UMC Utrecht
Nationaal Vergiftigingen Informatie Centrum

INTOXICATE STUDY

Universiteit Utrecht

Impact Of Country On Adverse Outcome In Acutely Intoxicated Patients In ICU: A Generalized Linear Mixed Models Analysis

General Research Profile Thesis

Author: Stavri Karasiali[1] **Supervisors:** Samanta Zwang[2] **Examiners:** Claudine Hunault[2]

Claudine Hunault[2] Said el Bouhaddani[3]

1. Master Students at Toxicology & Environmental Health, Utrecht University , The Netherlands
2. Nationaal Vergiftigingen Informatie Centrum, UMC Utrecht, Utrecht, The Netherlands
3. Dept. Data science & Biostatistics, Julius Center for Health Sciences and Primary Care, UMC Utrecht, The Netherlands

AUGUST 24, 2023

Table of Contents

Abstract

Our study aimed to investigate poor outcomes among ICU patients with acute poisoning, with a specific focus on discerning international variations. Additionally, we assessed the efficacy of Generalized Linear Mixed Models (GLMM) when compared to Generalized Linear Models (GLMs). We found that patient outcomes significantly diverge across countries, even after accounting for various factors. Notably, countries introduced as random effects within the GLMM exerted a substantial influence on patient outcomes, reflecting the influence of distinct healthcare systems, sociocultural factors, and resource availability. In terms of modelling, GLMM outperformed GLMs. The inclusion of laboratory results alongside patient characteristics notably improved the ability to discriminate poor outcomes. The most effective model incorporated a comprehensive set of parameters, encompassing physiological indicators, laboratory findings, and exposure types. Our analysis unveiled the complexity of the data, indicating that relying solely on physiological and laboratory parameters does not suffice to explain patient outcomes fully. Exposure types, as revealed in this study, significantly contribute to our understanding, aligning with prior research on the mortality implications of intoxication. However, while exposures remained relevant, their impact was comparatively modest. The study also highlighted challenges related to categorizing exposures due to methodological disparities. To tackle this challenge, we must explore and evaluate new exposure categories. Furthermore, concerns emerged regarding patients exposed to multiple substances, warranting the need for a refined approach to exposure categorization and prudent consideration of exposure dosages. To enhance our models, we should consider transforming physiological and laboratory parameters to establish linear relationships with outcomes, thus addressing challenges posed by right-skewed distributions and outliers. In conclusion, our study sheds light on the varying patient outcomes following intoxication, emphasizing the effectiveness of GLMM as an analytical tool, emphasizing the importance of considering exposure types alongside physiological and laboratory parameters. This comprehensive approach emphasizes the need for future research to refine exposure categorization, consider exposure dosages, and carefully handle other parameters to establish linear relationships with outcomes.

Layman's Summary

In our study, we embarked on an exploration to unravel the factors influencing the outcomes of ICU patients who had experienced acute poisoning, with a particular emphasis on understanding why these outcomes differed across various countries. Additionally, we sought to determine whether employing a statistical approach known as Generalized Linear Mixed Models (GLMM) yielded more insightful results than the conventional Generalized Linear Models (GLMs).

What emerged from our investigation was a notable discrepancy in patient outcomes between different countries, even when accounting for variables like age and gender. It became evident that the variations were closely linked to the disparities in healthcare systems, cultural norms, and available resources across these nations. The GLMM approach proved its mettle by providing superior results compared to the GLMs. By incorporating data from laboratory tests in conjunction with patient-specific details, our predictive accuracy for adverse outcomes significantly improved. The most effective model we developed incorporated an extensive range of parameters, encompassing not only laboratory findings and patient characteristics but also the specific type of poison involved.

Our journey through the data revealed its inherent complexity. It became increasingly clear that relying solely on information from laboratory tests and patient records was insufficient to fully elucidate the intricacies of the outcomes. The type of poison to which individuals were exposed emerged as a significant factor, albeit with varying degrees of impact. Categorizing these poisons presented challenges due to the diverse methods employed by different individuals, and the added complexity of some patients being exposed to multiple poisons further compounded the issue. To enhance the robustness of our models, we recognized the need to revamp our approach to interpreting laboratory results and patient data, simplifying the process while taking into account data anomalies that deviated from the expected patterns or displayed unusual characteristics.

In summary, our study illuminated the stark divergence in patient outcomes following poisoning incidents across different countries. The application of GLMM as our analytical tool proved to be a more effective approach compared to GLMs. We learned that the comprehensive evaluation of patient outcomes necessitates not only an examination of laboratory results and patient profiles but also a careful consideration of the specific type of poison involved. While this presented its own set of challenges, it underscored the importance of adapting our methodology to accommodate these complexities in the pursuit of more accurate models and results.

Introduction

Acute poisoning represents a critical medical emergency with potential life-threatening consequences. The prediction of prognosis in patients with acute poisoning bears significant clinical importance by facilitating timely and appropriate treatment interventions (Han, et al., 2021). Variations in admission rates are evident across hospitals and countries, particularly regarding the admission of intoxicated patients to the ICU. Intoxication refers to the manifestation of harmful effects in humans caused by the exposure to a single or repeated instances of a mixture, whether it is a naturally occurring or synthetic substance, commonly found in the market or existing in the surrounding environment (de Lange, Hunault, & Zwaag, 2021)

Developed countries tend to prioritize ICU admissions for intoxicated patients primarily for observational purposes rather than immediate treatment. This approach acknowledges the critical importance of closely monitoring these patients, as severe symptoms may still arise if the xenobiotic substances have not yet reached their peak concentration (Brandenburg, et al., 2016). The patients with acute intoxication comprise approximately 14% of ICU admissions, however, the in-hospital mortality rate among these patients is relatively modest, estimated at around 2% (Hondebrink, et al., 2021)

The mortality rate among patients with acute poisoning is contingent upon their physiological condition and the distinctive attributes of the poisoning episode. Various factors, including the type of substance, route of exposure, and intent of poisoning, significantly impact the observed outcomes in these individuals. Consequently, an effective mortality prediction model for patients with acute poisoning must encompass both poisoning-related characteristics and the patients' physiological status, while also accommodating individuals across all age groups (Han, et al., 2021). Distinguishing between patients who do not necessitate immediate treatment and those at risk of experiencing severe symptoms and complications poses a significant challenge.

Previous studies have investigated different factors and different clinical systems utilized in critical care. The Acute Physiology and Chronic Health Evaluation (APACHE) score and Simplified Acute Physiology Score (SAPS) are widely employed tools in the intensive care unit, specifically for forecasting outcomes in specific poisoning cases (Brandenburg, Brinkman, de Keizer, Meulenbelt, & de Lange, 2014). The established association between certain laboratory and physiological parameters and an extended ICU stay highlights their significance in clinical outcomes. However, the investigation into the relationship between exposure type and outcomes, beyond mortality, remains inadequately explored in the existing scientific literature (Liisanantti, Ohtonen, Kiviniemi, Laurila, & Ala-Kokko, 2011; Clark, Binswanger, & Moss, 2014).

Moreover, extensive research has been conducted to investigate the outcomes of acute poisoning observed among different countries worldwide, including studies conducted in regions such as Australia (Henderson, Wright, & Pond, 1993; Cretikos & Parr, 2003), Europe (Brien, Murphy, Conrick-Martin, & Marsh, 2009; Liisanantti, Ohtonen, Kiviniemi, Laurila, & Ala-Kokko, 2011) and Asia (Fathelrahman, Ab Rahman, & Zain, 2008; Lam, Lau, & Yan, 2010; Han, et al., 2021). This examination across diverse geographic regions highlights the urgent need for a comprehensive understanding of the multifactorial determinants contributing to the divergent outcomes in cases of acute poisoning. By examining these variations in outcomes, valuable insights can be gained to improve our understanding of the factors influencing the prognosis and management of acute poisoning globally. As well the difference in outcome between different countries across the globe, from Australia to UK, including countries from Asia, Middle East, Africa and Europe.

Given the intricate nature of our dataset, drawn from intoxicated patients in diverse countries, we adopted a mixed model approach for our analysis. This decision was prompted by the inherent hierarchical or clustered structure of our data, stemming from the complex interplay of contextual factors that influence patient outcomes. Patients residing within the same country or region often share common characteristics and risk profiles related to intoxication. These shared traits are shaped by a multitude of determinants, encompassing healthcare infrastructure, cultural norms, and access to medical services (Mulyaningsih, et al., 2021). Additionally, individual patient attributes, comorbidities, and exposure types can vary significantly across different countries. The utilization of mix effect modelling provides a robust framework to accommodate these intricacies, both within and between countries. This approach enables a comprehensive exploration of the factors contributing to poor patient outcomes.

The primary objective of this study is to examine the association between the prevalent poor outcomes observed in ICU patients across diverse geographical areas and the risk factors related to exposure type and patients' characteristics. To achieve this, we perform a Generalized linear mixed model (GLMM). The objectives of the study:

1. Does the risk of poor outcome in ICU intoxicated patients differ between countries after adjustment for demographic, clinical examination, lab results and exposures variables?
2. To assess the potential improvement in predictive accuracy for poor outcomes among ICU patients following acute poisoning through the implementation of GLMM, in comparison to the Generalized Linear Models (GLM) utilized in previous investigations.

