by ICD-10-AM category. The categories 'Diseases of the respiratory system' (J00–J99) and 'Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism' (D50–D89) consist of 8.36% of the ED presentation in Australia in 2017/2018. In comparison, this is 8.63% and 8.07% for NSW and ACT, respectively (AIH, 2018). This strongly indicates that a relatively high proportion of ED patients in South Australia has cardiac complaints. ED presentations and overcrowding are associated with higher mortality, longer hospital stays afterwards and excess costs. Patient outcomes like service usage, representation/readmission to the ED, length of stay and costs related to the ED presentation differ greatly for patients.

Since ED presentation, a higher ED occupancy and the consequences of a higher ED occupancy are associated with certain patients characteristics, it can be valuable to build a statistical (i.e. a Bayesian causal) model, for the patient outcomes of cardiac patients to the emergency department, for the following two reasons:

- The found described causalities can be used to support the RAPIDx AI predictive models. Models predicting using cardiac patient data often have a bias because an observation bias exists. In the context of the RAPIDx AI project, the following is an example of an observation bias: Patients receive multiple troponin tests when the first test result is abnormal and/or an overall abnormal cardiac profile. Hence, the values (and missingness) of later troponin measures are dependent on other variables in the data. In conclusion, it can be of great added value for the RAPIDx AI predictive models to gain more insights into the actual causal structures.

- Clinical data in general and data on cardiac patients/heart attacks, in particular, is very complex of nature and thus it is easy to build prediction models based on false likelihood specifications. Particularly, this is problematic in case the data contains many missing observations, as is the case with patient data. Namely, missingness of data leads to worse performance of prediction models, especially when the proportion of missingness increases (Gill et al., 2007). Hence, it can be of great added value to study a model that solely focuses on describing the

causal structures of the complex data sets, in order to support the RAPIDx AI predictive models.

Two examples of the practical applications of the results of this research are:

- Cardiologists with less field experience (i.e. junior cardiologists) tend to be over-conservative in their clinical decision-making and thus be hesitant to discharge patients. Consequently, these patients are admitted to the hospital/ED at costs of 3000$ per day instead of being discharged. Apart from the costs, this also leads to more access block. The found causalities in this research can improve the predictive models of the RAPIDx AI project. Consequently, these prediction models can improve the ability to identify cardiac patients to discharge because the clinicians will be supported in their decision making, based on the predictions.
- An ED clinician (i.e. cardiologist) can admit a patient to the ED whereas the patient should be discharged, due to incorrect identification of the cardiac patients' clinical outcome. Apart from the extra direct costs, this can be harmful for the cardiac patient as well. Namely, iatrogenic complications (i.e. complications caused by medical treatment) may arise, such as adverse drug events, bed sickness and acquired infections. Apart from these complications leading to higher costs, they are also associated with higher mortality (Laskou et al., 2006). The predictive models by RAPIDx AI can support the decision making by the ED clinicians and, as previously mentioned, the found causalities in this paper can support the predictive models by RAPIDx AI.

Apart from studying the causal structures, a solution for the problem of missing data is explored. Missing data in patients' datasets is a very common problem, for example due to patients dying or not all tests being done on every patient (Altman and Bland, 2007). Hence, both SAHMRI and other medical research institutes can profit from a proper solution of the problem of missing data. As can be seen in the methodology (section 4) the fully Bayesian approach as solution for the problem of missing data is explored. To my knowledge, relatively few researchers have used the fully Bayesian approach in the medical world of research. In general, there is few research on applications of the fully Bayesian approach as solution for missingness of data. A possible explanation is the fact that few researchers are familiar with Bayesian inference. As previously mentioned, missingness in patient data is a very common problem and thus exploration of this approach can lead to more insights. According to Ibrahim et al. (2012), the fully Bayesian approach can be perceived as the most powerful and general approach to solving the problem of missing data. Namely, a widely used method, of simply omitting observations with missing values, leads to a decrease in power and an increase in standard errors (Dong and Peng, 2013). Besides, Bayesian inference after multiple imputation (another widely used solution for the problem of missing data) generally leads to unreliable results. Namely, the results are solely reliable in case the posterior

distribution is Gaussian (Zhou and Reiter, 2010). Contrarily, the fully Bayesian approach is feasible when the posterior distribution is not Gaussian (Ibrahim et al., 2012).

1.3 Research questions

In this subsection the research questions are presented. They are formulated as follows:

- Which causal structures are found by modelling patient characteristics on patient outcomes of chest pain patients in the ED?
- How effective is the fully Bayesian approach in solving the problem of missingness in patient data?
- How is the Bayesian workflow effectively used in modelling patient data?

In this research, the following patient outcome is considered:

- Classification of cardiac outcomes. See table 1 for the categories and the meaning of the categories.

The remainder of the thesis is structured as follows: Firstly, the literature related to this topic is presented. Secondly, the data/data structure is described. Thirdly, the proposed methods for answering the research questions are presented. Fourthly, the results of the models are presented in the results section. Consequently, the results are discussed. At last, a conclusion is given, the limitations of the research are discussed and recommendations for further research are presented.

2 Related Literature

Similar research has previously been conducted by Kim et al. (2014). They have used data from the emergency department (ED) patients at Flinders Medical Centre between January 2010 and March 2012. By using logistic regression they have predicted the need of a patient for hospital admission after the ED presentation. The model with the most accurate prediction accuracy had an accuracy of 76%, as opposed to a nurse accuracy of 67.7%. The research by Kim et al. (2014) is different from this paper since they focus on the chance of a patient leaving the ED or not instead of describing the causal structures on the patient outcomes mentioned in the introduction.

Aboagye-Sarfo et al. (2015) forecasted total ED demand in Western Australia with time-series techniques (VARMA, ARMA and Winters' method). Using the VARMA model they managed to get the most accurate prediction. The VARMA model predicted 1,143,812 ED presentations over a five year period in Western Australia, about 60.8% of these would be in metropolitan hospitals. Approximately 24.63% of the ED presentations are predicted to result in admission to the hospital

afterwards (Aboagye-Sarfo et al., 2015). Their study is different from this paper since Aboagye-Sarfo et al. (2015) used time-series techniques to forecast ED demand instead of describing causal structures on patient outcomes.

Cardona et al. (2018) evaluated the predictive validity of an Australian clinical prediction tool called 'Criteria for Screening and Triaging to Appropriate aLternative care (CRISTAL)' to identify short-term mortality risk (within three months) among older patients. Logistic regression regression was used in order to model death prediction. This was done based on two clinical outcome criteria: The Clinical Frailty Score (CFS) and the Fried Frailty Score. The Area Under the Receiving Operating Curve (AUROC) in Australia were 0.825 and 0.81, respectively. The variables age, male, advanced malignancy and nursing home residence had a positive significant effect on the short-term mortality (Cardona et al., 2018). Even though the study by Cardona et al. (2018) gives valuable insights, their study is different from this paper since their focus was on evaluating the predictive validity of a tool instead of predicting patients outcomes per se.

Liew et al. (2003) evaluated whether or not the length-of-stay in the ED predicts the excess inpatient length-of-stay (i.e. longer than the state average for the relevant admission problem). Based on logistic regression the researchers have shown that an 8-12 hours stay in the ED and stay of more than 12 hours in the ED has a positive significant effect on the excess inpatient length-of-stay. Apart from that, being older than 65 years, being female and being presented to the ED outside working hours (18:00-08:00) has a positive significant effect on excess inpatient length of stay (Liew et al., 2003). Their study focused on predicting whether or not a patient had an excess stay in the hospital instead of describing the causal structures on the patient outcomes mentioned in the introduction.

Apart from the reasons mentioned in each paragraph above, the studies used frequentist methods whereas this paper will focus on Bayesian methods. Besides, their research focuses on ED presentations in general instead of solely cardiac patients. However, they all focused on Australia as well. As to my knowledge, the Bayesian approach is completely new in this context. The Bayesian approach is beneficial in this research for the following reasons:

- Combined information from multiple sources can be incorporated in Bayesian models by using a prior. Besides, flexibility is introduced in Bayesian models by incorporation of multiple levels of randomness, while incorporating reasonable sources of uncertainty in inferential summaries. This allows for more realistic parameter estimates in complex data structures (Gelman et al., 2013), like the data structure of this research (see subsection data).
- In general, interval estimates are interpreted as Bayesian intervals by specialists in the medical field. More specifically, the interval estimates are interpreted as the probability that the interval contains a value of an unknown quantity, conditional on the data (Gelman et al., 2013).

Besides, my research uses the fully Bayesian approach as solution to missingness of data. As to my knowledge, this method has never been used in the context of (cardiac) patient data, let alone cardiac patient data in South Australia. A possible explanation for this is the fact that Bayesian inference gained much popularity due to exponential increase in computational power. In the past, Bayesian methods were less interesting due to the time it takes to run a desirable number of iterations.

3 Data

First of all, a textual description of the data is given. Furthermore, the descriptive statistics of the variables are presented. At last, the degree of missingness of every variable is evaluated.

3.1 Data description

In this subsection a textual description of the data is given. For this research a collection of the RAPIDx AI datasets is used. The datasets are the result of a study conducted on 9,600 patients in total at 12 South Australian Hospitals, of which 6 urban hospitals and 6 rural hospitals. [2] The inclusion criteria, for patients presented to the ED, are the following:

- Clinical features of chest pain or suspected acute coronary syndrome as the principal cause.
- At least one Troponin measure is conducted.
- 18 years or older.

All patients admitted to the participating hospitals are considered part of the trial, unless they decide to opt-out. The data is partly collected by taking measures (e.g. Creatine or Troponin level) in the Emergency Department whereas the rest of the data already exists in the database (e.g. Age or Gender).

These datasets have been provided by Flinders University and SAHMRI in South Australia. In total, 13 of the available datasets are used for this research. The Pandas library in Python is used for data processing and merging of the datasets. After merging and filtering of this data 3388 observations are left. An observation represents a chest pain patient that has been admitted to the ED, the information contained in one observation is described below. Duplicate admissions for one patient have been omitted for simplicity reasons. The final dataset contains 42 variables, that can (roughly) be divided in the following categories:

[2] Flinders Medical Centre, Royal Adelaide Hospital, The Queen Elizabeth Hospital, The Lyell McEwen Hospital, Noarlunga Hospital, Modbury Hospital, Mount Gambier Hospital, Murray Bridge Hospital, South Coast District Hospital, Berri Hospital, Port Augusta Hospital and Whyalla Hospital

- Clinical outcomes: The cardiac outcomes (Normal, Acute, Chronic, Type 1 Myocardial Infection (commonly known as: a heart attack) and Type 2 Myocardial Infection) and the length of stay in the ED (in minutes). See table 1 for a comprehensive description of the definitions of the cardiac outcomes.
- Clinical observations: troponin levels, white blood cell count, creatine level, lactate level etcetera. The first and second measurement are considered of the troponin levels of patients being admitted to the ED. In order to estimate the velocity of the troponin levels, the difference between the first and second troponin measurement is taken. Troponin is known to be highly correlated with the likelihood of having experienced a myocardial infarction (Fathil et al., 2015).
- Risk factors: Variables that are known to influence the chance of getting a myocardial infarction, like age, gender, the kidney function level, history of a heart attack, an ECG etcetera.

For a full description of all 42 variables, please refer to the table 8 Appendix .

Clinical outcome	**Description**
Normal	Rule-out myocardial injury - Recommendation for early discharge. Evidence for routine functional testing is limited. Consider involvement in clinical trial of early outpatient computed tomography coronary angiography (CTCA).
Type I MI	Type 1 MI likely - Consider early coronary angiography. Commence aspirin, statin and P2Y 12 inhibitor (choice of agent to be determined by co-morbidities).
Type II MI	Type 2 MI, in case cardiac likely - Consider management of concomitant heart failure or arrhythmia. Consider aspirin, but do not commence P2Y 12 inhibition. Consider echocardiogram. Cardiology team will determine coronary investigation. In case non-cardiac is likely, Treat primary presenting condition. Do not commence aspirin or P2Y 12 inhibition. Consult cardiology team for possible involvement in ongoing trials of coronary investigation for Type 2 MI.
Acute injury	Acute myocardial injury likely - Consider other non-coronary diagnoses including pulmonary embolus and aortic dissection. Consider echocardiogram. Do not commence aspirin or P2Y 12 inhibition. Consult cardiology team.
Chronic injury	Chronic myocardial injury likely - Consider other non-coronary diagnoses including pulmonary embolus and aortic dissection. Consider echocardiogram. Do not commence aspirin or P2Y 12 inhibition. Consult cardiology team for possible ongoing optimization of heart failure investigation and management.

Table 1: Description of cardiac outcomes and corresponding advice for treatment by SAHMRI. These textual descriptions are taken from the HeartAI website

3.2 Descriptive statistics

In this subsection the descriptive statistics are presented. Additionally, the effects that are expected (based on the literature) are presented here. Even though the cardiac outcome is categorical (and will be modelled as such), the expected effects are evaluated as binary in order to keep it comprehensive. Hence, the effects evaluated as either having a positive or negative effect on the probability of having a myocardial injury (i.e. T1MI Cardiac Outcome).

In figure 1 the distribution of the cardiac outcome variable is presented. As can be seen, for approximately 50% of the patient a heart attack is ruled-out (classified as Normal). Apart from that, a relatively big proportion of the patients is classified as either Acute or Chronic. A relatively small and approximately equal proportion of patients is classified as T1MI or T2MI. The exact counts and proportions of the Cardiac Outcome categories can be seen in table 2, along with the descriptives for Onset and Smoking.

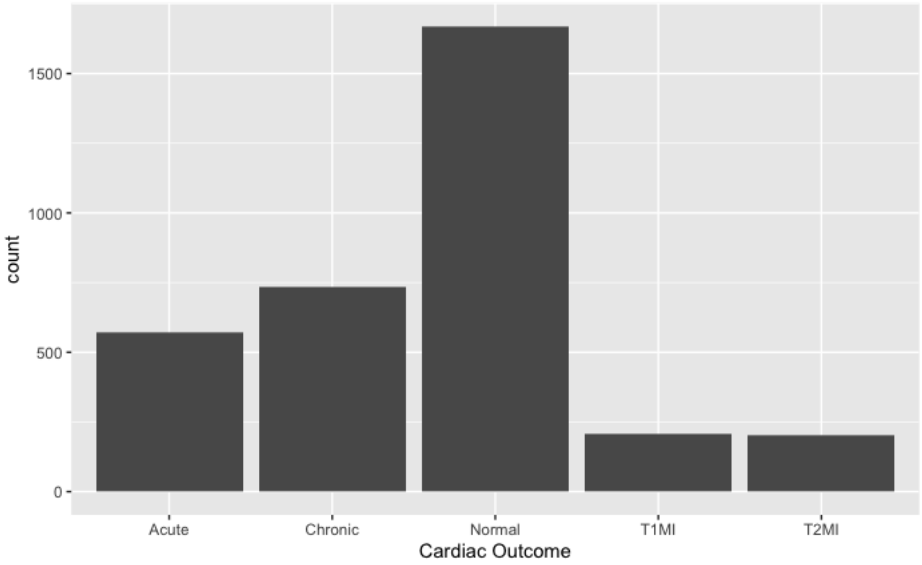


Figure 1: Distribution of the cardiac outcomes

Cardiac outcome		**Onset**		**Smoking**	
Acute	572 (16.88%)	*<1 hour*	117 (3.45%)	*Never*	896 (26.45%)
Chronic	733 (21.64%)	*1-4 hours*	346 (10.21%)	*Past*	491 (14.49%)
Normal	1670 (49.29%)	*4-6 hours*	170 (5.02%)	*Currently*	820 (24.2%)
T1MI	209 (6.17%)	*6-12 hours*	123 (3.63%)	*NA*	1181 (34.86%)
T2MI	204 (6.02%)	*12-24 hours*	119 (3.51%)		
NA	0 (0%)	*Above 24 hours*	270 (7.97%)		
		NA	2243 (66.2%)		

Table 2: Descriptive statistics of the categorical variables

Apart from that, the continuous risk factor are presented in table 3. The average age of the patients is 66.77 years. Age is expected to have a positive effect on the chance of having a myocardial injury (Members et al., 2014). The average Heart Rate falls within the 'normal' range of 60 to 100 beats per minute. The Heart Rate is expected to have a positive effect on the chance of having a myocardial injury (Perret-Guillaume et al., 2009). However, the average Systolic Blood Pressure is higher than what is considered to be 'normal' (less than 120 mm Hg). Moreover, the average

Diastolic Blood Pressure is just below the 'normal' range (less than 80 mm Hg). At last, the mean of the Kidney Function is within the 'normal' range (higher than 60). Both the Systolic Blood Pressure and the Diastolic Blood Pressure are expected to have a positive effect on the chance of having a myocardial injury (Everson-Rose and Lewis, 2005). Kidney Function is known to have a positive effect on the probability of having a myocardial infarction (Sarnak et al., 2003).

	NA	*Min*	*Max*	*Range*	*Median*	*Mean*	*Variance*
Length of Stay	0	0.67	3020.22	3019.55	243.40	523.69	490448.57
Age	0	14.09	105.00	90.91	67.34	65.77	318.05
Heart Rate	1178	36.00	182.00	146.00	75.00	77.94	338.80
Systolic Blood Pressure	1179	78.00	226.00	148.00	137.00	138.99	495.55
Diastolic Blood Pressure	2376	40.00	138.00	98.00	79.00	79.07	203.37
Kidney Function	1191	6.60	274.70	268.11	82.51	86.14	986.53

Table 3: Descriptive statistics length of stay & continuous risk factors

In figure 2 the distribution for the first and second troponin measure are presented. Please note that troponin values higher than 1000 are not included to keep the figure comprehensive. As can be seen in the figure, the density is higher for low Troponin First Measure values, as opposed to the Troponin Second Measure values. Moreover, the mean of the second troponin measure is higher compared to the first troponin measure (131.85 vs 85.67). This is most likely caused by the fact that patients with a low troponin level on the first measure are more often discharged (because the low troponin level indicates a lower risk of a myocardial infarction) and thus the patients remaining are more likely to have a higher troponin value. Hence, it can be stated that patients with a low troponin value are underrepresented in Troponin Second Measure. Whereas Troponin First Measure only contains 20 missing values, Troponin Second Measure contains 470 missing values.

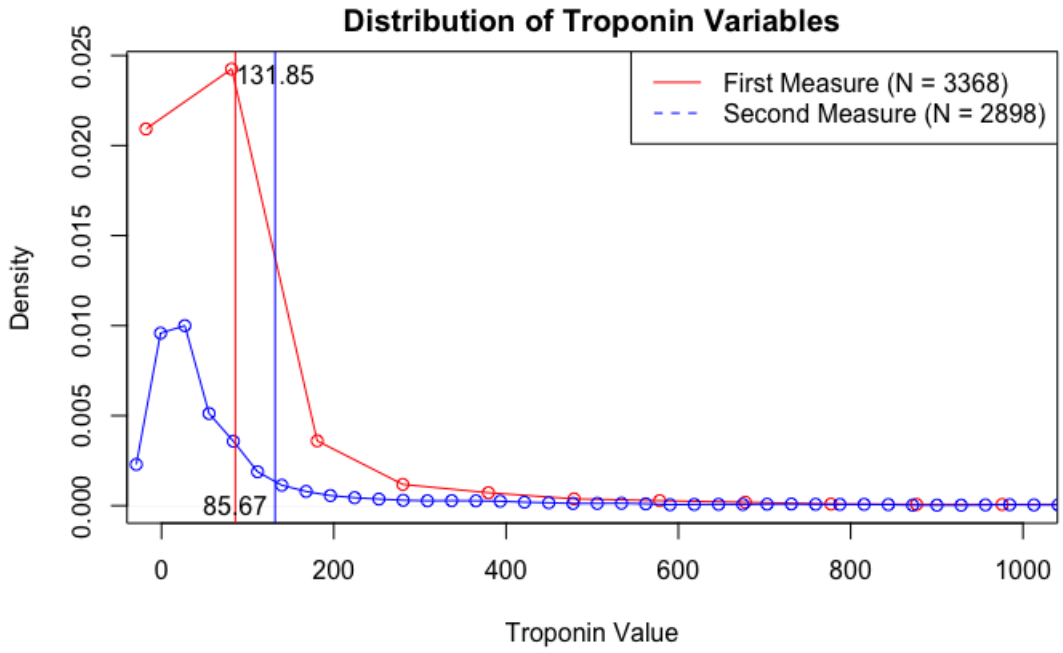


Figure 2: Distribution of the Troponin Variables

In figure 3 the distribution of Troponin Difference is presented. As can be seen in the figure, approximately 20% of the values are 0. The mean of the Troponin Difference values is 54.06. Troponin Difference contains 468 missing values. All Troponin First Measure, Troponin Second Measure and Troponin Difference are expected to have a positive effect on the chance of having a myocardial injury (Members et al., 2014).