Methods

Design of the Study

This study is a part of a multinational and multi-centre observational study, the INTOXICATE study, which focuses on examining the outcomes and prognosis of patients admitted to intensive care units (ICUs) following incidents of acute intoxication. In the scope of the study, an intensive care unit (ICU) is characterized as a specialized facility capable of performing endotracheal intubation and delivering mechanical ventilation to patients. It encompasses a range of specialties, including medical, surgical, toxicology, and other specialized units. Furthermore, high-dependency units (HDUs) or high-care units (HCUs) that offer comprehensive care for critically ill patients and possess the capability to provide mechanical ventilation, if necessary, are considered as part of the ICU framework within the INTOXICATE study (de Lange, Hunault, & Zwaag, 2021).

The study encompasses multiple sites across Europe and other continents, including Asia, Australia, UK and Africa. Clinicians of the participating ICUs were referred to the study through its website (www.toxicstudy.org). While participation is not limited by geographic boundaries, it is worth noting that the majority of patients included in the study are primarily sourced from ICUs located in Europe, particularly the Netherlands (***Table 1***). Because of the limited patient inclusion, Greece (N=2), Sudan (N=1), and United Stated (N=2) will be excluded from the data analysis.

Table 1: *The number (N) and the percentage (%) of the patients that included in the study per country The total number of patients that included in the study is 1915.*

COUNTRY	NUMBER OF PATIENTS (N)	% OF TOTAL PATIENTS
NETHERLANDS	873	45.75
SPAIN	214	11.22
SWEDEN	200	10.48
TURKEY	119	6.24
BELGIUM	114	5.97
AUSTRALIA	63	3.30
AUSTRIA	53	2.78
JORDAN	42	2.20
GERMANY	34	1.78
LIBYA	33	1.73
ROMANIA	33	1.73
PORTUGAL	31	1.62
LITHUANIA	28	1.47
EGYPT	20	1.05
UNITED KINGDOM	19	1.00
ITALY	10	0.52
PALESTINE	10	0.52
CROATIA	7	0.37
BRUNEI	5	0.26
GREECE	2	0.10
UNITED STATES	2	0.10
SUDAN	1	0.05

Patients and Data Collection

From January 2021 to July 2023, which signifies the completion of the inclusion period, participating units diligently provide data from acutely intoxicated patients admitted to their respective ICUs. Those data are integrated into the online data aggregation platform, Castor EDC. This integration is achieved through the use of a comprehensive questionnaire, which captures crucial information including patient characteristics, exposure details, clinical assessment, administered treatments, complications, and vital status at the time of hospital discharge.

The study employs specific criteria to ensure that the selected patient population focuses on those who meet the necessary requirements for ICU/HDU admission due to intoxication. These criteria include:

(1) Patients must have been directly admitted to the intensive care unit (ICU) or high-dependency unit (HDU) from an ambulance or the Emergency Department, or transferred from a medical or surgical ward to the ICU/HDU.

(2) The primary reason for ICU/HDU admission should be intoxication.

(3) Patients must have stayed in the ICU/HDU for a minimum duration of at least 4 hours.

(4) The study includes patients (males/females) who are 18 years of age or older.

The study applied strict exclusion criteria to ensure a well-defined patient population, enhancing the relevance and validity of the findings regarding the outcomes and prognosis of ICU/HDU patients specifically related to intoxication. The following criteria were implemented:

(1) The outcome used in this study could not be determined.
(2) The patient had no ICU admission date
(3) The data entry was too empty to use. We removed the patients that had >50% of missing data (Marino, Lucas, Latour, & Heintzman, 2021; Lee & Carlin, 2012; Haji-Maghsoudi, Haghdoost , Rastegari, & Baneshi, 2013; Wulff & Jeppesen, 2017)

The data for this study were extracted from the Castor EDC database as of July 31, 2023. The patient selection process involved applying a predefined exclusion criteria, which has been represented in ***Figure 1***. The total patients in Castor were 2056, which after the use of the exclusion criteria the final patients that used in the analysis were 1907.

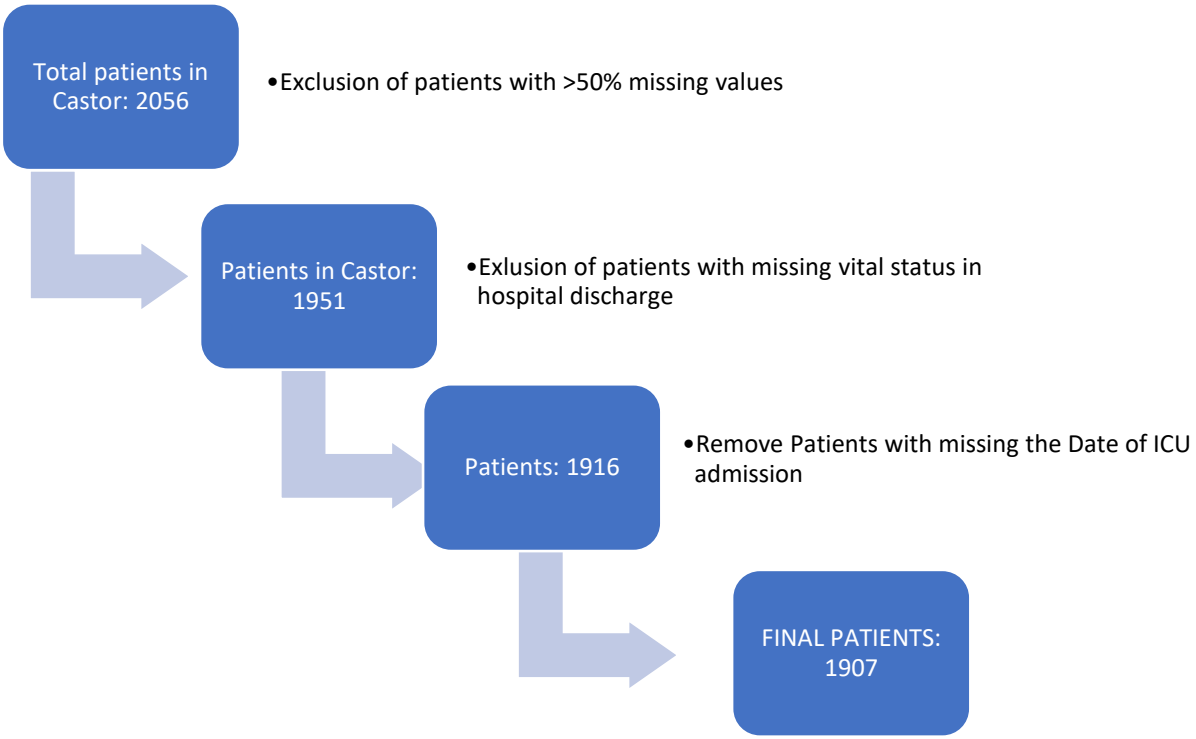


Figure 1: *The number of patients in Castor in the start of the analysis and after the exclusion of patients with missing important information. The exclusion criteria used show in each step.*

Definitions of Variables

A. Poor outcome:

It is defined in this study as a patient who either passes away before hospital discharge, as indicated by "Deceased at ICU/HDU" or "Deceased at the ward following ICU/HDU discharge," or experiences complications during their ICU stay. The complications recorded in the Castor database include the following:

1. Acute liver failure
2. Acute renal failure
3. Anoxic brain injury
4. Aspiration pneumonitis
5. Coma
6. Hospital acquired infection
7. Hypertensive crisis
8. Hypoxic-ischemic brain injury
9. Respiratory failure
10. Other
11. No complications

In cases where a specific complication is not listed in the database, the investigator has the flexibility to select the "Other" option and manually input the complication. This allows for the inclusion of additional complications that may not be predefined in the database.

Furthermore, it is important to note that the definition of a poor outcome extends beyond the duration of hospital discharge. Specifically, it encompasses the assessment of patient status up to 30 days following hospital discharge, taking into account whether the patient remains alive or not during this period. Conversely, a good outcome is defined as the patient being "Alive" without any documented complications upon and after 30 days of hospital discharge.

B. Comorbidities:

Comorbidity refers to the presence of additional medical conditions that coexist alongside a primary health condition. These comorbidities can encompass both physical and mental health disorders and are typically chronic in nature. In the context of intoxication, the presence of comorbidities can lead to an elevated risk of hospitalization, heightened susceptibility to adverse treatment effects, increased demands on both patients and healthcare professionals, escalated healthcare expenses, diminished quality of life, and higher mortality rates. It is crucial to consider comorbidities when evaluating the overall health status and outcomes of individuals affected by intoxication. Comorbidities used in this study and listed in Castor dataset as follows:

1. Acute Myocardial Infarction (AMI)
2. Acute airway obstruction
3. Addiction
4. Arrythmia
5. Chronic cardiovascular insufficiency
6. Chronicled haemodialysis
7. Chronic Kidney Disease (CKD)
8. Chronic Obstructive Pulmonary Disease (COPD)
9. Cirrhosis
10. Delirium
11. Demetria
12. Diabetes
13. Haematological malignancy
14. High intracranial pressure
15. Immunodeficiency
16. Maltutrition/weight loss
17. Metabolic/endocrine disease
18. Metastasized cancer
19. Paralysis
20. Primary Epilepsy
21. Psychiatric disease
22. Sepsis
23. Stroke
24. Other
25. Severe respiratory disease with e.g. oxygen use or mechanical ventilation at home
26. None

In the analysis of the data, physicians had the option to specify comorbidities in the "Other" category. This allowed for the identification of various comorbidities present among the patients. However, it was observed that some of the conditions initially categorized as comorbidities were, in fact, symptoms directly related to the intoxication itself. Therefore, to ensure accuracy, these symptoms were not considered as separate comorbidities and were excluded from the final analysis.

To classify the comorbidities appropriately, two distinct categories were established. The first category focused on somatic comorbidities, encompassing conditions such as Acute Myocardial Infarction (AMI), Chronic Kidney Disease (CKD), and others that primarily relate to physical health. The second category specifically addressed psychiatric comorbidities, which included conditions like Dementia, Psychiatric Diseases, Alzheimer's Disease, Addiction, and other mental health disorders. By segregating comorbidities into these two categories, the analysis aimed to capture and

differentiate the various types of comorbidities experienced by the patients, considering both their physical and mental health aspects.

C. Clinical assessment and Laboratory Results:

Glasgow Coma score (GCS): In Castor, the GCS is recorded in two different ways, allowing for flexibility in score calculation. Users have the option to record the individual sub-scores and Castor computes the total GCS score, or directly input the total GCS score. By incorporating both approaches, Castor enables the derivation of a comprehensive and accurate final score that serves as a crucial component for statistical analysis. In order to capture the complete spectrum of patient conditions, specific limits for GCS calculation, GCS manual, and GCS admission have been set at 3. This choice ensures that patients who are intubated during their admission at ICU are still included in the analysis. By establishing the minimum GCS final score as 3, the system accounts for these cases and avoids the exclusion of important data.

Platelet count: To ensure the integrity of the analysis, a minimum limit for platelet counts has been established at 150 mmol/L in Castor. Any value below this threshold is considered indicative of thrombocytopenia, which could potentially introduce confounding factors. Consequently, patients falling below this limit are excluded from the analysis.

Potassium concentration: The variable is categorized as bad potassium levels, which are defined as values above 5.0 and less than 3.0 mmol/L. "Bad potassium levels" encompass both hyperkalemia (elevated potassium levels) and hypokalemia (low potassium levels).

Sodium concentration: The variable is categorized as either hyponatremia or hypernatremia, with specific limits set at 145 mmol/L and 135 mmol/L, respectively. Hyponatremia refers to a lower than normal concentration of sodium in the blood, while hypernatremia indicates an elevated sodium level.

Heart rate: The variable is categorized as tachycardia which is defined as a heart rate exceeding 109 beats per minute. In this analysis, bradycardia is not considered as it has been established that it does not exhibit a significant relationship with prolonged ICU stay (Liisanantti, Ohtonen, Kiviniemi, Laurila, & Ala-Kokko, 2011).

Systolic Blood Pressure: The variable is categorized as "low systolic blood pressure," defined by a threshold of less than 90 mmHg. In this analysis, high systolic blood pressure is not considered, as previous research (Liisanantti, Ohtonen, Kiviniemi, Laurila, & Ala-Kokko, 2011) has indicated that it does not exhibit a significant relationship with prolonged ICU stay.

Temperature: The variable is defined as “hyperthermia” and “hypothermia” with thresholds as >38.3 and <35 C^{o}, respectively. Hyperthermia refers to an elevated body temperature, while hypothermia indicates a lower than normal body temperature.

During the analysis of the remaining laboratory variables including **Arterial pH, Creatinine, and White Blood Cells (WBC),** a thorough quality check of the data was conducted. Values that did not make sense or appeared to be inconsistent were identified and corrected to ensure the accuracy and reliability of the dataset. In addition to addressing erroneous data, particular attention was given to the units used for these variables. All values were standardized to a uniform unit of measurement prior to conducting the analysis. By standardizing the units, potential discrepancies stemming from variations in unit measurements were effectively eliminated. This critical step enabled a consistent and meaningful comparison of the laboratory results.

D. Exposures:

The categorization of exposure types utilized in the study is derived from various sources, including APACHE IV, the annual reports of the American Association of Poison Control Centres, and relevant scientific articles. Notable references in this regard include studies conducted by Brandenburg et al. (2016), and (Rezar, et al., 2022).

- **Cardiopulmonary:** Includes anticoagulants, blood pressure medication, antiarrhythmics, cholesterol medications, antithrombotic.
- **Antipsychotics:** As benzodiazepines, antipsychotics and sedatives, including lithium and methylphenidate.
- **Antidepressants:** Includes TCA's, SSRI's, SNRI's, MAO-inhibitors
- **Opioids:** Includes both medical used opioids like morphine and drugs of abuse like heroin
- **Analgesics:** Includes Paracetamol and NSAID's
- **Antidiabetics**: Includes insulin, metformin, sulfonylurea derivatives, SGLT2-inhibitors
- **Alcohol:** Non-Ethanol alcohols, including methanol and ethylene glycol.
- **Ethanol:** Drinks containing ethanol
- **Street Drug:** Including cocaine, amphetamine, GHB, 2-CB, 3-MMC, Cannabis
- **Chemicals:** Non-drug chemicals, e.g. industrial chemicals, cleaning products, pesticides and the 'Other Toxins' type: Arsenic, carbon monoxide, cyanide
- **Toxins not otherwise specified:** Exposure does not belong to any of the categories above
- **Two or more types:** A combination of 2 or more types above
- **Unknown:** No exposure data was available

Data analysis

The INTOXICATED study encompasses a substantial amount of data for each patient, resulting in a large number of variables. However, conducting an analysis with such a vast number of variables may pose challenges for the Generalized Linear Mixed Model (GLMM), and more accurately the Generalized Regression Mixed Model (GLMER), which is part of GLMM. To address this issue, the analysis will focus solely on the variables that have been previously identified as important based on relevant studies. These variables have demonstrated significance in relation to the research question, and their relevance has been established through previous investigations. Furthermore, the selection of these specific variables is guided by the principle of avoiding overly specific or context-specific findings.

Assumptions of the model:

For Generalized Linear Mixed Models (GLMM) the assumptions that must be fulfilled are the follow (Schielzeth, et al., 2020; Grilli & Rampichini, 2015; Brown, 2021):

- Validity of the model
- Independence of the data points
- Absence of measurement error in the predictor
- Occurrence of data missing completely at random
- Linearity of the relationship between predictor and response and collinearity between variables. This can be illustrated by scatterplots showing a linear or curvilinear relationship.
- Normality of the residuals (Q-Q plots). Mixed models are flexible, in principle, in their use of various distributions (Lee & Nelder, 2004), but normal distributions are by far the most commonly used.

- Homogeneity of the residuals. Plotting residuals against fitted values allows an assessment of heteroscedasticity—although again observing the predicted pattern does not guarantee the absence of heteroscedasticity (Grilli & Rampichini, 2015).

As per the definition of GLMER, the dependent variable in our study is appropriately categorized as "poor outcome," aligning with the binary nature required for this kind of analysis. The outcome variable has two levels, indicating the presence or absence of a poor outcome, making it suitable for this type of modelling. Additionally, the study design ensures the independence of variables. Each patient's data is unique and distinct, with no duplicated measurements or observations, thus preserving the necessary independence of variables for accurate analysis.

To ensure the absence of measurement errors in the predictors, we diligently corrected values and units before proceeding with the analysis. To address potential biases from missing data, the study's design was carefully crafted, and thorough data cleaning was performed to identify and rectify any inconsistencies or errors. Additionally, multiple imputation techniques were employed to estimate missing values based on observed data patterns, significantly reducing bias and improving the statistical efficiency of the analysis. Moreover, the use of mixed-models in this study proves advantageous in handling missing data compared to general linear models, further enhancing the robustness of the results.

In order to assess the presence of collinearity among the continuous variables intended for inclusion in the model, a pair-wise correlation analysis was conducted (refer to the provided FIGURE displaying the collinearity**) (Figure 2).** This analysis aimed to evaluate the strength of the relationships between these variables and identify any potential multicollinearity issues that may arise from high correlations among them, which may cause faulty results in the model.