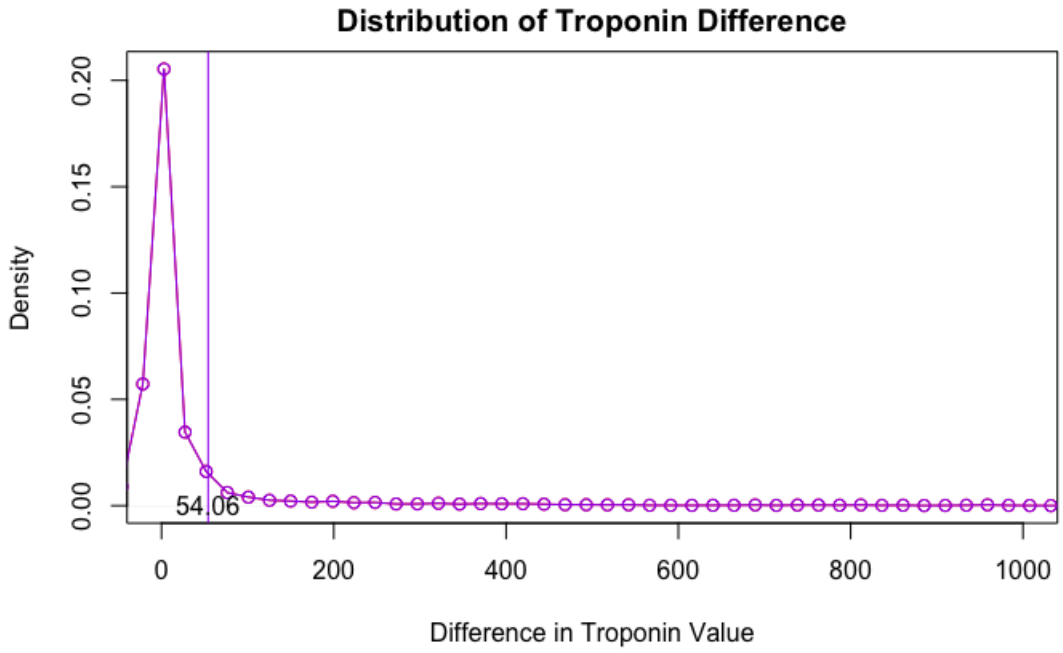
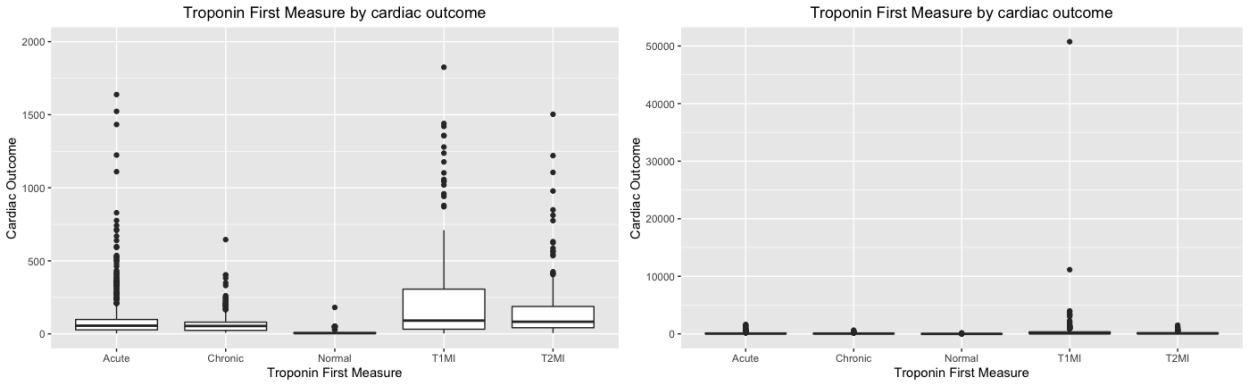


Figure 3: Distribution of Troponin difference

In subfigure 4a boxplots per Cardiac Outcome are presented for Troponin First Measure. The y-axis is cut off at a troponin value of 2000 in order to keep it orderly, which led to some observations being omitted. However, in subfigure 4b the full scale figure is presented. As can be seen in the figures, the Troponin First Measures are higher on average for the T1MI Cardiac Outcome group compared to the other categories. Specifically, the Troponin First measures values are relatively low for the Normal Cardiac Outcome group. Moreover, the interquartile range is relatively high for the T1MI group.

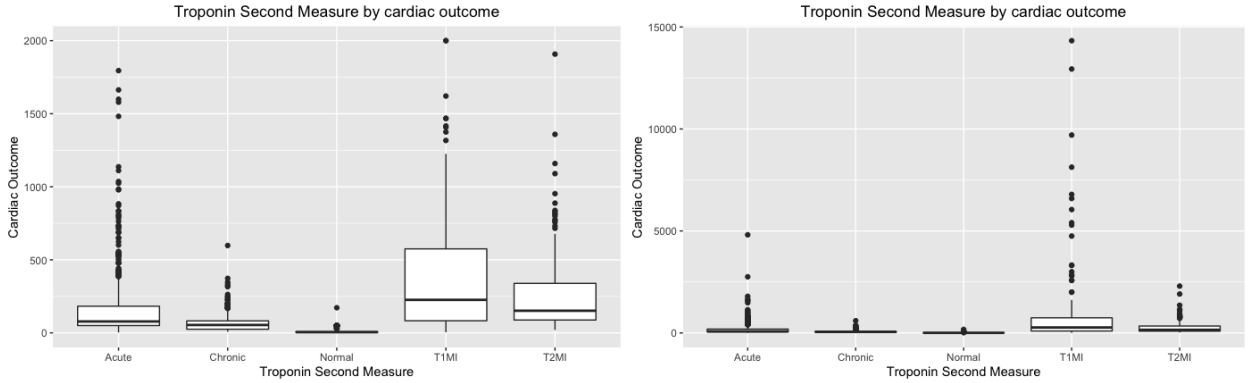


(a) Y-axis cut off at 2000

(b) Full scale

Figure 4: Distribution of Troponin First Measure

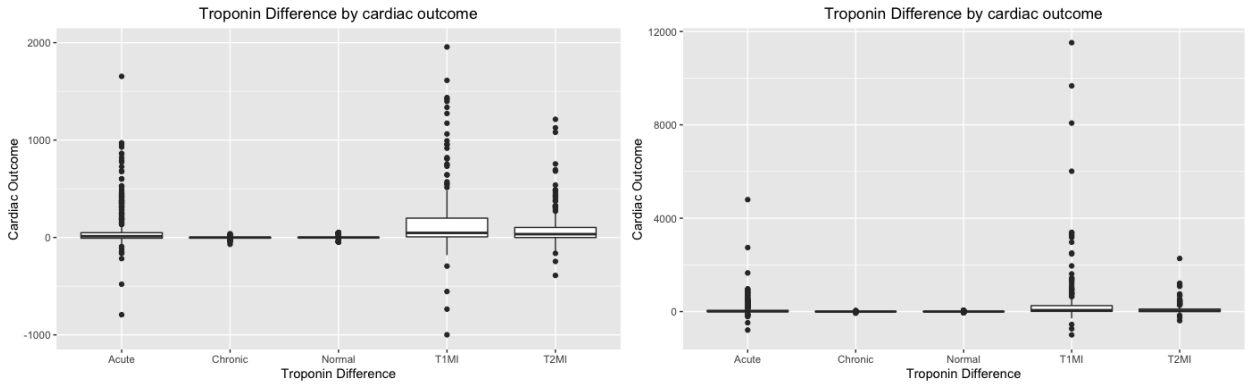
As can be seen in subfigure 5a, the Troponin Second Measure is relatively high for the T1MI Cardiac Outcome group. The previously mentioned observation bias (i.e. the second measure being higher because patients not at risk are being discharged) is also present in the boxplots, when comparing subfigure 5a to subfigure 4a.



(a) Y-axis cut off at 2000 (b) Full scale

Figure 5: Distribution of Troponin Second Measure

In figure 6 boxplots of the Tropinin Difference by Cardiac Outcome (y-axis cut off at 2000 and full-scale) are presented. Interestingly, there were a few observations with a substantial decrease in Troponin in the T1MI Cardiac Outcome group.



(a) Y-axis cut off at 2000 (b) Full scale

Figure 6: Distribution of Troponin Difference

The descriptive statistics for the physiological variables and the binary risk factors can be found in table 9 and table 10 in the appendix (section 8). Apart from that, the expected effects of the physiological variables and the binary risk factors can be found in table 8 in 8.

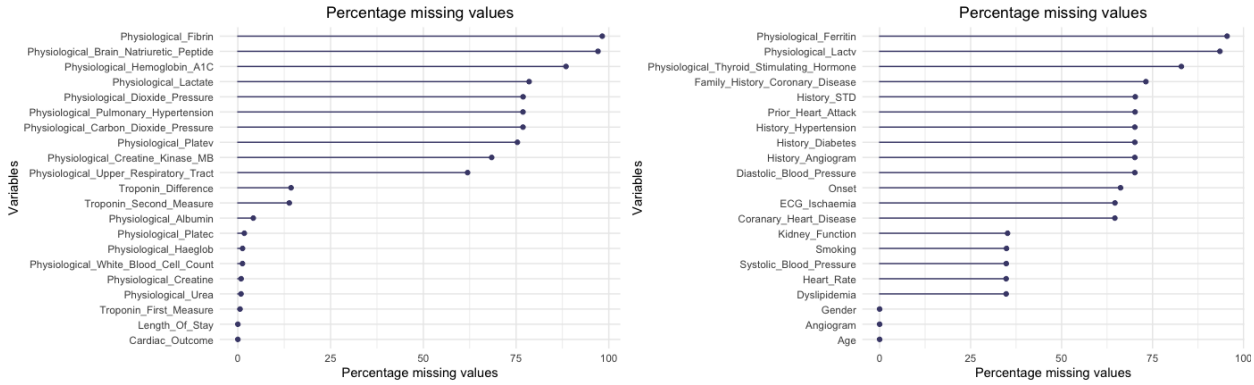
3.3 Missingness of data

3.3.1 Overall missingness

In this subsection the missingness of the variables is analysed. The percentage of missing values is presented for all 42 variables in figure 7. As can be seen in figure 7a, the dependent variable, Cardiac Outcome, does not contain any missing values.

The variables of interest, the troponin values, contain only 20 ($< 1\%$) missing values for the first troponin measure but 470 missing values ($> 10\%$) for the second troponin measure (see figure 7b). This is most likely caused by the fact that patients with a low tropinin value for the first measure (and thus at less risk for a heart attack) are discharged before getting a second troponin measure. This assumption is supported by the fact that the mean of the Troponin Second Measure is substantially higher compared to the mean of the Troponin First Measure (see figure 2).

Whereas some risk factor variables contain no missing values (i.e. Gender, Angiogram and Age), some risk factor variables (e.g. Prior Heart Attack, Diastolic Blood Pressure or History Hypertension) contain approximately 70% missing values (see figure 7b). The physiological variables generally contain a relatively high proportion of missing values, this can be seen in figure 7.



(a) First half of variables

(b) Second half of variables

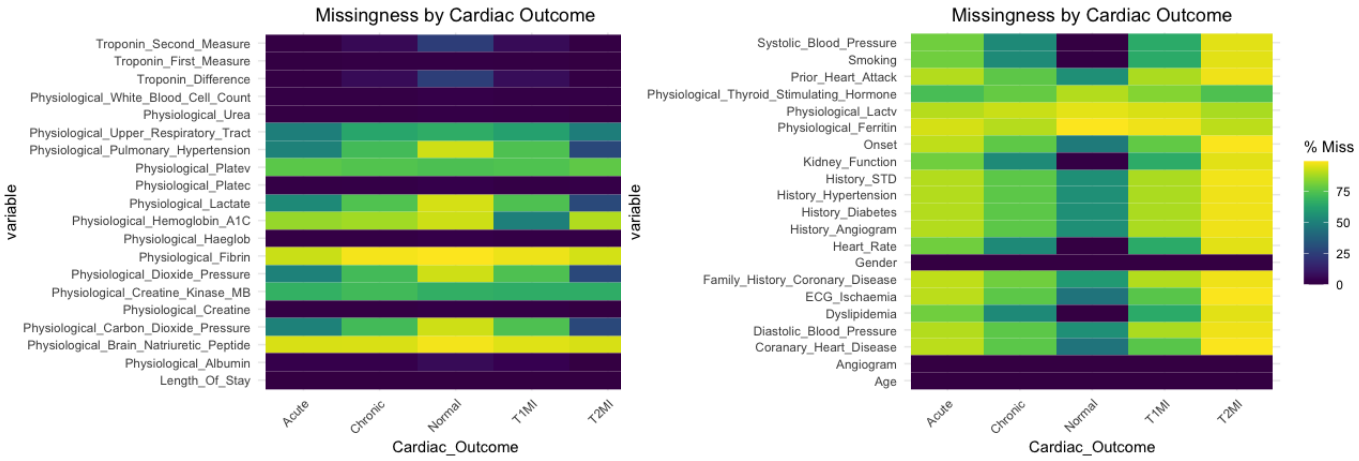
Figure 7: Percentage of missing values

3.3.2 Missingness by Cardiac Outcome

In figure 8 the percentage of missing values by Cardiac Outcome factor is presented [3]. Interestingly, the Troponin Second Measure (and thus, Troponin Difference) mostly contains missing values for the 'Normal' group. This is likely caused by the fact that patients with a low troponin value for the first measure are discharged and thus their second troponin measure is missing. In combination with

[3]This figure only contains 41 variables on the y-axis because their missingness is evaluated by cardiac outcome, one of the variables from the original dataset

the increase in the mean of the troponin value (see figure 2), this is a strong indication for Missing Not at Random (MNAR) since the missingness of Troponin Second Measure seems to depend on both the y-variable (Cardiac Outcome) and the missing values (the value would have been lower than average if the patient wasn't discharged). In subsection 4.4 an overview of the missing data mechanisms can be seen.



(a) First half of variables

(b) Second half of variables

Figure 8: Percentage of missing values by Cardiac Outcome factor

Surprisingly, as can be seen in figure 8b, the risk factor variables (e.g Heart Rate or Kidney Function) generally contain less missing values for the Normal category compared to the other Cardiac Outcome categories. Contrarily, most of the physiological variables (e.g. Pulmonary Hypertension or Dioxide Pressure) contain more missing values for the Normal category compared to the other Cardiac Outcome categories. A possible explanation for this is the fact that these measures are only taken (either during admittance in the ED or prior to the admittance) for patients that are considered to be at risk of a myocardial infarction.

4 Methodology

In this section of the paper the models used for analysis are discussed. Firstly, the Bayesian GLMM and the corresponding settings in R are discussed. Secondly, the variable and model selection methods are discussed. Thirdly, the missing data mechanisms and corresponding possible solutions are discussed. Fourthly, the scenarios that will be evaluated are discussed. At last, sensitivity analysis is discussed.

4.1 Bayesian GLMM

4.1.1 The model

The first model that will be used for prediction of patients outcomes in cardiac patients admitted to the Emergency Department is the Bayesian generalized linear mixed model (GLMM). One can think of GLMMs as an extension of generalized linear models (GLM) by also incorporating varying effects. In the Bayesian context, the varying effects can be incorporated within the (hyper-)parameters. A common way to do this for a GLMM is by the usage of an 'adaptive prior' for the intercept, meaning that the intercept β_j is a function of parameters as well (e.g. the intercept being $\beta_0 + b_{0j}$ with $b_{0j} \sim N(\hat{b}_{0j}, \sigma)$). Consequently, the prior adaptively pools the information across the groups (McElreath, 2020). The GLM itself is an alternative of Ordinary Least Squares (OLS) that allows for a non-normal error distribution (Agresti, 2015). A linear mixed model contains both non-varying regressors and varying effects (varying regressors) (Cameron and Trivedi, 2005). Linear mixed models are commonly used for cross-sectional data on subjects nested in hospitals (Chung et al., 2013).

Mathematically, a GLMM looks as follows (McElreath, 2020):

$$Y_{ij} = f(\pi_{ij}) = \beta_0 + b_{0i} + (\beta_1 + b_{1i})X_{ij} \tag{1}$$

Where $f(\pi_{ij})$ is the link function, β_1 is a vector of parameters and b_{1i} a vector of group-level effects. This function relates the expected value of the response of the linear predictors in the model. It is determined independently from the distribution choice (McElreath, 2020).

In this paper, the categorical link is used. Mathematically, this looks as follows (Hadfield et al., 2010):

$$\begin{aligned} f(\pi_{ij}) = P[y_i = j] &= \frac{1}{1 + \sum_{h=2}^{m} exp(\beta_0 + b_{0i} + (\beta_1 + b_{1i})X_{ij})} \; for \; j = 1 \\ f(\pi_{ij}) = P[y_i = j] &= \frac{exp(\beta_0 + b_{0i} + (\beta_1 + b_{1i})X_{ij})}{1 + \sum_{h=2}^{m} exp(\beta_0 + b_{0i} + (\beta_1 + b_{1i})X_{ij})} \; for \; j = 2, ..., m \end{aligned} \tag{2}$$

Where β_1 is a vector of parameters.

The Bayesian variant of the GLMM allows for modelling of varying effects in longitudinal data (Zhao et al., 2006). For interpretation of the parameter estimates, the probability density function of the posterior distribution of the Bayesian GLMM can be evaluated. The credible intervals (C.I.) are evaluated. The credible intervals contain a certain posterior probability mass (McElreath, 2020). One advantage of a Bayesian approach over its frequentist counterpart (GLMM) is the fact that it better accounts for uncertainty of the variance component for the prior, likelihood and the posterior distribution (Handcock and Stein, 1993) (Zhao et al., 2006) (Diggle et al., 1998). An additional

advantage of the Bayesian GLMM is the fact that it is computationally more simple to obtain variance estimates of the random effect predictions (Zhao et al., 2006).

4.1.2 Model settings

STAN is a probabilistic programming language that implements full Bayesian statistical inference via Markov Chain Monte Carlo (Carpenter et al., 2017). STAN in R makes use of the No-U-Turn sampler (NUTS), which is an extension of Hamiltonian Monte Carlo (HMC) simulation, for sampling from an estimation of the posterior. HMC uses derivatives of the target probability density w.r.t. the parameters, the sampling can be done in less steps and thus HMC is faster as opposed to the simple random walk methods (Neal et al., 2011). Hence, the random walk behaviour, that comes with Gibbs sampling, is mostly eliminated (Betancourt and Girolami, 2015). However, the desired number of steps L has to be specified by the user. Values of L that are too small lead to undesirable random walk behaviour and large values of L is a waste of computation (Hoffman and Gelman, 2014). Hence, the NUTS sampler was introduced which does not require the use to specify L. The NUTS sampler searches for a set of likely candidate points (Hoffman and Gelman, 2014).

For the non-varying effect parameters the $student(nu, 0, \sigma)$ prior is used. nu represents the degrees of freedom, 0 is the location parameter and σ is the scale parameter. σ is related to the standard deviation, namely, $SD = \frac{nu}{nu-2}\sigma$. Hence, for large nu the standard deviation is equal to σ (Fonseca et al., 2008). This prior is commonly used in the literature as weakly informative prior for models with a discrete outcome (Gelman et al., 2008). Hence, it is a suitable choice for the models with the Cardiac Outcome as dependent variable. Specifically, the prior used for the non-varying effect parameters is the $student(3, 0, 5)$. Gelman (2020) advises to use degrees of freedom between 3 and 7. However, no conclusive results exist on what specific value should be chosen. The default degrees of freedom in the 'brms' package is also 3 (Bürkner, 2017b). The 'brms' package can be used to imlement Bayesian multilevel models in R using STAN (Bürkner, 2017b).