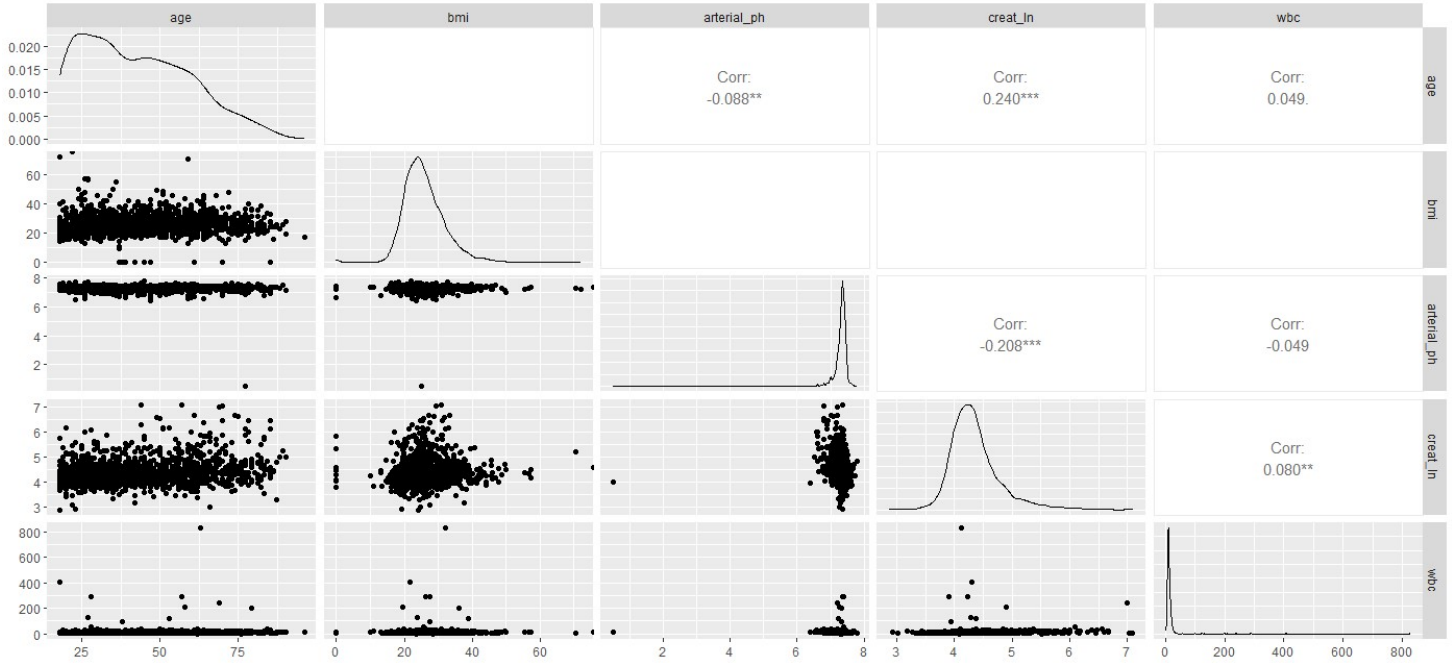


Figure 2: Pair-wise correlation analysis of the continuous variables to check collinearity of the data.

According to the literature on collinearity, for two variables to be considered collinear, their correlation must be close to 1 or -1. In our study, the highest correlation is observed between creatine levels and age, with a value of 0.240, while the lowest correlation is between arterial pH and white blood cells, with a value of -0.049. Based on these results, it can be assumed that there is no strong collinearity between the variables in this study. However, it's important to note that the

presence of missing values, particularly in arterial pH and white blood cells (WBC) variables, may influence the correlation values when imputed data is introduced in the analysis. This highlights the need to handle missing data appropriately and be cautious in interpreting the correlation results in the presence of imputed data.

Boxplots were generated to compare the distribution of continuous variables used in the study between the two "poor_outcome" groups (presence or absence of a poor outcome) **(Figure 3**). The boxplots reveal the variability and central tendency of each variable across the two groups. In order to better visualize certain variables, a log transformation was applied to the "value" variable for some of the variables. The log transformation can help address skewness in the data and highlight patterns that might not be apparent in the original scale.

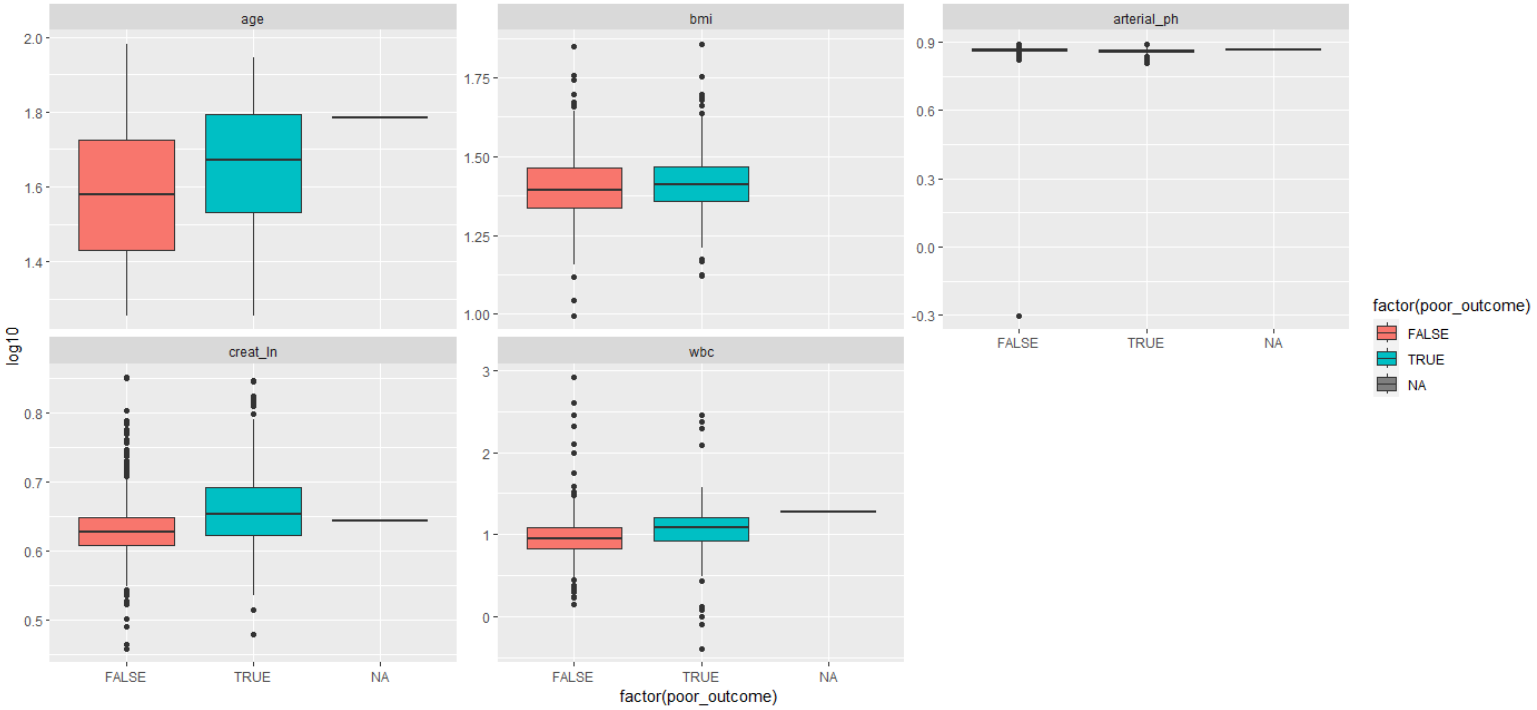


Figure 3:The distribution of continuous variables at each level of the binary outcome (poor outcome).

The analysis of the boxplots reveals some interesting patterns between patients with a poor outcome and those with a good outcome. Specifically, there are slight differences in the distribution of age and creatine levels among patients who died compared to those who survived. However, no significant variations are observed for BMI, white blood cell count (WBC), and arterial pH between the two groups. These findings suggest that age and creatine levels might play a role in distinguishing between poor and good outcomes, while BMI, WBC, and arterial pH appear to have similar distributions across both patient groups. It is important to note that these observations are based on the data visualizations, and further statistical analyses would be required to confirm the significance of these differences. (**Appentix: 1.Distribution of variables)**

The normality of the residuals, were checked using the Q-Q plots for continuous variables used in the model, while their distribution using histograms. The Q-Q plots show a normal distribution for age, arterial pH and creatinine (***Figure 4***). The rest variables shown in Appendix, especially the binary and categorical variables.