The most preferred prior option for the covariance matrix is the Lewandowski, Kurowicka and Joe (LKJ) correlation distribution, which can be interpreted as a symmetric Beta distribution (Bürkner, 2017b). The advantage of this prior for the covariance matrix is the fact that it requires less computational time compared to the so called 'onion-method' by Ghosh and Henderson (2003) and the 'D-vine method' by Joe (2006). Mathematically, it looks as follows (Bürkner, 2017b):

$$\Omega \sim LKJ(\eta) \tag{3}$$

Where Ω is the correlation matrix and the parameter $\eta > 0$.

If $\eta = 1$, then the probability density function (PDF) is uniform over correlation matrices of order K. For values of η higher than 1, the PDF concentrates around the identity matrix. Hence, less

correlation is favored (Bürkner, 2017b). Since the value of η can be adjusted, the expected amount of correlation among the parameters β_j can be controlled with the use of the LKJ prior (Bürkner, 2017b). In other words, flexibility exists in expressing prior beliefs about the correlations between variables, independently of the variances itself, by using the LKJ prior. This has proven to lead to a reduced absolute bias of the coefficients (Follett and Vander Naald, 2020). The value of η is usually set to 2 and this is also the value that is used in this research. In figure 9 the density of the LKJ prior for different values of η is presented.

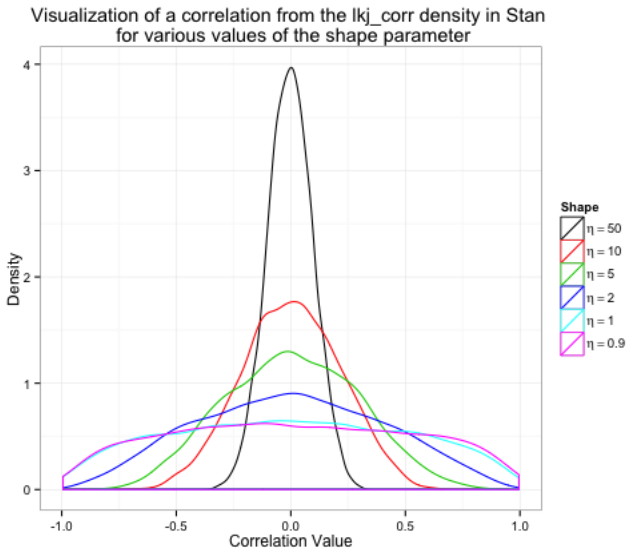


Figure 9: LKJ prior for different values of η

One disadvantage of the LKJ prior is the fact that the LKJ prior places strong assumptions on the correlation matrix. Another disadvantage is the fact that the non-conjugancy of the LKJ prior leads to a posterior distribution which is less convenient in terms of deriving closed-form analytical solutions. However, the conjugacy is not a problem in terms of computation in STAN (Follett and Vander Naald, 2020).

Another widely used prior for the covariance matrix is the Inverse Wishart (IW) prior. However, this prior is not used in this research since it is known to cause dependencies between correlations and variations (Akinc and Vandebroek, 2018b). Even though the Scaled IW prior reduces the dependencies between correlations and standard deviations, it does not fully eliminate them (Akinc and Vandebroek, 2018a).

The Bayesian GLMM can be implemented with the 'brm' function from the 'brms' package, which uses STAN on the back end. The family parameter is set to 'categorical' because the dependent variable (Cardiac Outcome) is categorical. The 'Normal' category of Cardiac Outcome is used as

reference point because this is the most logical in terms of interpretation. The number of iterations is set to the default setting in 'brms', namely, 2000, and this will be increased if the convergence criteria are unacceptable (see subsection 4.1.3). The number of Markov chains and cores are set to 4 because my laptop has 16 available threads. With 4 chains and 4 cores parallelisation, over all threads is possible (4 times 4). The 'control' argument can be used to control the behavior of the NUTS sampling, it contains the 'adapt delta' and 'max treedepth' parameters. The 'adapt delta' parameter is set to 0.96 in order to decrease the number of divergent transitions (which lead to a bias in the posterior sample), even though this slows down the NUTS sampler. The 'max treedepth' parameter is set to 12 in order to prevent the evaluated three depth in each iteration being exceeded (Bürkner, 2017a).

Since the models can take quite long to run, it can be useful to run the models parallel. With use of the 'future' package, the models are run in parallel by inserting 'plan(multisession(workers = 4))' into the Rscript. This reduces the running time to the running time of the slowest model, which is most likely the most complex model.

4.1.3 Evaluation of the convergence

The Gelman-Rubin statistic, also known as the potential scale reduction factor (PSRF) or $\hat{R}$, is evaluated as indicator for likely convergence of the MCMC (Brooks and Gelman, 1998). The Gelman-Rubin statistic is an approximation of convergence of the MCMC by using the between-chain variation and the within-chain variation (Vats and Knudson, 2018). Mathematically, the between-chain variation looks as follows:

$$B = \frac{n}{m-1} \sum_{j=1}^{n} (\theta_j - \overline{\theta_j})^2 \tag{4}$$

Where j represents the j-th chain, m the number of chains and n the number of draws. The within-chain variation looks as follows:

$$W = \frac{1}{m} \sum_{j=1}^{m} s_j^2 \tag{5}$$

where

$$s_j^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\theta_j - \overline{\theta_j})^2 \tag{6}$$

The variance of the stationary distribution is then estimated as follows:

$$\hat{Var}(\theta) = (1 - \frac{1}{n})W + \frac{1}{n}B \tag{7}$$

Consequently, the univariate PSRF or $\hat{R}$ can be calculated as follows (Gelman et al., 1992):

$$\hat{R} = \sqrt{\frac{\hat{Var}(\theta)}{W}} \tag{8}$$

At last, the multivariate $\hat{R}$ or multivariate potential scale reduction factor (MPSRF) is calculated as follows:

$$MPSRF = \frac{n-1}{n} + (\frac{m+1}{m})\lambda_1 \tag{9}$$

Where λ_1 is the largest eigenvalue of the symmetric, positive definite matrix:

$$\frac{W^{-1}B}{n} \tag{10}$$

In this research, a threshold of the multivariate $\hat{R}$ of 1.1 is used, as recommend by Gelman and Rubin (Gelman et al., 1992). For the univariate $\hat{R}$ a treshold of 1.05 is used, which is the second most used threshold in the existing literature (Vats and Knudson, 2018).

4.2 Variable selection

4.2.1 Modular construction

Variable selection is done based on the methods proposed by Gelman et al. (2020) in their 'Bayesian Workflow' paper. For variable selection, Gelman et al. (2020) use a method called 'modular construction'. Based on domain knowledge and literature, an initial likelihood function and priors are chosen. Consequently, this base model is extended by adding variables (i.e. modules) and these extensions are compared to the base model (or an improved extension). When a local optimum is found, this likelihood specification is used to proceed with the modelling with a higher number of iterations. This variable selection method is chosen because, with the guidance of SAHMRI and Flinders University, a proper initial likelihood function and prior can be chosen based on domain knowledge. In a context where the researcher has such limited knowledge about the context of the data, it might be better to resort to Bayesian Lasso, the Horseshoe prior or Stochastic Search Variable Selection (see below).

In this paper, the initial model is specified as follows:

$$\begin{aligned} Cardiac\,Outcome = log(Troponin\,First\,Measure) \\ + (Troponin\,First\,Measure|ECG\,Ischaemia) \end{aligned} \tag{11}$$

with the priors mentioned above. Consequently, variables are like Troponin Difference or Creatine are iteratively added. The order of adding these variables is done based on domain knowledge and

literature.

ECG Ischaemia is chosen as varying effect component because cardiac patients receiving a ECG can be considered as a separate sub-population, similarly to cardiac patients not receiving an ECG.

The models found with the iterative approach are presented in a directed acyclic graph, which is a graphical representation of causal structures (McElreath, 2020). An arrow from X to Y implies a causal effect of X on Y.

4.2.2 Alternative methods

Apart from modular construction, other variable selection methods exist in the Bayesian context. The following widely-used Bayesian variable selection methods [4] were considered but are not the most preferred choice in this research:

- Bayesian Lasso. This method was not the most preferred due to the fact that, in case a group of explanatory variables is correlated, Bayesian Lasso only selects one of these explanatory variables (Van Erp et al., 2019).
- Horseshoe prior. With this method it is not possible to specify a prior on the non-varying effect parameters (Piironen et al., 2017).
- Stochastic Search Variable Selection (SSVS): This method is known to be computationally very intensive (Srivastava and Chen, 2009).

4.3 Model selection

For the selection of the models, the following methods are proposed: the Watanabe-Akaike Information Criterium (WAIC) and the Leave-one-out Cross Validation (LOOCV). Different models (see subsection 4.5 for the scenarios) are estimated and these are compared by abovementioned model selection criteria. Apart from that, the models are evaluated based on their predictive accuracy, specificity and sensitivity.

4.3.1 Information criteria

The model fit is based on the Leave-One-Out Cross Validation (LOOCV). If the LOOCV is indecisive (i.e. <1% difference), the Watanabe–Akaike information criterion (WAIC) is used. The Bayes Factor is not considered in this paper since research has shown that it is strongly dependent on irrelevant aspects of the model (Gelman and Yao, 2020).

LOOCV

LOOCV is an estimation of the out-of-sample log posterior predictive density (lppd). In case of N

[4] Bayesian Lasso and the Horseshoe prior are technically shrinkage priors

observations the model is fit N times and one observation y_i each time. The LOOCV can then be seen as the sum of the average accuracy for each omitted y_i (McElreath, 2020).

Mathematically, the LOOCV looks as follows:

$$LOOCV = \sum_{i=1}^{N} \frac{1}{S} \sum_{s=1}^{S} log\ \Pr(y_i|\theta_{-i,s}) \tag{12}$$

The disadvantage of LOOCV is the fact that it is computationally expensive to compute N posterior distributions in case of N observations. Luckily, methods exist to approximate the CV score without having to run the model over and over again (McElreath, 2020). Pareto Smoothing Importance sampling assigns a weight to observations that have a larger influence on the posterior distribution. An estimation of the model's out-of-sample accuracy can be done based on these weights. It can be seen as the Bayesian version of importance sampling with a prior on the largest importance ratios (Vehtari et al., 2015).

WAIC

The AIC is not useful in hierarchical modelling since the prior effectively restricts the freedom of the model parameters. Hence, the appropriate number of parameters is generally unclear (Spiegelhalter et al., 2014). An innovation on the AIC is the Watanabe-Akaike information criterion as proposed by Watanabe and Opper (2010).

Mathematically, the WAIC looks as follows (McElreath, 2020):

$$\begin{gathered} WAIC(y,\theta) = -2(lppd(y,\theta) - \sum_i var_\theta\ log\ p(y_i|\theta)) \\ \text{with } lppd(y,\theta) = \sum_i log \frac{1}{S} \sum_s p(y_i|\theta_s) \end{gathered} \tag{13}$$

Where S is the number of samples

The log-pointwise-predictive-density (lppd) can be considered as the Bayesian version of the log-probability score (McElreath, 2020). The second part of the WAIC formula between brackets is the penalty term P_{WAIC}:

$$\sum_i var_\theta\ log\ p(y_i|\theta)) \tag{14}$$

P_{WAIC} can be considered as the effective number of parameters. In case of multilevel models, the effective number of parameters can reduce by the addition of parameters to the model (McElreath, 2020). The model with the lowest WAIC is selected.

4.3.2 Accuracy

As previously mentioned, the models are evaluated based on their predictive accuracy. The percentage of correctly classified observations is evaluated for each model and this is visualised with a confusion matrix. Apart from that, the specificity and sensitivity are evaluated. Even though the Cardiac Outcome consists of five categories, it can be divided into having a T1MI (Type 1 Myocardial Injury) or having another Cardiac Outcome (Normal, Acute, Chronic or T2MI). T1MI will be classified as a 'positive' and the other categories as a 'negative'. In this scenario, a false negative could have disastrous consequences whereas a false positive would only lead to higher costs because the patients has to be hospitalized.

Even though the accuracy, sensitivity and specificity are considered in this research, it is not the main goal of this paper to attain the highest possible accuracy. As previously mentioned, one of the goals of this research is to describe the causal structures of cardiac patient data on patient outcomes. The Bayesian GLMM, and Bayesian multilevel models in general, are not focused on attaining the highest prediction accuracy per se, but more on describing the causal structures, in contrary to machine learning methods like Neural Networks. However, from a point of view of the RAPIDxAI team at SAHMRI, the difference in accuracy amongst different likelihoods is interesting and thus it is considered in this research.

4.4 Missing data

Missingness of data is a common problem in cardiac patient data sets (Faris et al., 2002). Before discussing the multiple solutions to the problem of missing data, it is useful to analyse the nature of the missingness of the data. Denote $X = \{X_{miss}, X_{obs}\}$ as the variable with missing observations and Y as the dependent variable, without any missing values. Multiple missing data mechanisms exist in the Bayesian context, namely (Gelman et al., 2013):

- *Missing Completely At Random* (MCAR), in which the missing data mechanism is completely independent of the distribution of X and the distribution of Y.

$$P(X_{miss}|X, Y) = P(X_{miss}) \tag{15}$$

- *Missing At Random* (MAR),in which the missing data mechanism does not depend on the distribution of X but only on the distribution of Y.

$$P(X_{miss}|X, Y) = P(X_{miss}, Y) \tag{16}$$

- *Missing Not At Random* (MNAR), in which the missingness depends on both the distribution of X and the distribution of Y.

$$P(X_{miss}|X, Y) = P(X_{miss}|X_{miss}, X_{obs}, Y) \tag{17}$$

The missing data mechanism has to be determined based on the data. Most likely, the missingness mechanism is MAR or MNAR because patient data is lost due to patients dying. The latter is greatly dependent on variables like age, medical condition, reason for presentation etcetera. No statistical tests exist to differentiate between MAR and MNAR (Nakagawa, 2015). Hence, the missing data mechanism is determined based on domain knowledge.

The most common solution to solving the problem of missing data is simply dropping all observations containing missing values (complete case analysis) (McElreath, 2020). However, this can lead to statistical issues (i.e. biased estimations) if the data set is too sparse (which is probable with patient data) and it is a waste since patient data collection can be costly.
In the Bayesian context, there are two main ways of solving the problem of missing data (Gelman et al., 2013):

- Multiple imputation: Simulation of draws from the posterior predictive distribution $P(y_{mis}|y_{obs})$. Sampled values of parameter Θ are used to impute the missing data in $P(y_{mis}|y_{obs})$ M times. The pooled estimate is then based on the M estimates (Linero and Daniels, 2018).
- Directly taking draws from the posterior distribution of model parameters θ, this is considered to be the fully-Bayesian approach. Obtaining samples of Θ is usually done by MCMC. The disadvantage of the fully-Bayesian approach is the fact that it is computationally expensive. However, this problem can be avoided by using an informative prior (Linero and Daniels, 2018).

In practice, multiple imputation in combination with Bayesian inference is often used even though it is solely a viable option in case normality of the posterior distribution justifiable (Zhou and Reiter, 2010). Even though multiple imputation is easier to apply since the variables used in the imputation do not have to be explicitly specified, the fully Bayesian approach is preferred since it is still viable when normality of the posterior distribution is not justifiable (Zhou and Reiter, 2010). Besides, with the fully Bayesian approach the multilevel structure of the Bayesian GLMM can be used for the imputation of missing values within the 'brms' framework in R (Burkner, 2015). Hence, in this thesis the use of the fully Bayesian approach is used. The fully Bayesian approach is also referred to as the 'one-step' approach because the imputation is done during the fitting of the model. A disadvantage is the fact that discrete variables cannot be estimated since STAN does not support this (Burkner, 2015).

The problem of missing data in patient data is very common and has undesirable statistical consequences like estimation biases and reduction in power (Bell and Fairclough, 2014). Hence, it can be of great value to evaluate the effectiveness of the fully Bayesian approach as solution to the problem of missing patient data.

4.5 Scenarios

As mentioned in the previous subsection, the fully-Bayesian approach solution to missingness of data is implemented. In order to evaluate the effectiveness of this method, the following scenarios are evaluated (for every patient outcome):

- Model 1: No solution for the problem of missing data. STAN simply removes all observations that contain a missing value.
- Model 2: The fully Bayesian approach as solution to the problem of missingness of data.

The imputation of the variables in the second model is done as follows:

- Step 1: Estimate the missing *Troponin Second Measure* values in a 'sub-model'.
- Step 2: Calculate (not estimate) the *Troponin Difference* based on the *Troponin First Measure* (observed, this variable hardly contains missing values) and *Troponin Second Measure* (observed + imputed) values.
- Step 3: Estimate the main outcome model component.

Unfortunately, STAN in R does not support linear transformation (i.e. the subtraction mentioned in step 2) of imputed variables (i.e. *Troponin Second Measure*) [5]. Hence, the imputation of the variables has to be done in a sub-optimal way. Namely, as follows:

- Step 1: Calculate the observed *Troponin Difference* based on the difference between the *Troponin Second Measure* and *Troponin First Measure*.
- Step 2: Estimate the missing *Troponin Difference* values in a 'sub-model'.
- Step 3: Estimate the main outcome model component.

Estimating the missing values of *Troponin Difference* in a 'sub-model' is sub-optimal (in comparison to estimating missing *Troponin Second Measure* values) due to the nature of the data. Besides, modelling *Troponin Second Measure* is more suitable based on the domain knowledge/literature.

As previously mentioned, $ECG\ Ischaemia$ is modelled as varying effect component. However, this variable contains 2191 missing observations (out of 3388, see table 10). Unfortunately, STAN does not allow estimation of discrete parameters and thus $ECG\ Ischaemia$ cannot be imputed with the

[5] An attempt was done by adding both Troponin measures linearly and calculating as follows: -*Troponin First Measure* = *Troponin Second Measure*. However, this did not improve the model performance. This is most likely caused by perfect multicollinearity. Namely, the correlation between *Troponin First Measure* and *Troponin Second Measure* is approximately 0.7

fully Bayesian approach. Hence, 2191 observations are omitted. Even though $ECG\ Ischaemia$ contains relatively many missing observations, the models including the variable as varying effect component still performed the best, in comparison with models containing another variable as varying effect component.

The models are compared by the model selection criteria (LOOCV and WAIC) as described in subsection 4.3.1.

4.6 Sensitivity analysis

Prior research on Bayesian multilevel models has shown that the results can depend heavily on the prior assumptions (Geisser, 1993) (Roos et al., 2011). Hence, sensitivity analysis is performed in this research. With sensitivity analysis, the influence of changes in modelling assumptions (e.g. likelihoods, priors or link functions) on the Bayesian inferences are explored (McElreath, 2020). In this research, sensitivity analysis is conducted by adjustment of the parameter(s) of the $student_t(nu, 0, s)$ and the LKJ prior.

In the sensitivity analysis (for both model 1 and model 2), the $student_t(7, 0, 5)$ distribution is used as prior on the non-varying effect parameters and the $lkj(4)$ is used as prior for the correlation matrix. The degrees of freedom are changed to 7 since it is the other extreme of the range of recommended degrees of freedom for a $student_t$ prior, namely, $3 < nu < 7$ (Gelman et al., 2020).

5 Results

First of all, the final likelihood specifications and the corresponding directed acyclic graphs of the models are presented. Second of all, the model output and predictions of model 1 are evaluated. Furthermore, the model output and predictions of model 2 are discussed. Then, the results of the sensitivity analysis are evaluated and compared to the 'original' models. At last, the results of the two models are compared.

5.1 Directed acyclic graphs

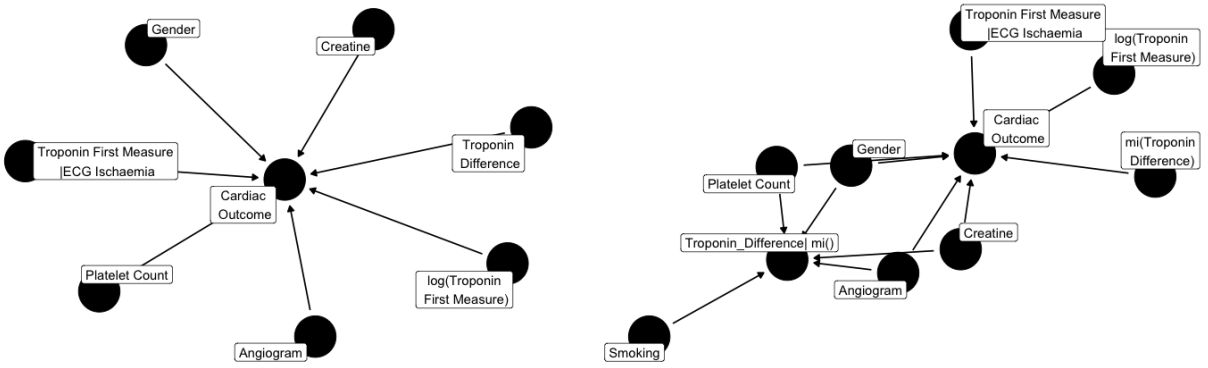
In figure 10 the directed acyclic graphs (DAGs) of model 1 and model 2 are presented. These likelihoods specifications are a result of the 'modular construction'. As can be seen in subfigure 10a, the final likelihood specification of model 1, the model without a solution for missingness of data, is as follows:

$$\begin{aligned} Cardiac\,Outcome \sim log(Troponin\ First\ Measure) + Troponin\ Difference + \\ Physiological\ Platec + Angiogram + Gender + Physiological\ Creatine + \\ (Troponin\ First\ Measure | ECG\ Ischaemia) \end{aligned} \tag{18}$$

The final likelihood specification of model 2, the model with the Fully Bayesian approach as solution for missing data, looks as follows (see subfigure 10b):

$$\begin{aligned} Cardiac\,Outcome \sim log(Troponin\ First\ Measure) + mi(Troponin\ Difference) + \\ Physiological\ Platec + Angiogram + Gender + Physiological\ Creatine + \\ (Troponin\ First\ Measure|ECG\ Ischaemia) \end{aligned} \tag{19}$$

$$\begin{aligned} Troponin\ Difference|mi() \sim Physiological\ Creatine + Physiological\ Platec + \\ Gender + Angiogram + Smoking \end{aligned}$$



(a) Directed acyclic graph of model 1

(b) Directed acyclic graph of model 2

Figure 10: Directed acyclic graphs

5.2 Model 1

5.2.1 Model output

In table 4 the output of model 1 (see formula 18 for the DGP) is presented. This is the model without any imputation of missing data. Furthermore, the performance criteria and the model settings are presented in the table. Moreover, the STAN code corresponding to the model can be seen in the technical appendix (see section 9).

As can be seen in table 4, all univariate R-hats are lower than 1.05, indicating convergence. Besides, the multivariate R-hat is below 1.10, which also indicates convergence. The Leave-One-Out Cross-Validation cross validation (LOOCV) of model 1 is 652.6.

Unsurprisingly, a positive significant effect exists (at a 5 % significance level) of the logarithm of the first troponin measure on the probability of having a Type 1 Myocardial Injury (heart attack). The effect can be interpreted as follows: if the Troponin First Measure goes up by 1%, the probability of having a Type 1 Myocardial Injury increases by approximately 3.89%, keeping other variables constant. Similarly, positive significant effects of the logarithm of Troponin First Measure on the probability of being in the Cardiac Outcome categories Acute, Chronic and T2MI. Moreover, a positive significant effect exists of the logarithm of Troponin Difference on the probability of having a T1MI. If Troponin Difference goes up by 1%, the probability of having a T1MI increases by approximately 0.13%, ceteris paribus. Similar effects exist of Troponin Difference on the probability of being in the Cardiac Outcome categories Acute, Chronic or T2MI. Besides, a positive significant effect exists of a patient having had an Angiogram on the probability of having a T1MI. Moreover, a positive significant effect exists of Gender on the probability of having a T2MI. Apart from that, no signifcant effects (except for the intercepts) exists. However, positive significant effects of the Platelet count and the Creatine level on the modelled Cardiac Outcome categories were expected.

Varying effects				
Group	*Variable*	*Estimate [Credible Interval (C.I.)]*		*Rhat*
Acute	sd(Acute Intercept)	3.0798 [0.34; 8.63]		1.01
	sd(Acute Troponin First Measure)	0.2903 [0.02; 1.42]		1.00
	cor(Acute Intercept,muAcute Troponin First Measure)	-0.0748 [-0.84; 0.75]		1.01
Chronic	sd(Chronic Intercept)	1.9682 [0.07; 6.98]		1.00
	sd(Chronic Troponin First Measure)	0.2647 [0.01; 1.64]		1.00
	cor(Chronic Intercept,muChronic Troponin First Measure)	-0.0586 [-0.84; 0.77]		1.00
T1MI	sd(T1MI Intercept)	2.1559 [0.05; 8.33]		1.00
	sd(T1MI Troponin First Measure)	0.4832 [0.05; 2.69]		1.01
	cor(T1MI Intercept,muT1MI Troponin First Measure)	0.0018 [-0.82; 0.81]		1.01
T2MI	sd(T2MI Intercept)	3.2025 [0.06; 13.9]		1.00
	sd(T2MI Troponin First Measure)	0.338 [0.01; 1.71]		1.01
	cor(T2MI Intercept,muT2MI Troponin First Measure)	-0.0674 [-0.84; 0.78]		1.00
Non-varying effects				
Group	*Variable*	*Estimate [Credible Interval (C.I.)]*	*P[C.I.][a]*	*Rhat*
Acute	Intercept	-22.3456 [-33.93; -10.71]	*	1.00
	Log Troponin First Measure	5.2511 [3.35; 7.11]	*	1.01
	Troponin Difference	0.1238 [0.07; 0.18]	*	1.00
	Physiological Platec	0.6983 [-0.69; 2.13]		1.00
	Angiogram	0.3307 [-1.19; 1.92]		1.00
	Gender	0.123 [-0.74; 1.02]		1.00
	Physiological Creatine	0.8309 [-0.48; 2.26]		1.00
Chronic	Intercept	-18.099 [-28.15; -7.97]	*	1.00
	Log Troponin First Measure	6.795 [4.82; 8.51]	*	1.00
	Troponin Difference	0.0772 [0.02; 0.14]	*	1.00
	Physiological Platec	-0.7597 [-1.98; 0.39]		1.00
	Angiogram	-0.3363 [-1.8; 1.21]		1.00
	Gender	0.5943 [-0.15; 1.33]		1.00
	Physiological Creatine	0.7504 [-0.45; 2.05]		1.00
T1MI	Intercept	-8.773 [-23.98; 6.5]		1.00
	Log Troponin First Measure	3.8949 [1.83; 5.72]	*	1.00
	Troponin Difference	0.1266 [0.07; 0.18]	*	1.00
	Physiological Platec	-0.7263 [-2.74; 1.25]		1.00
	Angiogram	4.1798 [2.62; 5.86]	*	1.00
	Gender	-0.9164 [-2.42; 0.47]		1.00
	Physiological Creatine	-0.79 [-2.89; 1.3]		1.00
T2MI	Intercept	-50.8471 [-105.44; -6.42]	*	1.00
	Log Troponin First Measure	11.3029 [5.19; 21.31]	*	1.00
	Troponin Difference	0.1439 [0.08; 0.21]	*	1.00
	Physiological Platec	1.9123 [-3.41; 7.98]		1.00
	Angiogram	-0.2741 [-4.49; 3.63]		1.00
	Gender	7.2632 [0.48; 20.04]	*	1.00
	Physiological Creatine	-1.1107 [-6.72; 3.15]		1.00
Performance criteria		***Model settings***		
LOOCV	652.6	**Priors**	student_t(3,0,5)	
Multivariate Rhat	1.04		lkj(2)	
		Reference category	Normal	
Significance level	0.05	**Iterations**	2000	
		Chains	4	
		Cores	4	
		Adapt_delta	0.96	
		Max_treedepth	12	

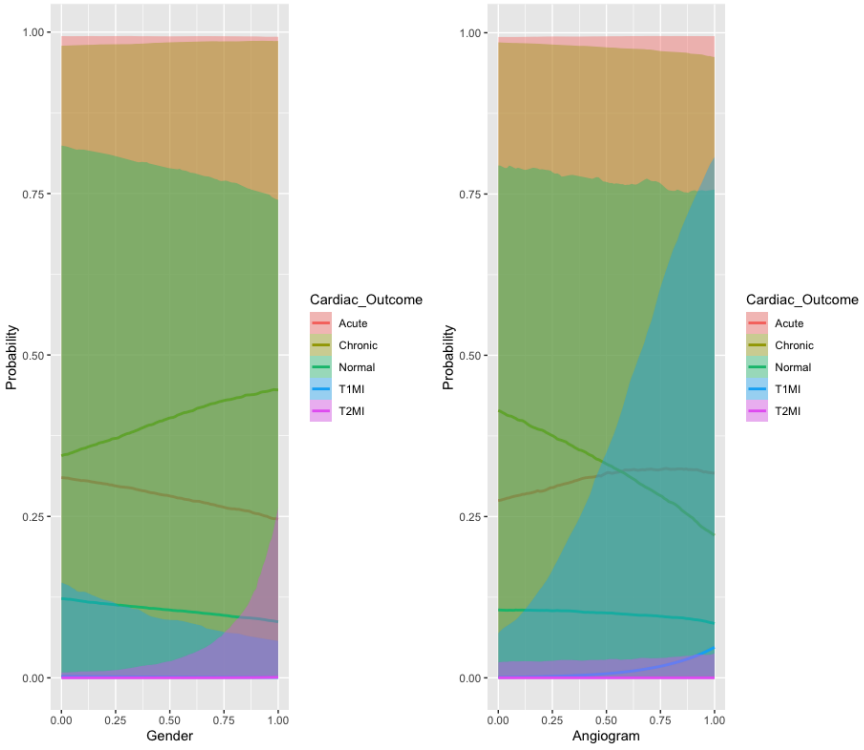
Table 4: Results model 1, significance level of 5%. See formula 18 for the DGP

[a]C.I. stands for credible interval, not confidence interval

Moreover, the marginal effects of Gender and Angiogram are evaluated. In figure 11 the marginal effects for Gender and Angiogram in model 1 are presented, respectively. Subfigure 11a can be read as follows: the effect size of Gender on the probability of a Cardiac Outcome is presented (on the y-axis) for different values of Gender. For example, the effect of being female (Gender = 0) on the probability of having a T2MI is significantly lower than the effect of being male (Gender = 1). The figures can solely be interpreted at X = 0 or X = 1 because both Gender and Angiogram cannot have values between 0 and 1. Only significant effects can be interpreted and it can be seen from the figures that the non-significant effects are stable when the X-value increases from 0 to 1 (e.g. the effect of Gender on the Normal Cardiac Outcome changes only slightly). Since Gender solely has a significant effect on the T2MI Cardiac Outcome in model 1 (see table 4), it is the only marginal effect that is evaluated here. As can be seen in subfigure 11a, the probability of a T2MI increases when Gender goes from 0 to 1, according to the results of model 1. In other words, if two patients share the exact same values for all modelled variables except Gender, the male patient will have a higher probability on a T2MI. [6]

Angiogram only has a significant effect on T1MI in model 1 (see table 4) and thus that is the only effect that is discussed here. As can be seen in subfigure 11b, the probability of a T1MI goes up substantially if Angiogram goes from 0 to 1, according to the results of model 1. In other words, when evaluating two patients with exactly the same characteristics except for the Angiogram, the patient with an Angiogram will have a substantially higher probability of having a T1MI, as opposed to the patient without the Angiogram.

[6] It has to be noted that this does not imply that surgically changing a patients' gender will affect its probability of a certain Cardiac Outcome



(a) Gender (b) Angiogram

Figure 11: Marginal effects of Gender and Angiogram in model 1

5.2.2 Prediction

In figure 12 the distribution of the predicted values based on the results of model 1 is presented. Moreover, the confusion matrix of model 1 is presented in table 5, along with the prediction accuracy values. Besides, the sensitivity and specificity of category T1MI are presented in table 5. As can be seen in both figure 12 and table 5, predictions of the categories Acute and T2MI are relatively few based on the results of model 1. The overall prediction accuracy of model 1 is 0.9077. The sensitivity of class T1MI is solely 0.7751, which is relatively low considering the consequences of discharging a patient who can potentially get a myocardial infarction. The specificity of class T1MI is 0.9823, this implies that the model does relatively well on detecting patients without a heart attack. As mentioned in section 1, the costs of admittance of cardiac patients to the ED are very high and thus properly discharging a relatively big proportion of cardiac patients can be very cost efficient.

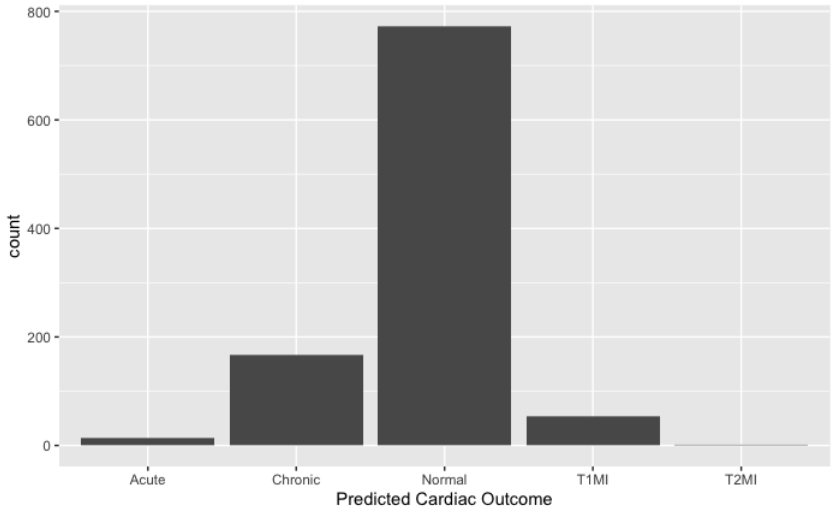


Figure 12: Predictions of Cardiac Outcome of model 1

		Observed				
		Acute	*Chronic*	*Normal*	*T1MI*	*T2MI*
Predicted	*Acute*	8	2	1	2	1
	Chronic	27	127	5	7	1
	Normal	10	17	742	3	0
	T1MI	9	6	2	37	0
	T2MI	0	0	0	0	1

Overal accuracy	**Sensitivity class T1MI**	**Specificity class T1MI**
0.9077	0.7551	0.9823

Table 5: Confusion matrix model 1

5.3 Model 2

5.3.1 Model output

In table 6 the output of the second model (see formula 19 for the DGP) is presented. As can be seen, the output contains the parameter mi(Troponin Difference) instead of Troponin Difference (as in table 4). mi(Troponin Difference) contains both observed and imputed values. The values are imputed in the sub-model, the results for this model are also presented in 6, below the non-varying effects. Furthermore, the performance criteria and model settings are presented. At last, the STAN code corresponding to model 2 is presented in the Technical Appendix (see section 9).

As can be seen in table 6, all univariate R-hats are lower than 1.05, indicating convergence. Besides, the multivariate R-hat is below 1.10, which also indicates convergence. The LOOIC of model 2 is 693.2.

Unsurprisingly, there is a positive significant effect of the logarithm of Troponin First Measure on the probability of having a T1MI Cardiac Outcome. Similarly, positive significant effects exists of the logarithm of Troponin First Measure on the probability of having a Acute, Chronic or T2MI Cardiac Outcome, respectively. Moreover, a positive significant effects exists of mi(Troponin Difference) on the probability of having a T1MI Cardiac Outcome. This can be interpreted as follows: If a patients' Troponin Difference (i.e. Troponin Second Measure minus Troponin First Measure) increases by 1, then the probability of having a T1MI Cardiac Outcome increases by approximately 0.07%, ceteris paribus. Similarly, there are positive significant effects of mi(Troponin Difference) on the probability of having a Acute or T2MI Cardiac Outcome, respectively. However, the effect of mi(Troponin Difference) on the probability of having a Chronic Cardiac Outcome is not significant at 5%. Apart from that, a patient having had an Angiogram leads to an increase of the probability of having a T1MI of approximately 4.45%, ceteris paribus. At last, being male leads to an increase of the probability of having a T2MI of approximately 4.45%, ceteris paribus.

Since the link of the sub-model is Gaussian, the interpretation of the results is different. As can be seen in table 6, a positive effect exists of both Creatine and Anigogram on Troponin Difference. In other words, if the Creatine level of a patient goes up by 1, the Troponin Difference increases by approximately 17.75, ceteris paribus. Besides, a patient having had an Angiogram leads to an increase of the patients' Troponin Difference of approximately 33.28, ceteris paribus.

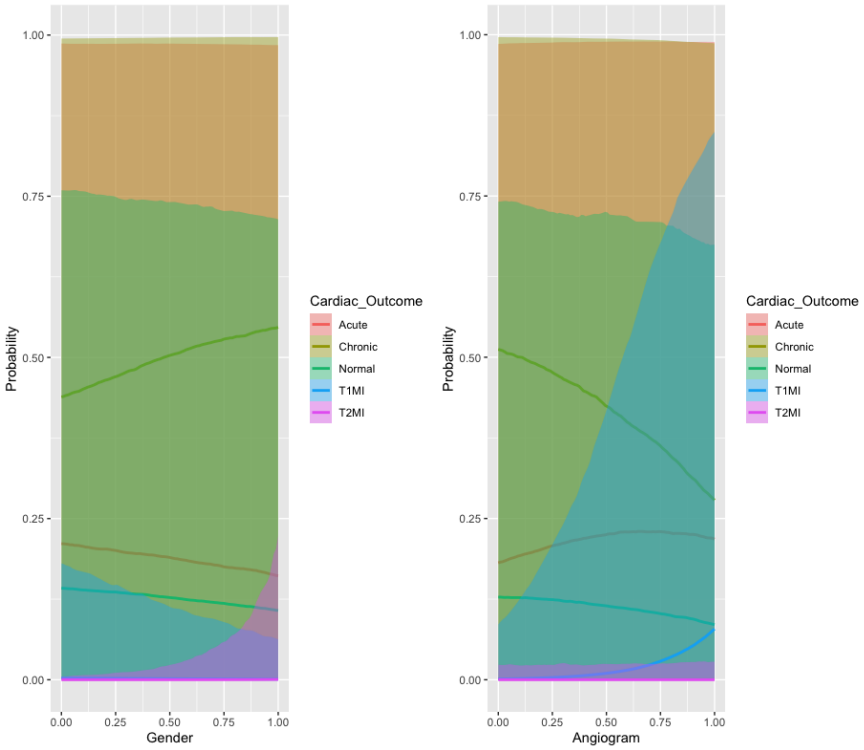
Varying effects				
Group	*Variable*	*Estimate [Credible Interval (C.I.)]*		*Rhat*
Acute	sd(Acute Intercept)	3.08 [0.26; 8.97]		1.00
	sd(Acute Troponin First Measure)	0.23 [0; 1.37]		1.00
	cor(Acute Intercept,muAcute Troponin First Measure)	-0.05 [-0.83; 0.77]		1.00
Chronic	sd(Chronic Intercept)	2.2 [0.07; 7.46]		1.00
	sd(Chronic Troponin First Measure)	0.21 [0; 1.31]		1.00
	cor(Chronic Intercept,muChronic Troponin First Measure)	-0.04 [-0.85; 0.79]		1.00
T1MI	sd(T1MI Intercept)	2.09 [0.05; 7.25]		1.00
	sd(T1MI Troponin First Measure)	0.33 [0.03; 1.53]		1.00
	cor(T1MI Intercept,muT1MI Troponin First Measure)	-0.02 [-0.84; 0.82]		1.00
T2MI	sd(T2MI Intercept)	3.26 [0.09; 12.74]		1.00
	sd(T2MI Troponin First Measure)	0.36 [0.01; 1.73]		1.00
	cor(T2MI Intercept,muT2MI Troponin First Measure)	-0.08 [-0.84; 0.78]		1.00
Non-varying effects				
Group	*Variable*	*Estimate [Credible Interval (C.I.)]*	*P[CI][a]*	*Rhat*
Acute	Intercept	-24.72 [-36.8; -13.54]	*	1.00
	Log Troponin First Measure	5.81 [3.61; 7.57]	*	1.00
	Physiological Platec	0.72 [-0.61; 2.06]		1.00
	Angiogram	0.62 [-0.87; 2.12]		1.00
	Gender	0.01 [-0.83; 0.82]		1.00
	Physiological Creatine	1.01 [-0.34; 2.35]		1.00
	mi(Troponin Difference)	0.07 [0.03; 0.11]	*	1.00
Chronic	Intercept	-20.21 [-30.34; -9.93]	*	1.00
	Log Troponin First Measure	7.35 [5.13; 9.07]	*	1.00
	Physiological Platec	-0.56 [-1.7; 0.56]		1.00
	Angiogram	-0.16 [-1.65; 1.32]		1.00
	Gender	0.46 [-0.21; 1.15]		1.00
	Physiological Creatine	0.85 [-0.32; 2.03]		1.00
	mi(Troponin Difference)	0.02 [-0.01; 0.07]		1.00
T1MI	Intercept	-14.67 [-29.23; -0.53]	*	1.00
	Log Troponin First Measure	4.49 [2.22; 6.36]	*	1.00
	Physiological Platec	-0.22 [-2.07; 1.64]		1.00
	Angiogram	4.45 [2.91; 6.2]	*	1.00
	Gender	-0.91 [-2.3; 0.43]		1.00
	Physiological Creatine	-0.28 [-2.3; 1.62]		1.00
	mi(Troponin Difference)	0.07 [0.04; 0.11]	*	1.00
T2MI	Intercept	-48.94 [-98.93; -8.57]	*	1.00
	Log Troponin First Measure	11.3 [5.32; 21.14]	*	1.00
	Physiological Platec	1.78 [-3.13; 7.58]		1.00
	Angiogram	-0.12 [-4.41; 3.8]		1.00
	Gender	7.14 [0.5; 20.84]	*	1.00
	Physiological Creatine	-1.31 [-7.26; 3.05]		1.00
	mi(Troponin Difference)	0.09 [0.04; 0.14]	*	1.00
Sub-model				
Y-Variable	*Variable*	*Estimate*	*P[C.I.][b]*	*Rhat*
Troponin Difference	Intercept	-74.03 [-124.92; -22.2]	*	1.00
	Physiological Creatine	17.75 [9.55; 25.5]	*	1.00
	Physiological Platec	-1.02 [-7.32; 5.26]		1.00
	Gender	3.93 [-0.47; 8.37]		1.00
	Angiogram	33.28 [25.43; 41.2]	*	1.00
	Smoking	0.27 [-1.98; 2.58]		1.00
Performance criteria		***Model settings***		
LOOCV	693.2	Priors		student_t(3,0,5)
Multivariate Rhat	1.06			lkj(2)
		Links		categorical
Sginificance level	0.05			Gaussian
		Reference category		Normal
		Iterations		2000
		Chains		4
		Cores		4
		Adapt_delta		0.96
		Max_treedepth		12