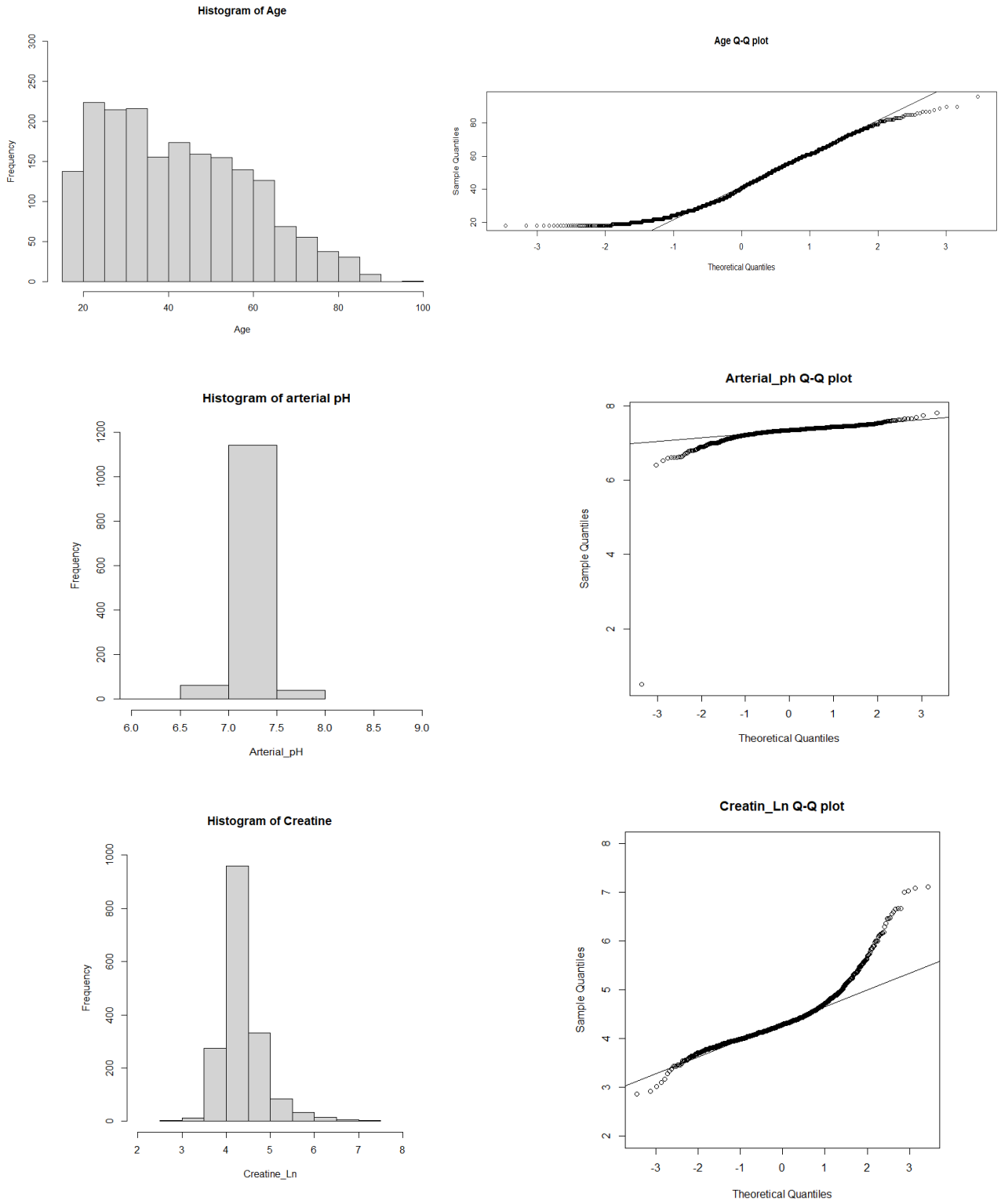


Figure 4: *The normality of the residuals and the distribution of the variables Age, Arterial Ph and Creatine Levels. On the left are the histograms of the distribution, and on the right the QQ plots for normality check.*

The homoscedasticity assumption in our analysis implies that the discrepancies between observed and predicted values should exhibit a roughly constant spread across all levels of the independent variables. In simpler terms, this means that the variability or dispersion of residuals should remain

consistent, without any discernible pattern as we traverse the range of predictor values. However, as depicted in ***Figure 5*** we observe a violation of the homoscedasticity assumption. This suggests that certain variables may be contributing to the inconsistent spread of residuals. To address this issue, further examination and potential transformations of these variables are warranted.

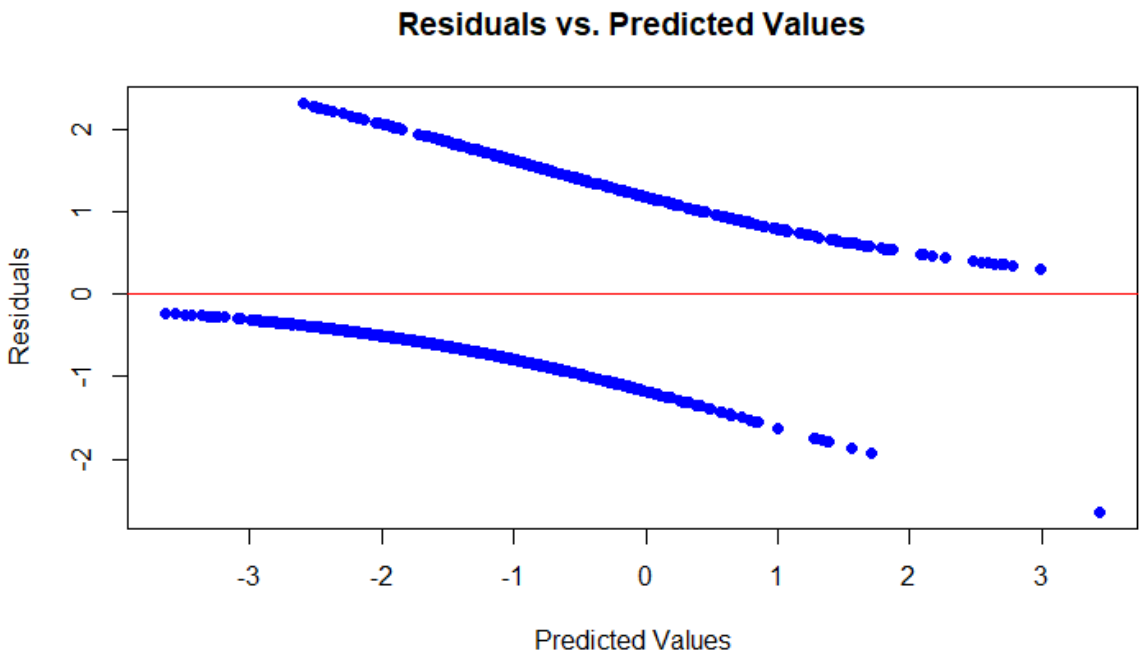


Figure 5: *The Homoscedasticity of the residuals. The dispersion of the residuals across the range of the the predicted values.*

Data extraction and cleaning

The data from Castor EDC was imported into the R programming language using the 'CastoRedc' R-package, developed by Castor EDC. In the study, certain variables in the dataset allowed site investigators to input data using multiple different units through the Castor interface. However, for the purpose of data analysis, all values were standardized and converted to a consistent unit to ensure uniformity and comparability. During the study, a thorough data validation process was carried out by a team member who carefully reviewed all data entries in Castor. If any typing errors, incorrect units, or other mistakes made by the site investigator were identified during data entry into the database, a query was raised to notify the investigator for necessary corrections. Data points flagged by queries underwent careful evaluation, and if they were deemed to be highly improbable or erroneous, they were excluded from the dataset during the data analysis phase.

Imputations of missing data

The use of clinically captured data for research often encounters the challenge of missing data, potentially introducing bias or affecting analytical results (Sterne, et al., 2009; Li, Stuart, & Allison, 2015). Missing data can arise from factors such as inconsistent staff documentation in electronic health records, technical malfunctions in data capture devices, and data stored in unstructured formats (Li, et al., 2014). As a result, these data are not readily accessible for analysis (Hegde, et al., 2019). One way to deal with missing values is to exclude the incomplete data from the subsequent analysis [9]. Excluding incomplete data from analysis is common (Manly & Wells , 2014), but widely considered suboptimal (Masconi, Matsha, Erasmus, & Kengne , 2015). This leads to information loss

and potential bias (Hegde, et al., 2019). Various techniques exist for handling missing data, but in this study, we will employ the Multiple Imputation method (MI).

Multiple imputation is a statistical technique employed to handle non-responses in surveys, which can potentially jeopardize the validity of survey results and subsequent statistical inferences (RUBIN, 1976)Rather than replacing missing values with a single value, this method utilizes the distribution of observed data or variables to estimate multiple potential values for the data points. This approach takes into account the uncertainty surrounding the true values, resulting in more robust and informative values for more comprehensive analysis. The multiply-imputed datasets can be analysed using standard complete-data methods, treating them as if they were the actual dataset, despite not containing any real data from nonrespondent individuals. (LITTLE & RUBIN, 1989).

The 'Multiple Imputation Using Chained Equations' (MICE) technique, which is an example of a MI (AZUR, STUART, FRANGAKIS, & LEAF, 2011; Raghunathan, Lepkowski, Van Hoewyk, & Solenberger, 2001), was chosen for the INTOXICATE study Our study utilized MICE operates with the assumption that the missing data is Missing At Random (MAR), implying that the probability of data being missing is solely dependent on observed values used in the imputation process, and not influenced by unobserved values (Schafer & Graham, Missing data: our view of the state of the art, 2002).The percentage of missing data points for each variable used is shown in ***Table 2***. The missing data points were replaced with substituted values through 100 imputation iterations and five imputed datasets. The multiple imputation process was conducted using the 'MICE' R-package (Zhang, 2015; Wulff & Jeppesen, 2017).