Table 6: Results model 2, significance level of 5%. See formula 19 for the DGP

[a] C.I. stands for credible interval, not confidence interval

[b] C.I. stands for credible interval, not confidence interval

Moreover, the marginal effects of Gender and Angiogram are evaluated. In figure 13 the marginal effects for Gender and Angiogram in model 2 are presented, respectively. The interpretation of theses figures is the same as the interpretation of figure 11. Whereas Gender only has a significant effect on T2MI in model 2, Angiogram solely has a significant effect on T2MI in model 2. As can be seen when comparing subfigure 13 and 11, the interpretation of the marginal effects is very similar.



(a) Gender (b) Angiogram

Figure 13: Marginal effects of Gender and Angiogram in model 2

5.3.2 Prediction

In figure 14 the distribution of the predicted outcomes based on model is presented. When compared to the distribution of the observed values (see figure 1), there are relatively few predicted values for the Acute and T2MI Cardiac Outcome, respectively. In table 7 the confusion matrix for the predictions of model 2 is presented. Moreover, the accuracy, the sensitivity of class T1MI and the specificity of class T1MI are presented. Model 2 predicts relatively few Acute and T2MI Cardiac Outcomes, in comparison the the distribution of the observed values (see figure 1). The overall accuracy of model 2 is 0.9183, as can be seen in table 7. The sensitivity of class T1MI is 0.7400, which is relatively low given the consequences of a false negative. The specificity of class T1MI is relatively high, namely, 0.9846.

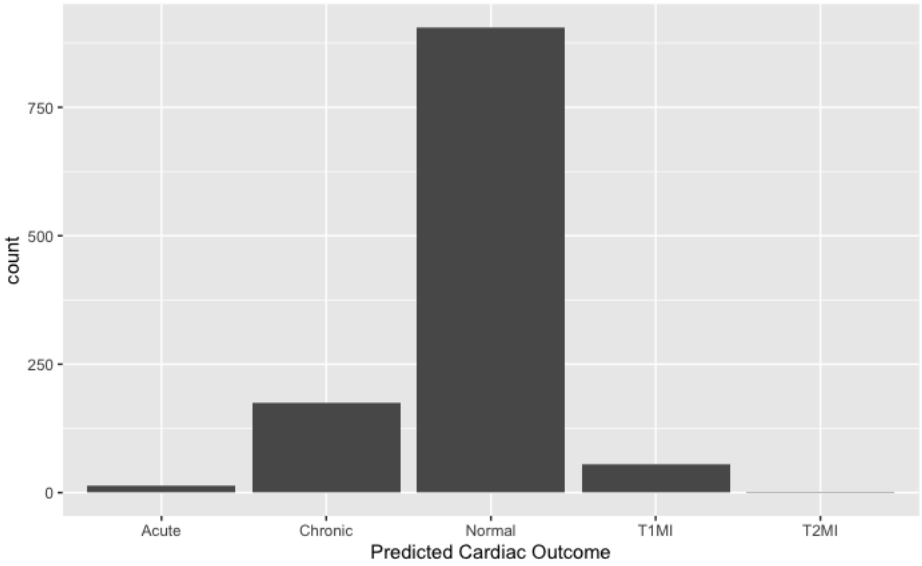


Figure 14: Predictions of Cardiac Outcome of model 2

		Observed				
		Acute	*Chronic*	*Normal*	*T1MI*	*T2MI*
Predicted	*Acute*	7	2	0	3	1
	Chronic	28	136	5	7	1
	Normal	9	18	876	3	0
	T1MI	10	6	1	37	0
	T2MI	0	0	0	0	1
Overall accuracy			**Sensitivity class T1MI**		**Specificity class T1MI**	
0.9183			0.7400		0.9846	

Table 7: Confusion matrix model 2

5.4 Sensitivity analysis

In table 11 and table 12 the results of the sensitivity analysis for model 1 and model 2, respectively, are presented. For both models, the prior has been changed to a $student_t(7, 0, 5)$ prior for the non-varying effects and $lkj(4)$ for the correlation matrix. Originally, the priors were $student_t(3, 0, 5)$ and $lkj(2)$, respectively.

5.4.1 Model 1

In terms of the LOOCV, the adjustment of the priors has not influenced the results of model 1 (difference $< 1\%$). With the robustness check the LOOCV is 654.1, whereas the value is 652.6 in the original model (see table 11). The model still convergences after adjustment of the priors (Multivariate Rhat < 1.10 & univariate Rhats < 1.05). However, the significance of the non-varying effect parameters changes. Specifically, the significance of the effect of the logarithm of Troponin

First Measure on the probability on having a Chronic or T1MI Cardiac Outcome, respectively, disappears. As previously mentioned in section 4, $lkj(\eta)$ priors with a high value for η assume few correlation between the parameters. Most likely, this assumption is not realistic for the likelihood specification of model 1 and thus adjusting the value of η distorts the significance of the results.

5.4.2 Model 2

For model 2, the adjustment of the priors has not influenced the LOOCV. The original value of the LOOCV was 693.2, whereas the value is 690.3 after adjustment of the priors (see table 12). Even though the model performance increases on paper due to the decrease of the LOOIC, the difference is so small that it can be considered indecisive (difference <1%). Model 2 still converges after adjustment of the priors (Multivariate Rhat <1.10 univariate Rhats <1.05)). Moreover, the sign and significance of the non-varying effect parameter (most importantly, the Troponin variables) remain the same after adjustment of the priors. In conclusion, model 2 is robust for changes in the model settings.

5.5 Model comparison

The LOOIC of model 1 is 652.6, whereas the LOOIC of model 2 is 693.2. The model with the lowest LOOIC is preferred. Hence, a strong indication exists for a better model performance of model 1.

The sign and significance of the non-varying effects of model 2 are almost the same as model 1. However, in model 2 no significant effect exists of mi(Troponin Difference) on the Cardiac Outcome 'Chronic', whereas a positive significant was expected. However, differences in the effect size exist. As can be seen in table 6, the effect of log Troponin First Measure on T1MI is 4.49, whereas this effect is 3.89 in model 1 (see table 4). Moreover, the effect of mi(Troponin Difference) on T1MI in model 2 is 0.07, whereas this effect is 0.13 in model 1. In other words, under the assumptions of model 1 the effect of an increase of Troponin Difference by 1 leads to an increase in the probability of having a T1MI of 0.13%, whereas this is 0.07% under the assumptions of model 2. Similar differences in effect size exist for other parameters. However, one cannot state which one of these effect sizes is the best due to lack of similar research in the field.

In terms of overall prediction accuracy, model 2 performs better compared to model 1 (0.9183 vs 0.9077). Besides, the specificity of class T1MI is slightly higher for model 2 compared to model 1 (0.9846 vs 0.9823). However, the sensitivity of class T1MI for model 1 is better than for model 2 (0.7551 vs 0.7400).

Whereas model 1 is sensitive to adjustments of the priors, model 2 is robust to adjustment of the priors. In the sensitivity analysis for model 1 the significance of the variables (of interest) was influenced by adjustments of the priors, whereas the significance was unaffected in model 2.

6 Conclusion

The aim of this paper is twofold. Firstly, causal structures in cardiac patient data, of those admitted to the ED in South Australia, were studied by the use of Bayesian multilevel models. Specifically, the Bayesian Generalized Linear Mixed Model (GLMM) was used. Secondly, the Fully Bayesian approach as solution to missingness of data was explored. The contribution of this paper is, first of all, the fact that causal structures of patient characteristics on the Cardiac Outcome of cardiac patients, admitted to the ED, were described. This was done despite (cardiac-) patient data being very complex by nature. Specifically, significant effects of the Troponin variables on the probability of having a myocardial infarction (T1MI Cardiac Outcome) were found. Moreover, this paper made a first attempt on using the Fully Bayesian approach as solution to missingness of cardiac patient data. Even though it did not substantially improve (nor impair) the model performance, the Fully Bayesian approach seems promising in this context. Despite the limitations of the probabilistic programming language used in this paper (STAN), the model performance amongst the models was similar. Hence, there is a strong indication that the Fully Bayesian approach would have performed substantially better if these limitations had not existed (see section 7 for more on this).

In comparison, model 1 performed better than model 2 based on the LOOIC. The sign and significance on the non-varying effects of model 2 are almost the same as model 1. However, some differences in the effect size existed. Though, it cannot be stated that one model performs best in terms of effect size due to lack of comparable results in the literature. Whereas model 2 performs better in terms of overall prediction accuracy, model 1 performs better in terms of sensitivity. The specificity was highly similar amongst models. Whereas the Fully Bayesian approach, as solution to missingness of data, led to a model robust to adjustment of priors, model 1 was sensitive to adjustments of the priors and and did therefore not pass the sensitivity analysis.

7 Discussion

Even though modelling cardiac patient data with Bayesian GLMMs has its advantages, there are some major drawbacks that hinder the ease of experimentation with this technique. The likelihood specifications used in this paper were found by using the 'modular construction' method, as proposed by Gelman et al. (2020) in their 'Bayesian Workflow' paper. Even though an optimum is found after experimenting with many likelihood specifications, this optimum is most likely a local optimum and not a global optimum. Specifically, there is most likely a combination of priors and explanatory variables that results in better model performance/results. However, attaining this global optimum was not possible within the timeline of my thesis and with the available computation power. Though, this is interesting for future research on this topic. In the context of the RAPIDx AI project, the RAPIDx AI dataset can be modelled by the use of Bayesian GLMMs with more extensive experimentation with likelihood specifications and choice of priors. Consequently, the model performance will improve, compared to my model performance. In a wider context, the

same can be done on (cardiac-) patient data. Since the computational power has been increasing exponentially, it becomes more and more interesting for researchers to use Bayesian models.

Apart from more experimentation with the variable selection method mentioned above, future research can focus on using Bayesian Lasso, the Horseshoe prior or Stochastic Search Variable Selection in the context of patient data. Namely, using these variable selection methods could have led to improved model performance but experimenting with all Bayesian variable selection methods was outside of the scope/timeline of this study. Besides, a comparison of all four methods in this context would be interesting. Due to limited time and computational power for this thesis, a choice had to be made for the variable selection method and thus only the 'modular construction' method, as proposed by (Gelman et al., 2020), was explored.

Apart from the drawback regarding the variable and prior selection, the probabilistic language/software engine used in this research, STAN, poses some serious drawbacks. Firstly, discrete parameters cannot be estimated in STAN. Despite $ECG\ Ischaemia$ being (by far) the best variable to use as varying-effect parameter, this leads to a relatively big proportion (about two-third) of the observations being dropped when including it in the likelihood specification. Hence, it would be interesting for future research to find a way to still estimate $ECG\ Ischaemia$. In other words, it could greatly improve model performance if researchers would find a way to impute the variable with the Fully Bayesian approach (i.e. in a sub-model). Possible solutions for this would be marginalizing discrete parameters (Yackulic et al., 2020) or the use of Discontinuous Hamiltonian Monte Carlo as sampling algorithm (Nishimura et al., 2017).

Moreover, linear transformations of the imputed variable are currently not possible in the 'brms' package. As described in the methodology 4, it would be optimal to impute *Troponin Second Measure* and calculate *Troponin Difference* based on the difference between *Troponin Second Measure* (observed imputed) and *Troponin First Measure* (observed). However, this linear transformation (i.e. calculating the difference) cannot be done directly with the 'brms' package and thus a sub-optimal solution was used. Even consulting the 'brms' forum on Discourse.org, on which Paul Bürkner himself is very active, did not result into a proper solution (within the scope of my thesis). However, the abovementioned linear transformation could be conducted by writing STAN code. Unfortunately, this was beyond the scope of this thesis. Yet, it would be interesting for further researchers to explore this solution and to model cardiac patient data in the optimal way.

8 Appendix

Variable	Description	Expected effect on myocardial injury	Reference expected effect
Cardiac Outcome	Classification of the clinical outcome, the categories are: Normal, Type 1 Mim Type 2 MI, Acute & Chronic	NA	NA
Length Of Stay	Duration of the ED visit in minutes	NA	NA
Troponin First Measure	First measurement of the troponin level, measured in ng/L	+	Members et al. (2014)
Troponin Second Measure	Second measurement of the troponin level, measured in ng/L	+	Members et al. (2014)
Troponin Difference	Difference between the troponin measurements, estimation of the velocity	+	Members et al. (2014)
Carbon Dioxide Pressure	Co2 level in a patients body	+	Atkinson et al. (2017)
Lactate	Measurement of the amount of lactic acid in the blood	+	Demers et al. (2000)
Dioxide Pressure	CO2 in a patients body	+	Atkinson et al. (2017)
Pulmonary Hypertension	Pulmonary artety systolic pressure in mm Hg	+	Members et al. (2014)
Brain Natriuretic Peptide	Brain Natriuretic Peptide in pg/L	+	Members et al. (2014)
Creatine Kinase MB	Creatine kinase myocardial band in IU/L	+	Levy et al. (2011)
Fibrin	Fibrinogen in mg/L	+	Members et al. (2014)
Urea	Urea nitrogen in mmol/L	-	Members et al. (2014)
Creatine	Creatine level	+	Members et al. (2014)
Upper Respiratory Tract	Upper respiratory tract level	+	Ruane et al. (2017)
Albumin	Albumin level in g/L	-	Kuller et al. (1991)
Haeglob	Hemoglobin level in g/L	+	Kim et al. (2013)
White Blood Cell Count	White blood cell count per cubib microliter of blood	+	Lee et al. (2001)
Platec	Platelet count per microliter	+	Gregg and Goldschmidt-Clermont (2003)
Platev	Platelet volume in femtolitre	+	Gregg and Goldschmidt-Clermont (2003)
Hemoglobin A1C	Hemoglobin A1C in %	+	Kim et al. (2013)
Thyroid Stimulating Hormone	Thyroid Stimulating Hormone in mIU/L	+	**?**
C-Reactive Protein	C-Reactive Protein test result in mg/L	+	Collaboration et al. (2010)
Ferritin	Ferritin in ng/mL	+	Liu et al. (2019)
D-Dimer	D-Dimer in ng/L	+	Mansour et al. (2020)
Lactv	Blood lactate level in mmol/L	+	Demers et al. (2000)
Smoking	0 = Never, 1 = Past, 2 = Currently	+	Members et al. (2014)
Age	Age of patient at time of admittance, not rounded	+	Members et al. (2014)
Gender	1 = Male, 0 = Female	+	Gao et al. (2019)
Prior Heart Attack	1 = Yes, 0 = No	+	Members et al. (2014)
History Diabetes	1 = Yes, 0 = No	+	Everson-Rose and Lewis (2005)
History Hypertension	1 = Yes, 0 = No	+	Members et al. (2014)
History STD	1 = Yes, 0 = No	+	Mussa et al. (2006)
ECG Ischaemia	1 = Yes, 0 = No	+	Members et al. (2014)
Coranary Heart Disease	1 = Yes, 0 = No	+	Members et al. (2014)
Dyslipidemia	1 = Yes, 0 = No	+	Miller (2009)
Family History Coronary Disease	1 = Yes, 0 = No	+	**?**
Heart Rate	Heart rate in beats per minute	+	Perret-Guillaume et al. (2009)
Systolic Blood Pressure	Systolic Blood Pressure in mm Hg	+	Everson-Rose and Lewis (2005)
Diastolic Blood Pressure	Diastolic Blood Pressure in mm Hg	+	Everson-Rose and Lewis (2005)
Kidney Function	Glomerular filtration rate	+	Sarnak et al. (2003)
Onset	Duration between onset of symptoms and admission: 1 = less than 1 hr, 2 = 1-3 hours, 3 = 4-6 hours, 4 = 6-12 hours, 5 = 12-24 hours, 6 = above 24 hours	+	Luepker et al. (2000)
Angiogram	1 = Yes, 0 = No	+ (minor effect)	Tavakol et al. (2012)
History Angiogram	1 = Yes, 0 = No	+ (minor effect)	Tavakol et al. (2012)

Table 8: Data dictionary of all variables, including expected effect on the chance of a myocardial injury

	NA	Min	Max	Range	Median	Mean	Variance
Carbon Dioxide Pressure	2604	14	137	123	43	44.49	14.14
Lactate	2660	0.3	23.79	23.49	1.5	2.22	2.24
Dioxide Pressure	2605	9	589	580	52	74.2	82.53
Pulmonary Hypertension	2604	1.01	7.64	6.63	7.39	7.36	0.24
Brain Natriuretic Peptide	3289	50	35000	34950	2348	5621.54	8382.91
Creatine Kinase MB	2318	0.3	1472	1471.7	2.7	7.54	49.98
Fibrin	3328	0.4	8.1	7.7	2.91	3.15	1.59
Urea	29	1.5	57.5	56	6.3	8.07	5.76
Creatine	30	30	1445	1415	82	101.93	80.72
Upper Respiratory Tract	2098	0.05	1.11	1.06	0.35	0.37	0.14
Albumin	141	14	54	40	37	36.38	5.13
Haeglob	43	50	207	157	136	133.69	20.13
White Blood Cell Count	42	0.12	259	258.88	8.2	9.47	6.84
Platec	59	6	2043	2037	235	243.59	87.04
Platev	2553	8.2	13.6	5.4	10.3	10.35	0.93
Hemoglobin A1C	2997	4.6	142	137.4	6.1	15.39	21.38
Thyroid Stimulating Hormone	2809	0.01	150	149.99	1.7	2.84	9.53
C-Reactive Protein	915	0.2	455.6	455.4	4.4	24.5	51.66
Ferritin	3235	5	2941	2936	98	205.62	362.61
D-Dimer	2951	0.2	80	79.8	0.5	2.74	9.17
Lactv	3168	0.5	14.7	14.2	1.7	2.12	1.68

Table 9: Descriptive statistics of the physiological variables

Variable	**Male**	**Female**	**NA**
Gender	1788 (52.77%)	1600 (47.23%)	0 (0%)
Variable	**Yes**	**No**	**NA**
History Diabetes	835	177	2376
Prior Heart Attack	831	179	2378
History Hypertension	497	515	2376
History STD	877	131	2380
ECG Ischaemia	1186	11	2191
Coranary Heart Disease	831	367	2190
Dyslipidemia	1125	1085	1178
Family History Coronary	355	554	2479
History Angiogram	930	82	2376
Angiogram	3032	356	0

Table 10: Descriptive statistics of the binary variables

Varying effects				
Group	*Variable*	*Estimate*		*Rhat*
Acute	sd(Acute Intercept)	3.1074 [0.4; 8.75]		1.00
	sd(Acute Troponin First Measure)	0.2612 [0.02; 1.37]		1.00
	cor(Acute Intercept,muAcute Troponin First Measure)	-0.055 [-0.66; 0.57]		1.00
Chronic	sd(Chronic Intercept)	1.9546 [0.07; 6.83]		1.00
	sd(Chronic Troponin First Measure)	0.2568 [0.01; 1.45]		1.00
	cor(Chronic Intercept,muChronic Troponin First Measure)	-0.0153 [-0.65; 0.62]		1.00
T1MI	sd(T1MI Intercept)	2.1146 [0.08; 7.62]		1.00
	sd(T1MI Troponin First Measure)	0.4194 [0.05; 2.33]		1.00
	cor(T1MI Intercept,muT1MI Troponin First Measure)	-0.0086 [-0.64; 0.62]		1.00
T2MI	sd(T2MI Intercept)	3.0271 [0.07; 13.24]		1.00
	sd(T2MI Troponin First Measure)	0.3004 [0.01; 1.53]		1.00
	cor(T2MI Intercept,muT2MI Troponin First Measure)	-0.0418 [-0.65; 0.6]		1.00
Non-varying effects				
Group	*Variable*	*Estimate*	*Star*	*Rhat*
Acute	Intercept	-22.1322 [-33.92; -10.55]	*	1.00
	Log Troponin First Measure	5.2355 [3.33; 7.01]	*	1.00
	Troponin Difference	0.1235 [0.07; 0.18]		1.01
	Physiological Platec	0.6696 [-0.7; 2.05]	*	1.00
	Angiogram	0.4026 [-1.15; 1.91]	*	1.00
	Gender	0.1199 [-0.77; 0.98]	*	1.00
	Physiological Creatine	0.8347 [-0.58; 2.25]		1.00
Chronic	Intercept	-17.995 [-27.82; -8.78]		1.00
	Log Troponin First Measure	6.7675 [4.82; 8.56]		1.00
	Troponin Difference	0.0771 [0.02; 0.14]		1.01
	Physiological Platec	-0.7848 [-2.04; 0.35]	*	1.00
	Angiogram	-0.2606 [-1.76; 1.2]	*	1.00
	Gender	0.5919 [-0.15; 1.32]		1.00
	Physiological Creatine	0.7616 [-0.46; 2.06]		1.00
T1MI	Intercept	-8.5847 [-23.78; 6.9]		1.00
	Log Troponin First Measure	3.9254 [1.9; 5.76]		1.00
	Troponin Difference	0.1263 [0.08; 0.18]	*	1.01
	Physiological Platec	-0.7631 [-2.7; 1.19]	*	1.00
	Angiogram	4.2293 [2.62; 5.92]		1.00
	Gender	-0.8946 [-2.31; 0.53]	*	1.00
	Physiological Creatine	-0.7791 [-3.03; 1.29]		1.00
T2MI	Intercept	-46.0778 [-93.25; -3.8]		1.00
	Log Troponin First Measure	10.0618 [4.77; 17.23]	*	1.00
	Troponin Difference	0.1436 [0.08; 0.21]	*	1.01
	Physiological Platec	1.9093 [-3.02; 7.74]		1.00
	Angiogram	-0.0424 [-4.26; 4.03]		1.00
	Gender	6.0442 [0.32; 14.66]	*	1.00
	Physiological Creatine	-1.114 [-6.68; 3.23]		1.00
Performance criteria		***Model settings***		
LOOCV	654.1	**Priors**		student_t(7,0,5)
Multivariate Rhat	1.02			lkj(4)
		Link		Categorical
Significance level	0.05	**Reference category**		Normal
		Iterations		2000
		Chains		4
		Cores		4
		Adapt_delta		0.96
		Max_treedepth		12

Table 11: Results sensitivity analysis model 1

Varying effects				
Group	*Variable*	*Estimate*		*Rhat*
Acute	sd(Acute Intercept)	2.8939 [0.28 ;8.73]		1.00
	sd(Acute Troponin First Measure)	0.2017 [0.01 ;1.22]		1.00
	cor(Acute Intercept,muAcute Troponin First Measure)	-0.033 [-0.64 ;0.61]		1.00
Chronic	sd(Chronic Intercept)	2.0698 [0.06 ;7.08]		1.00
	sd(Chronic Troponin First Measure)	0.2094 [0 ;1.34]		1.00
	cor(Chronic Intercept,muChronic Troponin First Measure)	-0.0257 [-0.64 ;0.62]		1.00
T1MI	sd(T1MI Intercept)	2.1113 [0.06 ;7.86]		1.00
	sd(T1MI Troponin First Measure)	0.3445 [0.03 ;1.61]		1.00
	cor(T1MI Intercept,muT1MI Troponin First Measure)	-0.0068 [-0.64 ;0.65]		1.00
T2MI	sd(T2MI Intercept)	3.451 [0.09 ;14.65]		1.00
	sd(T2MI Troponin First Measure)	0.3319 [0.02 ;1.63]		1.00
	cor(T2MI Intercept,muT2MI Troponin First Measure)	-0.0344 [-0.66 ;0.6]		1.00
Non-varying effects				
Group	*Variable*	*Estimate*	*Star*	*Rhat*
Acute	Intercept	-24.5234 [-36.06; -13.08]	*	1.00
	Log Troponin First Measure	5.8137 [3.75; 7.56]	*	1.00
	Physiological Platec	0.7078 [-0.62; 2.09]		1.00
	Angiogram	0.5878 [-0.91; 2.09]		1.00
	Gender	0.0092 [-0.86; 0.87]		1.00
	Physiological Creatine	1.0115 [-0.38; 2.34]		1.00
	m(iTroponin Difference)	0.067 [0.03; 0.11]	*	1.00
Chronic	Intercept	-20.2888 [-29.69; -10.44]	*	1.00
	Log Troponin First Measure	7.3398 [5.38; 8.96]	*	1.00
	Physiological Platec	-0.5671 [-1.65; 0.51]		1.00
	Angiogram	-0.1894 [-1.65; 1.26]		1.00
	Gender	0.4595 [-0.23; 1.17]		1.00
	Physiological Creatine	0.8663 [-0.32; 2.06]		1.00
	mi(Troponin Difference)	0.0255 [-0.01; 0.06]		1.00
T1MI	Intercept	-14.7851 [-29.79; -0.22]	*	1.00
	Log Troponin First Measure	4.4681 [2.27; 6.35]	*	1.00
	Physiological Platec	-0.2359 [-2.14; 1.62]		1.00
	Angiogram	4.4251 [2.89; 6.08]	*	1.00
	Gender	-0.8952 [-2.22; 0.43]		1.00
	Physiological Creatine	-0.2658 [-2.35; 1.75]		1.00
	mi(Troponin Difference)	0.0698 [0.04; 0.11]	*	1.00
T2MI	Intercept	-46.0122 [-93.36; -4.44]	*	1.00
	Log Troponin First Measure	10.4648 [4.98; 17.86]	*	1.00
	Physiological Platec	1.8258 [-3.14; 7.38]		1.00
	Angiogram	-0.1358 [-4.38; 3.68]		1.00
	Gender	5.9787 [0.09; 15.35]	*	1.00
	Physiological Creatine	-1.3627 [-6.92; 2.99]		1.00
	mi(Troponin Difference)	0.0875 [0.04; 0.14]	*	1.00
Sub-model				
Y-Variable	*Variable*	*Estimate*	*Star*	*Rhat*
Troponin Difference	Intercept	-67.481 [-117.22; -18.32]	*	1.00
	Physiological_Creatine	16.2408 [8.77; 23.96]	*	1.00
	Physiological_Platec	-0.9442 [-6.91; 5.17]		1.00
	Gender	3.4936 [-0.83; 7.76]		1.00
	Angiogram	31.732 [23.89; 39.72]	*	1.00
	Smoking	0.3189 [-2.01; 2.66]		1.00
Performance criteria		***Model settings***		
LOOCV	690.3	**Priors**		student_t(7,0,5)
Multivariate Rhat	1.05			lkj(4)
		Links		categorical
Sginificance level	0.05			Gaussian
		Reference category		Normal
		Iterations		2000
		Chains		4
		Cores		4
		Adapt_delta		0.96
		Max_treedepth		12

Table 12: Results sensitivity analysis model 2

9 Technical Appendix

Listing 1: Stan code model 1

```
// generated with brms 2.14.0
functions {
  /* turn a vector into a matrix of defined dimension
   * stores elements in row major order
   * Args:
   *   X: a vector
   *   N: first dimension of the desired matrix
   *   K: second dimension of the desired matrix
   * Returns:
   *   a matrix of dimension N x K
   */
  matrix as_matrix(vector X, int N, int K) {
    matrix[N, K] Y;
    for (i in 1:N) {
      Y[i] = to_row_vector(X[((i - 1) * K + 1):(i * K)]);
    }
    return Y;
  }
 /* compute correlated group-level effects
  * Args:
  *   z: matrix of unscaled group-level effects
  *   SD: vector of standard deviation parameters
  *   L: cholesky factor correlation matrix
  * Returns:
  *   matrix of scaled group-level effects
  */
  matrix scale_r_cor(matrix z, vector SD, matrix L) {
    // r is stored in another dimension order than z
    return transpose(diag_pre_multiply(SD, L) * z);
  }
}
data {
  int<lower=1> N;  // total number of observations
  int<lower=2> ncat;  // number of categories
  int Y[N];  // response variable
  int<lower=1> K_muAcute;  // number of population-level effects
```

```
matrix[N, K_muAcute] X_muAcute;  // population-level design matrix
int<lower=1> K_muChronic;  // number of population-level effects
matrix[N, K_muChronic] X_muChronic;  // population-level design matrix
int<lower=1> K_muT1MI;  // number of population-level effects
matrix[N, K_muT1MI] X_muT1MI;  // population-level design matrix
int<lower=1> K_muT2MI;  // number of population-level effects
matrix[N, K_muT2MI] X_muT2MI;  // population-level design matrix
// data for group-level effects of ID 1
int<lower=1> N_1;  // number of grouping levels
int<lower=1> M_1;  // number of coefficients per level
int<lower=1> J_1[N];  // grouping indicator per observation
// group-level predictor values
vector[N] Z_1_muAcute_1;
vector[N] Z_1_muAcute_2;
int<lower=1> NC_1;  // number of group-level correlations
// data for group-level effects of ID 2
int<lower=1> N_2;  // number of grouping levels
int<lower=1> M_2;  // number of coefficients per level
int<lower=1> J_2[N];  // grouping indicator per observation
// group-level predictor values
vector[N] Z_2_muChronic_1;
vector[N] Z_2_muChronic_2;
int<lower=1> NC_2;  // number of group-level correlations
// data for group-level effects of ID 3
int<lower=1> N_3;  // number of grouping levels
int<lower=1> M_3;  // number of coefficients per level
int<lower=1> J_3[N];  // grouping indicator per observation
// group-level predictor values
vector[N] Z_3_muT1MI_1;
vector[N] Z_3_muT1MI_2;
int<lower=1> NC_3;  // number of group-level correlations
// data for group-level effects of ID 4
int<lower=1> N_4;  // number of grouping levels
int<lower=1> M_4;  // number of coefficients per level
int<lower=1> J_4[N];  // grouping indicator per observation
// group-level predictor values
vector[N] Z_4_muT2MI_1;
vector[N] Z_4_muT2MI_2;
int<lower=1> NC_4;  // number of group-level correlations
```

```
  int prior_only;  // should the likelihood be ignored?
}
transformed data {
  int Kc_muAcute = K_muAcute - 1;
  matrix[N, Kc_muAcute] Xc_muAcute;  // centered version of X_muAcute without an in
  vector[Kc_muAcute] means_X_muAcute;  // column means of X_muAcute before centerin
  int Kc_muChronic = K_muChronic - 1;
  matrix[N, Kc_muChronic] Xc_muChronic;  // centered version of X_muChronic without
  vector[Kc_muChronic] means_X_muChronic;  // column means of X_muChronic before ce
  int Kc_muT1MI = K_muT1MI - 1;
  matrix[N, Kc_muT1MI] Xc_muT1MI;  // centered version of X_muT1MI without an inter
  vector[Kc_muT1MI] means_X_muT1MI;  // column means of X_muT1MI before centering
  int Kc_muT2MI = K_muT2MI - 1;
  matrix[N, Kc_muT2MI] Xc_muT2MI;  // centered version of X_muT2MI without an inter
  vector[Kc_muT2MI] means_X_muT2MI;  // column means of X_muT2MI before centering
  for (i in 2:K_muAcute) {
    means_X_muAcute[i - 1] = mean(X_muAcute[, i]);
    Xc_muAcute[, i - 1] = X_muAcute[, i] - means_X_muAcute[i - 1];
  }
  for (i in 2:K_muChronic) {
    means_X_muChronic[i - 1] = mean(X_muChronic[, i]);
    Xc_muChronic[, i - 1] = X_muChronic[, i] - means_X_muChronic[i - 1];
  }
  for (i in 2:K_muT1MI) {
    means_X_muT1MI[i - 1] = mean(X_muT1MI[, i]);
    Xc_muT1MI[, i - 1] = X_muT1MI[, i] - means_X_muT1MI[i - 1];
  }
  for (i in 2:K_muT2MI) {
    means_X_muT2MI[i - 1] = mean(X_muT2MI[, i]);
    Xc_muT2MI[, i - 1] = X_muT2MI[, i] - means_X_muT2MI[i - 1];
  }
}
parameters {
  vector[Kc_muAcute] b_muAcute;  // population-level effects
  real Intercept_muAcute;  // temporary intercept for centered predictors
  vector[Kc_muChronic] b_muChronic;  // population-level effects
  real Intercept_muChronic;  // temporary intercept for centered predictors
  vector[Kc_muT1MI] b_muT1MI;  // population-level effects
  real Intercept_muT1MI;  // temporary intercept for centered predictors
```

```
  vector[Kc_muT2MI] b_muT2MI;  // population-level effects
  real Intercept_muT2MI;  // temporary intercept for centered predictors
  vector<lower=0>[M_1] sd_1;  // group-level standard deviations
  matrix[M_1, N_1] z_1;  // standardized group-level effects
  cholesky_factor_corr[M_1] L_1;  // cholesky factor of correlation matrix
  vector<lower=0>[M_2] sd_2;  // group-level standard deviations
  matrix[M_2, N_2] z_2;  // standardized group-level effects
  cholesky_factor_corr[M_2] L_2;  // cholesky factor of correlation matrix
  vector<lower=0>[M_3] sd_3;  // group-level standard deviations
  matrix[M_3, N_3] z_3;  // standardized group-level effects
  cholesky_factor_corr[M_3] L_3;  // cholesky factor of correlation matrix
  vector<lower=0>[M_4] sd_4;  // group-level standard deviations
  matrix[M_4, N_4] z_4;  // standardized group-level effects
  cholesky_factor_corr[M_4] L_4;  // cholesky factor of correlation matrix
}
transformed parameters {
  matrix[N_1, M_1] r_1;  // actual group-level effects
  // using vectors speeds up indexing in loops
  vector[N_1] r_1_muAcute_1;
  vector[N_1] r_1_muAcute_2;
  matrix[N_2, M_2] r_2;  // actual group-level effects
  // using vectors speeds up indexing in loops
  vector[N_2] r_2_muChronic_1;
  vector[N_2] r_2_muChronic_2;
  matrix[N_3, M_3] r_3;  // actual group-level effects
  // using vectors speeds up indexing in loops
  vector[N_3] r_3_muT1MI_1;
  vector[N_3] r_3_muT1MI_2;
  matrix[N_4, M_4] r_4;  // actual group-level effects
  // using vectors speeds up indexing in loops
  vector[N_4] r_4_muT2MI_1;
  vector[N_4] r_4_muT2MI_2;
  // compute actual group-level effects
  r_1 = scale_r_cor(z_1, sd_1, L_1);
  r_1_muAcute_1 = r_1[, 1];
  r_1_muAcute_2 = r_1[, 2];
  // compute actual group-level effects
  r_2 = scale_r_cor(z_2, sd_2, L_2);
  r_2_muChronic_1 = r_2[, 1];
```

```
  r_2_muChronic_2 = r_2[, 2];
  // compute actual group-level effects
  r_3 = scale_r_cor(z_3, sd_3, L_3);
  r_3_muT1MI_1 = r_3[, 1];
  r_3_muT1MI_2 = r_3[, 2];
  // compute actual group-level effects
  r_4 = scale_r_cor(z_4, sd_4, L_4);
  r_4_muT2MI_1 = r_4[, 1];
  r_4_muT2MI_2 = r_4[, 2];
}
model {
  // likelihood including all constants
  if (!prior_only) {
    // initialize linear predictor term
    vector[N] muAcute = Intercept_muAcute + Xc_muAcute * b_muAcute;
    // initialize linear predictor term
    vector[N] muChronic = Intercept_muChronic + Xc_muChronic * b_muChronic;
    // initialize linear predictor term
    vector[N] muT1MI = Intercept_muT1MI + Xc_muT1MI * b_muT1MI;
    // initialize linear predictor term
    vector[N] muT2MI = Intercept_muT2MI + Xc_muT2MI * b_muT2MI;
    // linear predictor matrix
    vector[ncat] mu[N];
    for (n in 1:N) {
      // add more terms to the linear predictor
      muAcute[n] += r_1_muAcute_1[J_1[n]] * Z_1_muAcute_1[n] + r_1_muAcute_2[J_1[n]
    }
    for (n in 1:N) {
      // add more terms to the linear predictor
      muChronic[n] += r_2_muChronic_1[J_2[n]] * Z_2_muChronic_1[n] + r_2_muChronic_
    }
    for (n in 1:N) {
      // add more terms to the linear predictor
      muT1MI[n] += r_3_muT1MI_1[J_3[n]] * Z_3_muT1MI_1[n] + r_3_muT1MI_2[J_3[n]] *
    }
    for (n in 1:N) {
      // add more terms to the linear predictor
      muT2MI[n] += r_4_muT2MI_1[J_4[n]] * Z_4_muT2MI_1[n] + r_4_muT2MI_2[J_4[n]] *
    }
```

```
    for (n in 1:N) {
      mu[n] = transpose([muAcute[n], muChronic[n], 0, muT1MI[n], muT2MI[n]]);
    }
    for (n in 1:N) {
      target += categorical_logit_lpmf(Y[n] | mu[n]);
    }
  }
  // priors including all constants
  target += student_t_lpdf(b_muAcute | 3,0,5);
  target += student_t_lpdf(b_muChronic | 3,0,5);
  target += student_t_lpdf(b_muT1MI | 3,0,5);
  target += student_t_lpdf(b_muT2MI | 3,0,5);
  target += student_t_lpdf(sd_1 | 3, 0, 2.5)
    - 2 * student_t_lccdf(0 | 3, 0, 2.5);
  target += std_normal_lpdf(to_vector(z_1));
  target += lkj_corr_cholesky_lpdf(L_1 | 2);
  target += student_t_lpdf(sd_2 | 3, 0, 2.5)
    - 2 * student_t_lccdf(0 | 3, 0, 2.5);
  target += std_normal_lpdf(to_vector(z_2));
  target += lkj_corr_cholesky_lpdf(L_2 | 2);
  target += student_t_lpdf(sd_3 | 3, 0, 2.5)
    - 2 * student_t_lccdf(0 | 3, 0, 2.5);
  target += std_normal_lpdf(to_vector(z_3));
  target += lkj_corr_cholesky_lpdf(L_3 | 2);
  target += student_t_lpdf(sd_4 | 3, 0, 2.5)
    - 2 * student_t_lccdf(0 | 3, 0, 2.5);
  target += std_normal_lpdf(to_vector(z_4));
  target += lkj_corr_cholesky_lpdf(L_4 | 2);
}
generated quantities {
  // actual population-level intercept
  real b_muAcute_Intercept = Intercept_muAcute - dot_product(means_X_muAcute, b_muA
  // actual population-level intercept
  real b_muChronic_Intercept = Intercept_muChronic - dot_product(means_X_muChronic,
  // actual population-level intercept
  real b_muT1MI_Intercept = Intercept_muT1MI - dot_product(means_X_muT1MI, b_muT1MI
  // actual population-level intercept
  real b_muT2MI_Intercept = Intercept_muT2MI - dot_product(means_X_muT2MI, b_muT2MI
  // compute group-level correlations
```

```
  corr_matrix[M_1] Cor_1 = multiply_lower_tri_self_transpose(L_1);
  vector<lower=-1,upper=1>[NC_1] cor_1;
  // compute group-level correlations
  corr_matrix[M_2] Cor_2 = multiply_lower_tri_self_transpose(L_2);
  vector<lower=-1,upper=1>[NC_2] cor_2;
  // compute group-level correlations
  corr_matrix[M_3] Cor_3 = multiply_lower_tri_self_transpose(L_3);
  vector<lower=-1,upper=1>[NC_3] cor_3;
  // compute group-level correlations
  corr_matrix[M_4] Cor_4 = multiply_lower_tri_self_transpose(L_4);
  vector<lower=-1,upper=1>[NC_4] cor_4;
  // extract upper diagonal of correlation matrix
  for (k in 1:M_1) {
    for (j in 1:(k - 1)) {
      cor_1[choose(k - 1, 2) + j] = Cor_1[j, k];
    }
  }
  // extract upper diagonal of correlation matrix
  for (k in 1:M_2) {
    for (j in 1:(k - 1)) {
      cor_2[choose(k - 1, 2) + j] = Cor_2[j, k];
    }
  }
  // extract upper diagonal of correlation matrix
  for (k in 1:M_3) {
    for (j in 1:(k - 1)) {
      cor_3[choose(k - 1, 2) + j] = Cor_3[j, k];
    }
  }
  // extract upper diagonal of correlation matrix
  for (k in 1:M_4) {
    for (j in 1:(k - 1)) {
      cor_4[choose(k - 1, 2) + j] = Cor_4[j, k];
    }
  }
}
```