Table 2: *Percentage of Missing Data for Variables Used in Multiple Imputation*

Variables	Percentage of missing data
Arterial_ph	0.35%
Wbc	0.17%
Low_platelet	0.10%
Creatinine_ln	0.1%
Bad_potassium	0.08%
Bad_sodium	0.08%
GCS	0.05%
Bad_temp	0.02%
Low_sbp	0.01%
High_hr	0.007%
Sex_binary	0.004%
Poor_Outcome	0.001%

Model Selection

The INTOXICATE study includes data from various countries worldwide, with diverse protocols in managing intoxicated patients. Assessing variations in the relationship between exposure types and poor outcomes across countries is crucial to contextualize findings and consider the impact of diverse healthcare systems on research outcomes. In this study, binary, continuous, and categorical variables were employed to evaluate a Generalized Linear Mixed Model (GLMMs), aiming to investigate the potential differences in poor outcome associated with exposure types across different countries.

GLMERs extend the capabilities of GLMM, making them suitable for handling response variables from various distributions, including binary outcomes. These models are especially valuable when dealing with data that lacks independence due to a hierarchical structure. These models incorporate both fixed and random effects, providing flexibility in analysing different types of data. Fixed effects represent population-level effects that remain consistent across experiments, while random effects capture clusters of dependent data points with observations belonging to the same higher-level group (Brown, 2021). These random effects are discrete units sampled from a larger population (Winter & Wieling, 2016). GLMERs address issues like over-dispersion and account for population heterogeneity, making them valuable for statistical analysis (Takele, Zewotir, & Ndanguza, Understanding correlates of child stunting in Ethiopia using generalized linear mixed models, 2019).

$$y = X\beta + Zu + \varepsilon$$

Equation 1: The GLMER model equation. Y *is the outcome variable, X is a matrix of the predictor variables, β is the vector for fixed effects regression coefficients, Z is the matrix for the random effects, u is the vector for the random effects and ε is the vector of the residuals that are not explained by the model.*

In our study, in the ***Equation 1***, the dependent variable (Y) is defined as the binary variable representing the occurrence of the poor outcome. The independent variables encompass various factors, such as country, exposure types, laboratory results and clinical assessment findings. The independent variables in the study were categorized into two types: fixed effects (X) and random effects (Z). Fixed effects consisted of exposure types, clinical assessment variables and laboratory results. In our study, we treated countries as random effects. This decision was driven by the use of a specific subset of countries instead of all possible options. This approach allowed us to factor in the variability associated with different countries and consider how country-related factors might impact patient responses.

Potential confounders, including age, somatic or/and psychiatric comorbidities, and gender, were considered in the model. Confounding variables are associated with both the independent and dependent variables, meeting two conditions: correlation with the independent variable (causality not necessary) and causal relationship with the dependent variable. Addressing confounding variables ensures internal validity, as neglecting them may introduce bias and distort relationships between other variables. Age, for example, does not change due to exposure types, but it could influence the exposure type a patient experiences (e.g., younger individuals being more likely exposed to street drugs), making it a potential confounder if it also affects the outcome

The initial step involved identifying the variables used in constructing the model. The country was considered a constant random effect in the model. Subsequently, these variables were organized into three distinct blocks, facilitating the investigation into whether the exposure type can effectively distinguish between poor and good outcomes. The first block comprises physiological parameters that a physician can determine upon encountering an acutely intoxicated patient. The second block includes laboratory parameters that require blood tests for measurement. The third block consists of information about the exposure type (**Table 3**).

Table 3: *Blocks of variables used in the models.*

BLOCK	VARIABLES
1#	Age, Gender, Comorbidity category, Tachycardia (High_hr), Low systolic blood pressure (Low_spb), Glasgow Coma Score (GCS)
2#	Thrombocytopenia (Low_platelet), Arterial pH, White blood cells concentration (wbc), Creatine concentration (creatinine_ln), hypo-/hyperkalemia (bad_potassium), hypo-/hypernatremia (bad_sodium), hypo-/hyperthermia (bad_temp)
3#	Exposure types: Alcohol, Analgesics, Anti-diabetics, Antidepressants, Cardiopulmonary, Ethanol, Opioids, Street drugs, Chemicals, Toxins not otherwise specified, Unknown

Initially, we developed a model using covariates from the first block, which included patient characteristics. Next, we constructed a second model, incorporating covariates from both the first block and the second block, which consisted of laboratory results. Subsequently, the third model was constructed, incorporating covariates from both the first block (patient characteristics) and the third block (exposure categories). In this way, we assessed whether the inclusion of the third block improved the model fit compared to the models with covariates solely from the first block or the first and second blocks. Lastly, the fourth model was built, incorporating all three blocks, to examine whether the model with exposure types enhanced the goodness-of-fit compared to the models with the other covariates. Below on ***Table 4*** are the models developed in the study:

Table 4: *Models used in the analysis.*

MODEL	FIXED EFFECTS	RANDOM EFFECTS
1#	Block 1#	Countries
2#	Block 1# + Block 2#	Countries
3#	Block 1# + Block 3#	Countries
4#	Block 1# + Block 2# + Block 3#	Countries

As we navigate the intricacies of the GLMER model, a crucial decision arises—choosing between random/fixed intercepts and random/fixed slopes. In mixed modelling, the fixed-intercept estimate represents the average starting point, while random intercepts let individuals and items deviate from this average. Here, we've utilized random intercepts to uncover variations in baseline poor outcomes among subjects. This approach suggests that patients in each country have distinct starting points for poor outcomes, following a typical distribution of deviations with a zero mean and a variance estimated by the model (Brown, 2021). In contrast, a random slope model enhances this versatility by allowing explanatory variables to exert differing effects within each group, acknowledging the potential for distinct relationships between variables and responses across subsets. In our current study, we've embraced a model with fixed slopes. This choice implies that each explanatory variable

maintains a consistent impact across all patient groups, ensuring uniformity and enabling meaningful comparisons among countries in our analysis (University of Bristol).

Comparisons of the models

To compare the models and determine the parameters that best explain our data, we utilized the likelihood-ratio test through the anova() function in R. This test compares the goodness of fit between two nested models. A nested model contains a subset of the independent variables present in the full model. In our case, we compared the subset models (1, 2, 3) with the full model (4) to identify the one that provides a better explanation for our data. For the comparison of non-nested models, such as model 2 and model 3, which do not contain nested variables, the Akaike Information Criterion (AIC) is utilized.

The Likelihood ratio test is based on the following null and alternative hypotheses: **H0:** Both the full model and the nested model provide an equal fit to the data. Consequently, the nested model should be preferred for use. **HA:** The full model exhibits a significantly superior fit to the data compared to the nested model. Consequently, the full model should be favoured for utilization. If the p-value of the test is below a specific significance level (e.g., 0.2) (Steyerberg, 2019), we can reject the null hypothesis and infer that the full model provides a significantly better fit to the data. (Glen, 2023). The likelihood-ratio test was performed on the 5 imputed datasets, which were combined using the R package 'mice.'

The AIC is a mathematical approach for assessing how well a model fits the data from which it was derived. In statistics, AIC is employed to compare various models and identify the one that provides the best fit for the data. It takes into account the number of independent variables used in constructing the model and the maximum likelihood estimate, which measures how well the model replicates the data. The preferred model according to AIC is the one that explains the most variation while utilizing the fewest independent variables. Lower AIC scores are preferable as AIC penalizes models with more parameters. When two models explain the same amount of variation, the one with fewer parameters will have a lower AIC score and is considered the better-fit model (Cavanaugh & Neath, 2019; Bevans R. , 2023).

Generalized Linear Mixed Models VS Generalized Linear Model

The secondary objective of this study is to evaluate whether Generalized Linear Mixed Models offer a more comprehensive explanation for the relationship between exposures and poor outcomes across different countries compared to the Generalized Linear Models used previously. To achieve this, we will employ the AIC (Akaike Information Criteria) and likelihood-ratio test to compare the two types of models. The AIC is calculated from (1) the number of independent variable that used to build the model and (2) the maximum likelihood estimate of the model. The best-fit model according to AIC is the one that explains the greatest amount of variation using the fewest possible independent variables (Bevans R. , 2023).

Another method to compare the goodness of fit two regression models is the Bayesian Information

$$BIC=(RSS+ \log(n)d\hat{o}2) / n$$

Equation 2: *The BIC criteria for comparison of regression models. d: The number of predictors, n: Total observations, ô: Estimate of the variance of the error associate with each response measurement in a regression model, RSS: Residual sum of squares of the regression model, TSS: Total sum of squares of the regression.*

Criterion (BIC). BIC is calculated from ***Equation 2***

The (BIC) is more useful in selecting a correct model while the AIC is more appropriate in finding the best model for predicting future observations (Chakrabarti & Ghosh, 2011). Both models will encompass identical independent variables, covariates, and dependent variables, with the GLMM model uniquely incorporating the random effect of countries. This analysis aims to determine which model provides a more robust understanding of the association between exposures and poor outcomes, accounting for potential variations among different countries.