Listing 2: Stan code model 2

```
/ generated with brms 2.14.0
```

```
functions {
  /* turn a vector into a matrix of defined dimension
   * stores elements in row major order
   * Args:
   *   X: a vector
   *   N: first dimension of the desired matrix
   *   K: second dimension of the desired matrix
   * Returns:
   *   a matrix of dimension N x K
   */
  matrix as_matrix(vector X, int N, int K) {
    matrix[N, K] Y;
    for (i in 1:N) {
      Y[i] = to_row_vector(X[((i - 1) * K + 1):(i * K)]);
    }
    return Y;
  }
 /* compute correlated group-level effects
  * Args:
  *   z: matrix of unscaled group-level effects
  *   SD: vector of standard deviation parameters
  *   L: cholesky factor correlation matrix
  * Returns:
  *   matrix of scaled group-level effects
  */
  matrix scale_r_cor(matrix z, vector SD, matrix L) {
    // r is stored in another dimension order than z
    return transpose(diag_pre_multiply(SD, L) * z);
  }
}
data {
  int<lower=1> N;  // total number of observations
  int<lower=1> N_CardiacOutcome;  // number of observations
  int<lower=2> ncat_CardiacOutcome;  // number of categories
  int Y_CardiacOutcome[N_CardiacOutcome];  // response variable
  int<lower=1> K_muAcute_CardiacOutcome;  // number of population-level effects
  matrix[N_CardiacOutcome, K_muAcute_CardiacOutcome] X_muAcute_CardiacOutcome;
// population-level design matrix
  int<lower=1> Ksp_muAcute_CardiacOutcome;  // number of special effects terms
```

```
  int<lower=1> K_muChronic_CardiacOutcome;  // number of population-level effects
  matrix[N_CardiacOutcome, K_muChronic_CardiacOutcome] X_muChronic_CardiacOutcome;
// population-level design matrix
  int<lower=1> Ksp_muChronic_CardiacOutcome;  // number of special effects terms
  int<lower=1> K_muT1MI_CardiacOutcome;  // number of population-level effects
  matrix[N_CardiacOutcome, K_muT1MI_CardiacOutcome] X_muT1MI_CardiacOutcome;
// population-level design matrix
  int<lower=1> Ksp_muT1MI_CardiacOutcome;  // number of special effects terms
  int<lower=1> K_muT2MI_CardiacOutcome;  // number of population-level effects
  matrix[N_CardiacOutcome, K_muT2MI_CardiacOutcome] X_muT2MI_CardiacOutcome;
// population-level design matrix
  int<lower=1> Ksp_muT2MI_CardiacOutcome;  // number of special effects terms
  int<lower=1> N_TroponinDifference;  // number of observations
  vector[N_TroponinDifference] Y_TroponinDifference;  // response variable
  int<lower=0> Nmi_TroponinDifference;  // number of missings
  int<lower=1> Jmi_TroponinDifference[Nmi_TroponinDifference];  // positions of mis
  int<lower=1> K_TroponinDifference;  // number of population-level effects
  matrix[N_TroponinDifference, K_TroponinDifference] X_TroponinDifference;
// population-level design matrix
  // data for group-level effects of ID 1
  int<lower=1> N_1;  // number of grouping levels
  int<lower=1> M_1;  // number of coefficients per level
  int<lower=1> J_1_CardiacOutcome[N_CardiacOutcome];  // grouping indicator per obs
  // group-level predictor values
  vector[N_CardiacOutcome] Z_1_muAcute_CardiacOutcome_1;
  vector[N_CardiacOutcome] Z_1_muAcute_CardiacOutcome_2;
  int<lower=1> NC_1;  // number of group-level correlations
  // data for group-level effects of ID 2
  int<lower=1> N_2;  // number of grouping levels
  int<lower=1> M_2;  // number of coefficients per level
  int<lower=1> J_2_CardiacOutcome[N_CardiacOutcome];  // grouping indicator per obs
  // group-level predictor values
  vector[N_CardiacOutcome] Z_2_muChronic_CardiacOutcome_1;
  vector[N_CardiacOutcome] Z_2_muChronic_CardiacOutcome_2;
  int<lower=1> NC_2;  // number of group-level correlations
  // data for group-level effects of ID 3
  int<lower=1> N_3;  // number of grouping levels
  int<lower=1> M_3;  // number of coefficients per level
  int<lower=1> J_3_CardiacOutcome[N_CardiacOutcome];  // grouping indicator per obs
```

```
  // group-level predictor values
  vector[N_CardiacOutcome] Z_3_muT1MI_CardiacOutcome_1;
  vector[N_CardiacOutcome] Z_3_muT1MI_CardiacOutcome_2;
  int<lower=1> NC_3;  // number of group-level correlations
  // data for group-level effects of ID 4
  int<lower=1> N_4;  // number of grouping levels
  int<lower=1> M_4;  // number of coefficients per level
  int<lower=1> J_4_CardiacOutcome[N_CardiacOutcome];  // grouping indicator per obs
  // group-level predictor values
  vector[N_CardiacOutcome] Z_4_muT2MI_CardiacOutcome_1;
  vector[N_CardiacOutcome] Z_4_muT2MI_CardiacOutcome_2;
  int<lower=1> NC_4;  // number of group-level correlations
  int prior_only;  // should the likelihood be ignored?
}
transformed data {
  int Kc_muAcute_CardiacOutcome = K_muAcute_CardiacOutcome - 1;
  matrix[N_CardiacOutcome, Kc_muAcute_CardiacOutcome] Xc_muAcute_CardiacOutcome;
// centered version of X_muAcute_CardiacOutcome without an intercept
  vector[Kc_muAcute_CardiacOutcome] means_X_muAcute_CardiacOutcome;  // column mean
  int Kc_muChronic_CardiacOutcome = K_muChronic_CardiacOutcome - 1;
  matrix[N_CardiacOutcome, Kc_muChronic_CardiacOutcome] Xc_muChronic_CardiacOutcome
// centered version of X_muChronic_CardiacOutcome without an intercept
  vector[Kc_muChronic_CardiacOutcome] means_X_muChronic_CardiacOutcome;
// column means of X_muChronic_CardiacOutcome before centering
  int Kc_muT1MI_CardiacOutcome = K_muT1MI_CardiacOutcome - 1;
  matrix[N_CardiacOutcome, Kc_muT1MI_CardiacOutcome] Xc_muT1MI_CardiacOutcome;
// centered version of X_muT1MI_CardiacOutcome without an intercept
  vector[Kc_muT1MI_CardiacOutcome] means_X_muT1MI_CardiacOutcome;  // column means
  int Kc_muT2MI_CardiacOutcome = K_muT2MI_CardiacOutcome - 1;
  matrix[N_CardiacOutcome, Kc_muT2MI_CardiacOutcome] Xc_muT2MI_CardiacOutcome;
// centered version of X_muT2MI_CardiacOutcome without an intercept
  vector[Kc_muT2MI_CardiacOutcome] means_X_muT2MI_CardiacOutcome;  // column means
  int Kc_TroponinDifference = K_TroponinDifference - 1;
  matrix[N_TroponinDifference, Kc_TroponinDifference] Xc_TroponinDifference;
// centered version of X_TroponinDifference without an intercept
  vector[Kc_TroponinDifference] means_X_TroponinDifference;  // column means of X_
  for (i in 2:K_muAcute_CardiacOutcome) {
    means_X_muAcute_CardiacOutcome[i - 1] = mean(X_muAcute_CardiacOutcome[, i]);
    Xc_muAcute_CardiacOutcome[, i - 1] = X_muAcute_CardiacOutcome[, i] - means_X_m
```

```
  }
  for (i in 2:K_muChronic_CardiacOutcome) {
    means_X_muChronic_CardiacOutcome[i - 1] = mean(X_muChronic_CardiacOutcome[, i])
    Xc_muChronic_CardiacOutcome[, i - 1] = X_muChronic_CardiacOutcome[, i] - means_
  }
  for (i in 2:K_muT1MI_CardiacOutcome) {
    means_X_muT1MI_CardiacOutcome[i - 1] = mean(X_muT1MI_CardiacOutcome[, i]);
    Xc_muT1MI_CardiacOutcome[, i - 1] = X_muT1MI_CardiacOutcome[, i] - means_X_muT
  }
  for (i in 2:K_muT2MI_CardiacOutcome) {
    means_X_muT2MI_CardiacOutcome[i - 1] = mean(X_muT2MI_CardiacOutcome[, i]);
    Xc_muT2MI_CardiacOutcome[, i - 1] = X_muT2MI_CardiacOutcome[, i] - means_X_muT
  }
  for (i in 2:K_TroponinDifference) {
    means_X_TroponinDifference[i - 1] = mean(X_TroponinDifference[, i]);
    Xc_TroponinDifference[, i - 1] = X_TroponinDifference[, i] - means_X_TroponinI
  }
}
parameters {
  vector[Kc_muAcute_CardiacOutcome] b_muAcute_CardiacOutcome;  // population-level
  real Intercept_muAcute_CardiacOutcome;  // temporary intercept for centered predi
  vector[Ksp_muAcute_CardiacOutcome] bsp_muAcute_CardiacOutcome;  // special effect
  vector[Kc_muChronic_CardiacOutcome] b_muChronic_CardiacOutcome;  // population-le
  real Intercept_muChronic_CardiacOutcome;  // temporary intercept for centered pre
  vector[Ksp_muChronic_CardiacOutcome] bsp_muChronic_CardiacOutcome;  // special ef
  vector[Kc_muT1MI_CardiacOutcome] b_muT1MI_CardiacOutcome;  // population-level ef
  real Intercept_muT1MI_CardiacOutcome;  // temporary intercept for centered predic
  vector[Ksp_muT1MI_CardiacOutcome] bsp_muT1MI_CardiacOutcome;  // special effects
  vector[Kc_muT2MI_CardiacOutcome] b_muT2MI_CardiacOutcome;  // population-level ef
  real Intercept_muT2MI_CardiacOutcome;  // temporary intercept for centered predic
  vector[Ksp_muT2MI_CardiacOutcome] bsp_muT2MI_CardiacOutcome;  // special effects
  vector[Nmi_TroponinDifference] Ymi_TroponinDifference;  // estimated missings
  vector[Kc_TroponinDifference] b_TroponinDifference;  // population-level effects
  real Intercept_TroponinDifference;  // temporary intercept for centered predictor
  real<lower=0> sigma_TroponinDifference;  // residual SD
  vector<lower=0>[M_1] sd_1;  // group-level standard deviations
  matrix[M_1, N_1] z_1;  // standardized group-level effects
  cholesky_factor_corr[M_1] L_1;  // cholesky factor of correlation matrix
  vector<lower=0>[M_2] sd_2;  // group-level standard deviations
```

```
  matrix[M_2, N_2] z_2;  // standardized group-level effects
  cholesky_factor_corr[M_2] L_2;  // cholesky factor of correlation matrix
  vector<lower=0>[M_3] sd_3;  // group-level standard deviations
  matrix[M_3, N_3] z_3;  // standardized group-level effects
  cholesky_factor_corr[M_3] L_3;  // cholesky factor of correlation matrix
  vector<lower=0>[M_4] sd_4;  // group-level standard deviations
  matrix[M_4, N_4] z_4;  // standardized group-level effects
  cholesky_factor_corr[M_4] L_4;  // cholesky factor of correlation matrix
}
transformed parameters {
  matrix[N_1, M_1] r_1;  // actual group-level effects
  // using vectors speeds up indexing in loops
  vector[N_1] r_1_muAcute_CardiacOutcome_1;
  vector[N_1] r_1_muAcute_CardiacOutcome_2;
  matrix[N_2, M_2] r_2;  // actual group-level effects
  // using vectors speeds up indexing in loops
  vector[N_2] r_2_muChronic_CardiacOutcome_1;
  vector[N_2] r_2_muChronic_CardiacOutcome_2;
  matrix[N_3, M_3] r_3;  // actual group-level effects
  // using vectors speeds up indexing in loops
  vector[N_3] r_3_muT1MI_CardiacOutcome_1;
  vector[N_3] r_3_muT1MI_CardiacOutcome_2;
  matrix[N_4, M_4] r_4;  // actual group-level effects
  // using vectors speeds up indexing in loops
  vector[N_4] r_4_muT2MI_CardiacOutcome_1;
  vector[N_4] r_4_muT2MI_CardiacOutcome_2;
  // compute actual group-level effects
  r_1 = scale_r_cor(z_1, sd_1, L_1);
  r_1_muAcute_CardiacOutcome_1 = r_1[, 1];
  r_1_muAcute_CardiacOutcome_2 = r_1[, 2];
  // compute actual group-level effects
  r_2 = scale_r_cor(z_2, sd_2, L_2);
  r_2_muChronic_CardiacOutcome_1 = r_2[, 1];
  r_2_muChronic_CardiacOutcome_2 = r_2[, 2];
  // compute actual group-level effects
  r_3 = scale_r_cor(z_3, sd_3, L_3);
  r_3_muT1MI_CardiacOutcome_1 = r_3[, 1];
  r_3_muT1MI_CardiacOutcome_2 = r_3[, 2];
  // compute actual group-level effects
```

```
  r_4 = scale_r_cor(z_4, sd_4, L_4);
  r_4_muT2MI_CardiacOutcome_1 = r_4[, 1];
  r_4_muT2MI_CardiacOutcome_2 = r_4[, 2];
}
model {
  // likelihood including all constants
  if (!prior_only) {
    // vector combining observed and missing responses
    vector[N_TroponinDifference] Yl_TroponinDifference = Y_TroponinDifference;
    // initialize linear predictor term
    vector[N_CardiacOutcome] muAcute_CardiacOutcome = Intercept_muAcute_CardiacOutc
    // initialize linear predictor term
    vector[N_CardiacOutcome] muChronic_CardiacOutcome = Intercept_muChronic_CardiacO
    // initialize linear predictor term
    vector[N_CardiacOutcome] muT1MI_CardiacOutcome = Intercept_muT1MI_CardiacOutcom
    // initialize linear predictor term
    vector[N_CardiacOutcome] muT2MI_CardiacOutcome = Intercept_muT2MI_CardiacOutcom
    // linear predictor matrix
    vector[ncat_CardiacOutcome] mu_CardiacOutcome[N_CardiacOutcome];
    // initialize linear predictor term
    vector[N_TroponinDifference] mu_TroponinDifference = Intercept_TroponinDifferen
    Yl_TroponinDifference[Jmi_TroponinDifference] = Ymi_TroponinDifference;
    for (n in 1:N_CardiacOutcome) {
      // add more terms to the linear predictor
      muAcute_CardiacOutcome[n] += (bsp_muAcute_CardiacOutcome[1]) * Yl_TroponinDi
    }
    for (n in 1:N_CardiacOutcome) {
      // add more terms to the linear predictor
      muChronic_CardiacOutcome[n] += (bsp_muChronic_CardiacOutcome[1]) * Yl_Tropon
    }
    for (n in 1:N_CardiacOutcome) {
      // add more terms to the linear predictor
      muT1MI_CardiacOutcome[n] += (bsp_muT1MI_CardiacOutcome[1]) * Yl_TroponinDiffe
    }
    for (n in 1:N_CardiacOutcome) {
      // add more terms to the linear predictor
      muT2MI_CardiacOutcome[n] += (bsp_muT2MI_CardiacOutcome[1]) * Yl_TroponinDiffe
    }
    for (n in 1:N_CardiacOutcome) {
```

```
      mu_CardiacOutcome[n] = transpose([muAcute_CardiacOutcome[n], muChronic_Cardia
    }
    for (n in 1:N_CardiacOutcome) {
      target += categorical_logit_lpmf(Y_CardiacOutcome[n] | mu_CardiacOutcome[n]);
    }
    target += normal_lpdf(Yl_TroponinDifference | mu_TroponinDifference, sigma_Tro
  }
  // priors including all constants
  target += student_t_lpdf(b_muAcute_CardiacOutcome | 3,0,5);
  target += student_t_lpdf(bsp_muAcute_CardiacOutcome | 3,0,5);
  target += student_t_lpdf(b_muChronic_CardiacOutcome | 3,0,5);
  target += student_t_lpdf(bsp_muChronic_CardiacOutcome | 3,0,5);
  target += student_t_lpdf(b_muT1MI_CardiacOutcome | 3,0,5);
  target += student_t_lpdf(bsp_muT1MI_CardiacOutcome | 3,0,5);
  target += student_t_lpdf(b_muT2MI_CardiacOutcome | 3,0,5);
  target += student_t_lpdf(bsp_muT2MI_CardiacOutcome | 3,0,5);
  target += student_t_lpdf(b_TroponinDifference | 3,0,5);
  target += student_t_lpdf(sigma_TroponinDifference | 3, 0, 2.5)
    - 1 * student_t_lccdf(0 | 3, 0, 2.5);
  target += student_t_lpdf(sd_1 | 3, 0, 2.5)
    - 2 * student_t_lccdf(0 | 3, 0, 2.5);
  target += std_normal_lpdf(to_vector(z_1));
  target += lkj_corr_cholesky_lpdf(L_1 | 2);
  target += student_t_lpdf(sd_2 | 3, 0, 2.5)
    - 2 * student_t_lccdf(0 | 3, 0, 2.5);
  target += std_normal_lpdf(to_vector(z_2));
  target += lkj_corr_cholesky_lpdf(L_2 | 2);
  target += student_t_lpdf(sd_3 | 3, 0, 2.5)
    - 2 * student_t_lccdf(0 | 3, 0, 2.5);
  target += std_normal_lpdf(to_vector(z_3));
  target += lkj_corr_cholesky_lpdf(L_3 | 2);
  target += student_t_lpdf(sd_4 | 3, 0, 2.5)
    - 2 * student_t_lccdf(0 | 3, 0, 2.5);
  target += std_normal_lpdf(to_vector(z_4));
  target += lkj_corr_cholesky_lpdf(L_4 | 2);
}
generated quantities {
  // actual population-level intercept
  real b_muAcute_CardiacOutcome_Intercept = Intercept_muAcute_CardiacOutcome - dot_
```

```
// actual population-level intercept
real b_muChronic_CardiacOutcome_Intercept = Intercept_muChronic_CardiacOutcome -
// actual population-level intercept
real b_muT1MI_CardiacOutcome_Intercept = Intercept_muT1MI_CardiacOutcome - dot_pr
// actual population-level intercept
real b_muT2MI_CardiacOutcome_Intercept = Intercept_muT2MI_CardiacOutcome - dot_pr
// actual population-level intercept
real b_TroponinDifference_Intercept = Intercept_TroponinDifference - dot_product(
// compute group-level correlations
corr_matrix[M_1] Cor_1 = multiply_lower_tri_self_transpose(L_1);
vector<lower=-1,upper=1>[NC_1] cor_1;
// compute group-level correlations
corr_matrix[M_2] Cor_2 = multiply_lower_tri_self_transpose(L_2);
vector<lower=-1,upper=1>[NC_2] cor_2;
// compute group-level correlations
corr_matrix[M_3] Cor_3 = multiply_lower_tri_self_transpose(L_3);
vector<lower=-1,upper=1>[NC_3] cor_3;
// compute group-level correlations
corr_matrix[M_4] Cor_4 = multiply_lower_tri_self_transpose(L_4);
vector<lower=-1,upper=1>[NC_4] cor_4;
// extract upper diagonal of correlation matrix
for (k in 1:M_1) {
  for (j in 1:(k - 1)) {
    cor_1[choose(k - 1, 2) + j] = Cor_1[j, k];
  }
}
// extract upper diagonal of correlation matrix
for (k in 1:M_2) {
  for (j in 1:(k - 1)) {
    cor_2[choose(k - 1, 2) + j] = Cor_2[j, k];
  }
}
// extract upper diagonal of correlation matrix
for (k in 1:M_3) {
  for (j in 1:(k - 1)) {
    cor_3[choose(k - 1, 2) + j] = Cor_3[j, k];
  }
}
// extract upper diagonal of correlation matrix
```

```
  for (k in 1:M_4) {
    for (j in 1:(k - 1)) {
      cor_4[choose(k - 1, 2) + j] = Cor_4[j, k];
    }
  }
}
```

10 References

Emergency department care 2017–18, australian hospital statistics, 2018. URL https://www.aihw.gov.au/getmedia/9ca4c770-3c3b-42fe-b071-3d758711c23a/aihw-hse-216.pdf.aspx?inline=true.

2019. URL https://data.sa.gov.au/data/dataset/emergency-dept-presentations/resource/562af1ca-900e-493c-9318-6ab105bf8eb9.

Patrick Aboagye-Sarfo, Qun Mai, Frank M Sanfilippo, David B Preen, Louise M Stewart, and Daniel M Fatovich. A comparison of multivariate and univariate time series approaches to modelling and forecasting emergency department demand in western australia. *Journal of biomedical informatics*, 57:62–73, 2015.

Alan Agresti. *Foundations of linear and generalized linear models*. John Wiley & Sons, 2015.

Deniz Akinc and Martina Vandebroek. Bayesian estimation of mixed logit models: Selecting an appropriate prior for the covariance matrix. *Journal of choice modelling*, 29:133–151, 2018a.

Deniz Akinc and Martina Vandebroek. Bayesian estimation of mixed logit models: Selecting an appropriate prior for the covariance matrix. *Journal of choice modelling*, 29:133–151, 2018b.

Douglas G Altman and J Martin Bland. Missing data. *Bmj*, 334(7590):424–424, 2007.

Tamara M Atkinson, George D Giraud, Brandon M Togioka, Daniel B Jones, and Joaquin E Cigarroa. Cardiovascular and ventilatory consequences of laparoscopic surgery. *Circulation*, 135(7): 700–710, 2017.

David Banham, Jonathan Karnon, Kirsten Densley, and John W Lynch. How much emergency department use by vulnerable populations is potentially preventable?: A period prevalence study of linked public hospital data in south australia. *BMJ open*, 9(1):e022845, 2019.

Melanie L Bell and Diane L Fairclough. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. *Statistical methods in medical research*, 23(5):440–459, 2014.

Michael Betancourt and Mark Girolami. Hamiltonian monte carlo for hierarchical models. *Current trends in Bayesian methodology with applications*, 79(30):2–4, 2015.

Stephen P Brooks and Andrew Gelman. General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics*, 7(4):434–455, 1998.

Burkner. Handle missing values with brms. 2015.

P Bürkner. Bayesian regression models using stan. *R package version*, 1(0), 2017a.

Paul-Christian Bürkner. brms: An r package for bayesian multilevel models using stan. *Journal of statistical software*, 80(1):1–28, 2017b.

A Colin Cameron and Pravin K Trivedi. *Microeconometrics: methods and applications*. Cambridge university press, 2005.

Peter A Cameron, Anthony P Joseph, and Sally M McCarthy. Access block can be managed. *Medical Journal of Australia*, 190(7):364–368, 2009.

Magnolia Cardona, Ebony T Lewis, Mette R Kristensen, Helene Skjøt-Arkil, Anette Addy Ekmann, Hanne H Nygaard, Jonas J Jensen, Rune O Jensen, Jonas L Pedersen, Robin M Turner, et al. Predictive validity of the cristal tool for short-term mortality in older people presenting at emergency departments: a prospective study. *European geriatric medicine*, 9(6):891–901, 2018.

Bob Carpenter, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: a probabilistic programming language. *Grantee Submission*, 76(1):1–32, 2017.

Trevor Chan, Glenn Arendts, and Mary Stevens. Variables that predict admission to hospital from an emergency department observation unit. *Emergency Medicine Australasia*, 20(3):216–220, 2008.

Yeojin Chung, Sophia Rabe-Hesketh, Vincent Dorie, Andrew Gelman, and Jingchen Liu. A nondegenerate penalized likelihood estimator for variance parameters in multilevel models. *Psychometrika*, 78(4):685–709, 2013.

Emerging Risk Factors Collaboration et al. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. *The Lancet*, 375 (9709):132–140, 2010.

Louise Cullen, Jaimi Greenslade, Katharina Merollini, Nicholas Graves, Christopher J K Hammett, Tracey Hawkins, Martin P Than, Anthony F T Brown, Christopher B Huang, Seyed E Panahi, et al. Cost and outcomes of assessing patients with chest pain in an australian emergency department. *Medical Journal of Australia*, 202(8):427–432, 2015.

Philippe Demers, Stéphane Elkouri, Raymond Martineau, André Couturier, and Raymond Cartier. Outcome with high blood lactate levels during cardiopulmonary bypass in adult cardiac operation. *The Annals of thoracic surgery*, 70(6):2082–2086, 2000.

Peter J Diggle, Jonathan A Tawn, and Rana A Moyeed. Model-based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47(3):299–350, 1998.

Yiran Dong and Chao-Ying Joanne Peng. Principled missing data methods for researchers. *SpringerPlus*, 2(1):1–17, 2013.

Lukah Dykes. Heartai: Purpose and goals, Feb 2021. URL `https://www.heartai.net/overview/purposes-and-goals.html`.

Susan A Everson-Rose and Tené T Lewis. Psychosocial factors and cardiovascular diseases. *Annu. Rev. Public Health*, 26:469–500, 2005.

Peter D Faris, William A Ghali, Rollin Brant, Colleen M Norris, P Diane Galbraith, Merril L Knudtson, Approach Investigators, et al. Multiple imputation versus data enhancement for dealing with missing data in observational health care outcome analyses. *Journal of clinical epidemiology*, 55 (2):184–191, 2002.

MFM Fathil, MK Md Arshad, Subash CB Gopinath, U Hashim, R Adzhri, RM Ayub, AR Ruslinda, M Nuzaihan, AH Azman, M Zaki, et al. Diagnostics on acute myocardial infarction: Cardiac troponin biomarkers. *Biosensors and Bioelectronics*, 70:209–220, 2015.

Gerry FitzGerald, Sam Toloo, Joanna Rego, Joseph Ting, Peter Aitken, and Vivienne Tippett. Demand for public hospital emergency department services in australia: 2000–2001 to 2009–2010. *Emergency Medicine Australasia*, 24(1):72–78, 2012.

Lendie Follett and Brian Vander Naald. Explaining variability in tourist preferences: A bayesian model well suited to small samples. *Tourism Management*, 78:104067, 2020.

Thaís CO Fonseca, Marco AR Ferreira, and Helio S Migon. Objective bayesian analysis for the student-t regression model. *Biometrika*, 95(2):325–333, 2008.

Zujie Gao, Zengsheng Chen, Anqiang Sun, and Xiaoyan Deng. Gender differences in cardiovascular disease. *Medicine in Novel Technology and Devices*, 4:100025, 2019.

Gary C Geelhoed and Nicholas H de Klerk. Emergency department overcrowding, mortality and the 4-hour rule in western australia. *Medical Journal of Australia*, 196(2):122–126, 2012.

Seymour Geisser. *Predictive inference*, volume 55. CRC press, 1993.

AB Gelman, JB Carlin, HS Stern, DB Dunson, A Vehtari, and D Rubin. Bayesian data analysis third edition. boca raton. *FL: CRC Press.[Google Scholar]*, 2013.

Andrew Gelman. Prior choice recommendations, Apr 2020. URL `https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations`.

Andrew Gelman and Yuling Yao. Holes in bayesian statistics. *Journal of Physics G: Nuclear and Particle Physics*, 48(1):014002, 2020.

Andrew Gelman, Donald B Rubin, et al. Inference from iterative simulation using multiple sequences. *Statistical science*, 7(4):457–472, 1992.

Andrew Gelman, Aleks Jakulin, Maria Grazia Pittau, Yu-Sung Su, et al. A weakly informative default prior distribution for logistic and other regression models. *Annals of applied Statistics*, 2 (4):1360–1383, 2008.

Andrew Gelman, Aki Vehtari, Daniel Simpson, Charles C Margossian, Bob Carpenter, Yuling Yao, Lauren Kennedy, Jonah Gabry, Paul-Christian Bürkner, and Martin Modrák. Bayesian workflow. *arXiv preprint arXiv:2011.01808*, 2020.

Soumyadip Ghosh and Shane G Henderson. Behavior of the norta method for correlated random vector generation as the dimension increases. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 13(3):276–294, 2003.

M Kashif Gill, Tirusew Asefa, Yasir Kaheil, and Mac McKee. Effect of missing data on performance of learning algorithms for hydrologic predictions: Implications to an imputation technique. *Water resources research*, 43(7), 2007.

David Gregg and Pascal J Goldschmidt-Clermont. Platelets and cardiovascular disease. *Circulation*, 108(13):e88–e90, 2003.

Jarrod D Hadfield et al. Mcmc methods for multi-response generalized linear mixed models: the mcmcglmm r package. *Journal of statistical software*, 33(2):1–22, 2010.

Mark S Handcock and Michael L Stein. A bayesian analysis of kriging. *Technometrics*, 35(4): 403–410, 1993.

Matthew D Hoffman and Andrew Gelman. The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *J. Mach. Learn. Res.*, 15(1):1593–1623, 2014.

Joseph G Ibrahim, Haitao Chu, and Ming-Hui Chen. Missing data in clinical studies: issues and methods. *Journal of clinical oncology*, 30(26):3297, 2012.

Harry Joe. Generating random correlation matrices based on partial correlations. *Journal of Multivariate Analysis*, 97(10):2177–2189, 2006.

Moo-Young Kim, Sun Ha Jee, Ji Eun Yun, Soo Jin Baek, and Duk-Chul Lee. Hemoglobin concentration and risk of cardiovascular disease in korean men and women-the korean heart study. *Journal of Korean medical science*, 28(9):1316, 2013.

Susan W Kim, Jordan Y Li, Paul Hakendorf, David JO Teubner, David I Ben-Tovim, and Campbell H Thompson. Predicting admission of patients by their presentation to the emergency department. *Emergency Medicine Australasia*, 26(4):361–367, 2014.

Lewis H Kuller, June E Eichner, Trevor J Orchard, Gregory A Grandits, Lisa McCallum, Russell P Tracy, and Multiple Risk Factor Intervention Trial Research Group. The relation between serum albumin levels and risk of coronary heart disease in the multiple risk factor intervention trial. *American journal of epidemiology*, 134(11):1266–1277, 1991.

M Laskou, H Nikolaou, F Diamantea, A Vlahou, C Mathas, and N Maguina. Iatrogenic complications in icu patients. *Critical Care*, 10(1):1–1, 2006.

Chong D Lee, Aaron R Folsom, F Javier Nieto, Lloyd E Chambless, Eyal Shahar, and Douglas A Wolfe. White blood cell count and incidence of coronary heart disease and ischemic stroke, and mortality from cardiovascular disease in african-american and white men and women: The atherosclerosis risk in communities study. 2001.

Michael Levy, Diane Heels-Ansdell, Rajesh Hiralal, Mohit Bhandari, Gordon Guyatt, Salim Yusuf, Deborah Cook, Juan Carlos Villar, Matthew McQueen, Edward McFalls, et al. Prognostic value of troponin and creatine kinase muscle and brain isoenzyme measurement after noncardiac surgery: a systematic review and meta-analysis. *The Journal of the American Society of Anesthesiologists*, 114(4):796–806, 2011.

Don Liew, Danny Liew, and Marcus P Kennedy. Emergency department length of stay independently predicts excess inpatient length of stay. *Medical Journal of Australia*, 179(10):524–526, 2003.

Antonio R Linero and Michael J Daniels. Bayesian approaches for missing not at random outcome data: the role of identifying restrictions. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 33(2):198, 2018.

Jun-Ru Liu, Yang Liu, Fu-Zai Yin, and Bo-Wei Liu. Serum ferritin, an early marker of cardiovascular risk: a study in chinese men of first-degree relatives with family history of type 2 diabetes. *BMC cardiovascular disorders*, 19(1):82, 2019.

Russell V Luepker, James M Raczynski, Stravoula Osganian, Robert J Goldberg, John R Finnegan Jr, Jerris R Hedges, David C Goff Jr, Mickey S Eisenberg, Jane G Zapka, Henry A Feldman, et al. Effect of a community intervention on patient delay and emergency medical service use in acute coronary heart disease: The rapid early action for coronary treatment (react) trial. *Jama*, 284(1):60–67, 2000.

Hazem M Mansour, Yasmin N El-Sakhawy, et al. Initially presented acute coronary syndrome: does d-dimer imply any clinical significance? *The Egyptian Journal of Haematology*, 45(1):23, 2020.

Richard McElreath. *Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC press, 2020.

Authors/Task Force Members, Steen Dalby Kristensen, Juhani Knuuti, Antti Saraste, Stefan Anker, Hans Erik Bøtker, Stefan De Hert, Ian Ford, Jose Ramón Gonzalez-Juanatey, Bulent Gorenek, et al. 2014 esc/esa guidelines on non-cardiac surgery: cardiovascular assessment and management: The joint task force on non-cardiac surgery: cardiovascular assessment and management of the european society of cardiology (esc) and the european society of anaesthesiology (esa). *European heart journal*, 35(35):2383–2431, 2014.

M Miller. Dyslipidemia and cardiovascular risk: the importance of early prevention. *QJM: An International Journal of Medicine*, 102(9):657–667, 2009.

Firas F Mussa, Hong Chai, Xinwen Wang, Qizhi Yao, Alan B Lumsden, and Changyi Chen. Chlamydia pneumoniae and vascular disease: an update. *Journal of Vascular Surgery*, 43(6):1301–1307, 2006.

Shinichi Nakagawa. Missing data: mechanisms, methods and messages. *Ecological statistics: Contemporary theory and application*, pages 81–105, 2015.

Radford M Neal et al. Mcmc using hamiltonian dynamics. *Handbook of markov chain monte carlo*, 2(11):2, 2011.

Akihiko Nishimura, David Dunson, and Jianfeng Lu. Discontinuous hamiltonian monte carlo for sampling discrete parameters. *arXiv preprint arXiv:1705.08510*, 853, 2017.

Christine Perret-Guillaume, Laure Joly, and Athanase Benetos. Heart rate as a risk factor for cardiovascular disease. *Progress in cardiovascular diseases*, 52(1):6–10, 2009.

Juho Piironen, Aki Vehtari, et al. Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electronic Journal of Statistics*, 11(2):5018–5051, 2017.

Tobias Reichlin, Willibald Hochholzer, Stefano Bassetti, Stephan Steuer, Claudia Stelzig, Sabine Hartwiger, Stefan Biedert, Nora Schaub, Christine Buerge, Mihael Potocki, et al. Early diagnosis of myocardial infarction with sensitive cardiac troponin assays. *New England Journal of Medicine*, 361(9):858–867, 2009.

Drew Richardson, Anne-Maree Kelly, and Debra Kerr. Prevalence of access block in australia 2004–2008. *Emergency Medicine Australasia*, 21(6):472–478, 2009.

Drew B Richardson. The access-block effect: relationship between delay to reaching an inpatient bed and inpatient length of stay. *Medical Journal of Australia*, 177(9):492–495, 2002.

Małgorzata Roos, Leonhard Held, et al. Sensitivity analysis in bayesian generalized linear mixed models for binary data. *Bayesian Analysis*, 6(2):259–278, 2011.

Lorcan Ruane, Thomas Buckley, Soon YS Hoo, Peter S Hansen, Catherine McCormack, Elizabeth Shaw, Judith Fethney, and Geoffrey H Tofler. Triggering of acute myocardial infarction by respiratory infection. *Internal medicine journal*, 47(5):522–529, 2017.

Mark J Sarnak, Andrew S Levey, Anton C Schoolwerth, Josef Coresh, Bruce Culleton, L Lee Hamm, Peter A McCullough, Bertram L Kasiske, Ellie Kelepouris, Michael J Klag, et al. Kidney disease as a risk factor for development of cardiovascular disease: a statement from the american heart association councils on kidney in cardiovascular disease, high blood pressure research, clinical cardiology, and epidemiology and prevention. *Circulation*, 108(17):2154–2169, 2003.

David J Spiegelhalter, Nicola G Best, Bradley P Carlin, and Angelika Van der Linde. The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, pages 485–493, 2014.

Peter C Sprivulis, Julie-Ann Da Silva, Ian G Jacobs, George A Jelinek, and Amanda RL Frazer. The association between hospital overcrowding and mortality among patients admitted via western australian emergency departments. *Medical Journal of Australia*, 184(5):208–212, 2006.

Sudeep Srivastava and Liang Chen. Comparison between the stochastic search variable selection and the least absolute shrinkage and selection operator for genome-wide association studies of rheumatoid arthritis. In *BMC proceedings*, volume 3, pages 1–7. Springer, 2009.

Morteza Tavakol, Salman Ashraf, and Sorin J Brener. Risks and complications of coronary angiography: a comprehensive review. *Global journal of health science*, 4(1):65, 2012.

Sara Van Erp, Daniel L Oberski, and Joris Mulder. Shrinkage priors for bayesian penalized regression. *Journal of Mathematical Psychology*, 89:31–50, 2019.

Dootika Vats and Christina Knudson. Revisiting the gelman-rubin diagnostic. *arXiv preprint arXiv:1812.09384*, 2018.

Aki Vehtari, Daniel Simpson, Andrew Gelman, Yuling Yao, and Jonah Gabry. Pareto smoothed importance sampling. *arXiv preprint arXiv:1507.02646*, 2015.

Sumio Watanabe and Manfred Opper. Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of machine learning research*, 11(12), 2010.

Charles B Yackulic, Michael Dodrill, Maria Dzul, Jamie S Sanderlin, and Janice A Reid. A need for speed in bayesian population models: a practical guide to marginalizing and recovering discrete latent states. *Ecological Applications*, 30(5):e02112, 2020.

Yihua Zhao, John Staudenmayer, Brent A Coull, and Matthew P Wand. General design bayesian generalized linear mixed models. *Statistical science*, pages 35–51, 2006.

Xiang Zhou and Jerome P Reiter. A note on bayesian inference after multiple imputation. *The American Statistician*, 64(2):159–163, 2010.