Results

Variables

The age range of the patients enrolled in this study spans from 18 to 96 years old. The average age of the patients across the study cohort is 54.57 years. Notably, patients who experienced poor outcomes due to intoxication had a marginally higher mean age of 48.19 years compared to those who survived. A gender breakdown reveals that the study population consisted of 1024 females and 877 males. Interestingly, a greater proportion of females participated in the study. Specifically, 28.77% of males exhibited poor outcomes, whereas a slightly lower proportion, 18.46%, of females experienced poor outcomes. **Table 5** presents the patient demographics from the study, illustrating the average values (expressed as percentages) of the variables associated with poor or bad outcomes. These variables were subsequently utilized for the purpose of statistical analysis. Examining patients with poor outcomes reveals a distinct pattern: a higher occurrence of somatic comorbidities in contrast to psychiatric conditions. Interestingly, patients with both types of comorbidities exhibit an intermediate position. Furthermore, individuals who experienced poor outcomes exhibited noteworthy differences: their Glasgow Coma Scale (GCS) scores were notably lower, arterial pH levels were reduced, creatinine levels were higher, and white blood cell (WBC) counts were elevated.

Table 5: *Comparison of how various factors are distributed among three distinct patient groups: all patients, patients with good outcomes, and patients with poor outcomes. The p-values associated with continuous variables highlight differences in average values between patient groups, while p-values for categorical variables highlight associations between the variable and the likelihood of poor outcomes.*

Variables	**All patients** *(1907)*	**Good Outcome** *(N=1464)*	*Poor Outcome* *(N=443)*	**P -value**
Age (mean)	54.57	40.93	48.19	<0.00
Gender (Male/Female)	46.13/53.87 (N=877/N=1024)	71.23/81.54 (N=624/N=835)	28.77/18.46 (N=252/N=189)	0.05
Comorbidity (Somatic/Psychiatric/Both)	52.10/29.03/18.87 (N=554/N=994/N=360)	68.72/81.49/76.11 (N=380/N=810/N=274)	31.28/18.51/23.89 (173/184/86)	<0.0307
Glasgow Coma Score (GCS)	mean=9.71, median=10	mean=10.25, median=12	mean=7.93, median=7	<0.00
Low Systolic Blood pressure (SBP) (<90 mmHg)	18.72 (N=353)	60.34 (N=291)	39.66 (N=140)	0.1659
Tachycardia (Hr>109)	32.98 (N=625)	74.08 (N=463)	25.92 (N=162)	0.08252
Hypo-/hyperthermia *(<35 or >38.3 ℃)*	13.84 (N=259)	59.07 (N=153)	40.93 (N=106)	0.1882
Arterial Ph	mean=7.31, median=7.34	mean=7.33, median=7.36	mean=7.24, median=7.28	<0.00
Creatinine Levels (µmol/L)	mean=4.37, median=4.28	mean=4.29, median=4.25	mean=4.63, median=4.51	<0.00
Thrombocytopenia *(<150 x 109/L)*	6.67 (N=114)	77.29 (N=74)	22.71 (N=40)	0.222
White blood cells (10^9/L)	mean=12.21, median=9.49	mean=11.34, median=8.9	mean=14.97, median=12.12	0.0103
Hypo-/hyperkalemia *(<3.0 or > 5.0 mmol/L)*	11.36 (N=199)	47.23 (N=94)	52.76 (N=105)	0.2092
Hypo-/hypernatremia *(<135 or > 145 mmol/L)*	17.19 (N=303)	66.34 (N=201)	33.66 (N= 102)	0.1656

The risk of poor outcomes across diverse countries in the study was evaluated using a logistic regression model: glm(poor_outcome ~ country, data = data, family = binomial). As depicted in **Figure 6**, Brunei demonstrates the highest estimated risk for poor outcomes (0.9), though this could be influenced by limited patient participation. Following Brunei, Libya, Romania, and Egypt exhibit heightened risks of poor outcomes among their patients. In contrast, Spain and Belgium, with substantial patient representation, also reveal notable risks of poor outcomes, approaching 0.70. These findings suggest the persistence of distinct outcomes, irrespective of variations in patient participation rates among countries

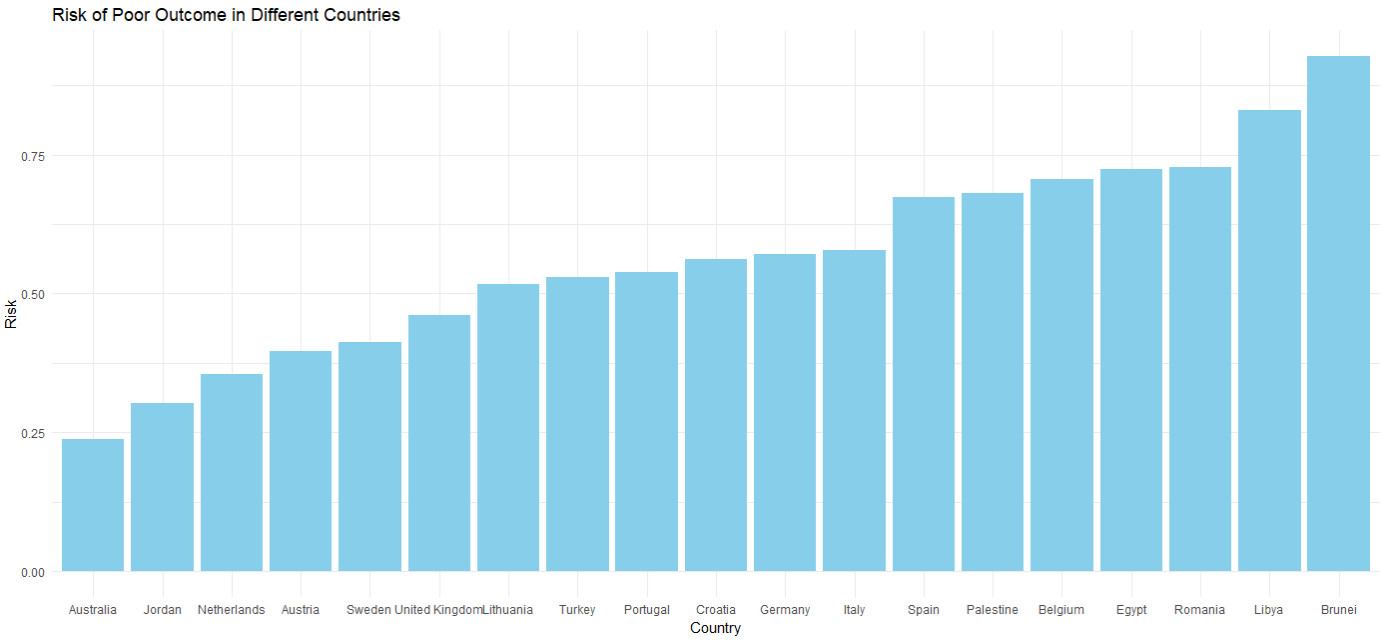


Figure 6: The plot that illustrates the risk of poor outcomes for each country included in the study. The risk assessments are derived from a simple logistic regression model, providing valuable insights into country-specific outcome patterns.

To reinforce the findings discussed earlier, **Table 6** and **Table 10** (Appendix: 2.Distribution of Exposure Variables in different countries) offer a more detailed examination of key variables, showcasing their variations in both average values and percentages. These tables are particularly relevant as we delve into the upcoming models. **Table 6** centres on patients' characteristics, lab results, and clinical assessments, whereas **Table 10** focuses on exposure-related variables. Together, they dissect these distinctions across five specific countries that played a role in our study.

Differences in parameter distributions among Belgium, Turkey, Spain, Sweden, and the Netherlands are evident. For instance, Belgium exhibits a higher percentage of tachycardia among patients compared to the Netherlands. Similarly, when it comes to somatic comorbidities, Turkey stands out with nearly 17% of its population affected, while only 5% of the Netherlands' population has somatic comorbidities. These disparities underscore the importance of considering regional variations in patient characteristics when conducting analyses and drawing conclusions.

***Table 6**: The distribution of patient characteristics, clinical assessment criteria, and laboratory results variables across five participating countries in the study. The results are presented as percentages or mean/median values for the number of patients in each country.*

VARIABLES	ALL PATIENTS (N=1907)	BELGIUM (N=114)	TURKEY (N=119)	SPAIN (N=214)	SWEDEN (N=200)	NETHERLANDS (N= 873)
Age (mean)	42.63	48.55263	38.07563	45.65888	40.89	41.71134
Gender (Male/Female) (%)	46.13(N=877)/ 53.87 (N=1024)	44.74(N=51)/ 55.26 (N=63)	52.94(N=63)/ 47.06(N=56)	54.72 (N=116)/ 45.28(N=97)	42.14(N=83)/ 57.86 (N=114)	40.37(N=352)/ 59.63(N=52)
Comorbidity (Somatic/Psychiatric/Both/No Comorbidity) (%)	8.17(N=156)/ 52.09(N=994)/ 20.86(N=398)/ 18.87(N=360)	5.26 (N=6)/ 52.63(N=60)/ 20.18 (N=23)/ 21.93(N=25)	16.81(N=20)/ 30.25(N=36)/ 15.13(N=18)/ 37.82(N=45)	8.88(N=19)/ 44.86(N=96)/ 24.77(N=53)/ 21.50(N=46)	7(N=14)/ 64.5(N=129)/ 17.5(N=35)/ 11(N=22)	4.47(N=39)/ 59.79(N=522/ 20.39(N=178/ 15.35(N=134)
Glasgow Coma Score (GCS)	mean=9.71, median=10	mean=9.79, median=11	mean=12.25, median=14	mean=8.4, median=8	mean=8.90, median=8	mean=10.03, median=11
Low Systolic Blood pressure (SBP) (<90 mmHg) (TRUE) (%)	18.71 (N=353)	20.54 (N=23)	21.93 (N=25)	23.33 (N=49)	23.47 (N=46)	15.67 (N=136)
Tachycardia (Hr>109) (TRUE) (%)	32.98 (N=625)	35.09 (N=40)	35.96 (N=41)	31.75 (N=67)	38.38 (N=76)	30.11 (N=262)
Hypo-/hyperthermia (<35 or >38.3 °C) (TRUE) (%)	13.84 (N=259)	18.75 (N=21)	0.88 (N=1)	15.64 (N=33)	19.70 (N=39)	12.90 (N=110)
Arterial Ph	mean=7.31, median=7.34	mean=7.34, median=7.35	mean=7.29, median=7.35	mean=7.30, median=7.33	mean=7.32, median=7.34	mean=7.33, median=7.36
Creatinine Levels (µmol/L)	mean=4.37, median=4.28	mean=4.38, median=4.33	mean=4.24, median=4.15	mean=4.40, median=4.30	mean=4.37, median=4.23	mean=4.34, median=4.27
Thrombocytopenia (<150 x 109/L) (TRUE) (%)	6.67 (N=114)	4.50 (N=5)	14.56 (N=15)	9.66 (N=20)	6.15 (N=12)	4.84 (N=38)
White blood cells (10^9/L)	mean=12.21, median=9.49	mean=12.50, median=10.4	mean=11.83, median=10.44	mean=10.62, median=9.9	mean=11.75, median=8.55	mean=11.63, median=8.9
Hypo-/hyperkalemia (<3.0 or > 5.0 mmol/L) (TRUE) (%)	11.36 (N=199)	6.48 (N=7)	16.50 (N=17)	12.50 (N=25)	11.22 (N=22)	8.50 (N=69)
Hypo-/hypernatremia (<135 or > 145 mmol/L) (TRUE) (%)	17.19 (N=303)	11.82 (N=13)	18.45 (N=19)	18.63 (N=38)	18.97 (N=37)	14.41 (N=119)

Analysis of **Table 10** (Appendix: 1.Distribution of variables) reveals significant disparities in the distribution of exposure categories within the study population. For instance, the "Street_drug" category has no patients in Turkey and Belgium, whereas approximately 5% of patients in Spain and the Netherlands have experienced street drug intoxication. Conversely, the "2_or_More" category stands out with a higher number of patients compared to other categories in all countries. This prevalence of the "2_or_More" category may lead to inaccurate exposure estimations. It's noteworthy that a patient exposed to both "Street drug" and "Alcohol" is categorized under "2_or_More," rather than in both "Alcohol" and "Street drug." This raises questions about the accuracy of exposure categorization, highlighting the need for further investigation in this area we can conclude that the exposures categories are not equally distributed among the population of the study.

Multiple Imputations (MI)

To assess the missing data before the statistical analysis, multiple imputation using the mice() package were contacted with 100 iterations and 5 different imputed datasets. **Figure 7**, shows the density plots of the real data (blue line) across the imputed data (red lines) for patient vitals during the ICU stay and their laboratory results. The graphs show that the imputed and real data trends are quite similar. Most imputed curves follow the same patterns with the real data, however, some imputed values have higher or lower data distribution. It's interesting to note that heart rates and systolic blood pressure (SBP) values often go above the usual range of the real data, while the sodium levels as well as temperature and platelet values follow under the curve of real data. The rest data shown in Appendix: 3. Multiple Imputations

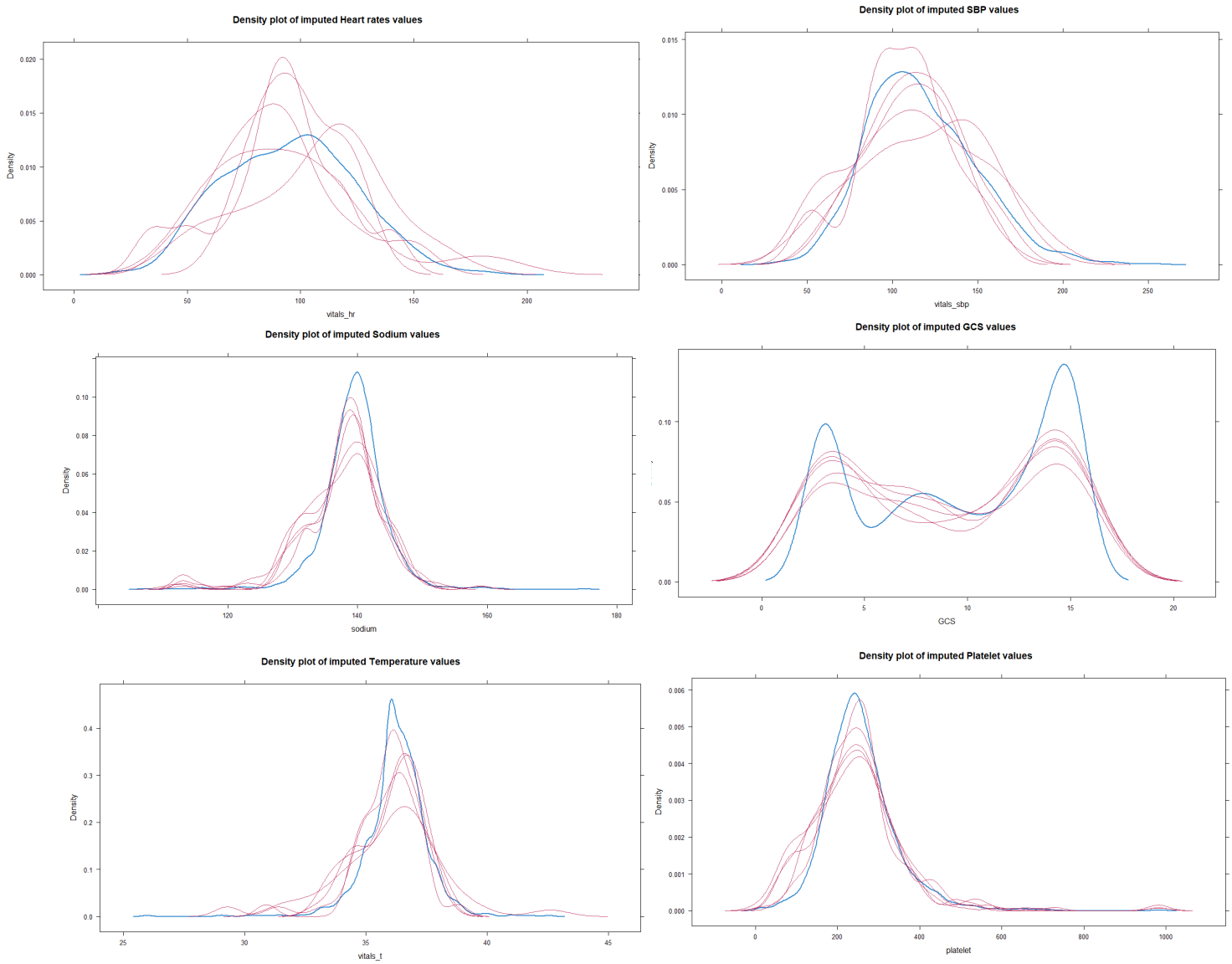


Figure 7: *The density plots of comparison of imputed data (depicted in red) with real data (shown in blue) pertaining to vital signs (heart rates, systemic blood pressure, temperature and GCS) and laboratory results (platelet counts, sodium levels.*

Detecting the convergence of the MICE algorithm can be a challenging task, often requiring specific diagnostic methods. One approach is to visualize the behavior of imputed parameters against the iteration number. In **Figure 8**, the mean and variance of imputations for each parallel stream are displayed. When convergence is achieved, these streams should blend together seamlessly, without displaying distinct patterns. Convergence is typically established when the variance across different sequences is comparable to or smaller than the variance within each individual sequence. This criterion indicates that imputed values from different streams have coalesced into a consistent result, implying that the imputation process has reached a stable state (Abayomi, Gelman, & Levy, 2007). The rest data shown in Appendix: 3. Multiple Imputations .

As shown in the figure, the differences in imputed values across various datasets don't follow any

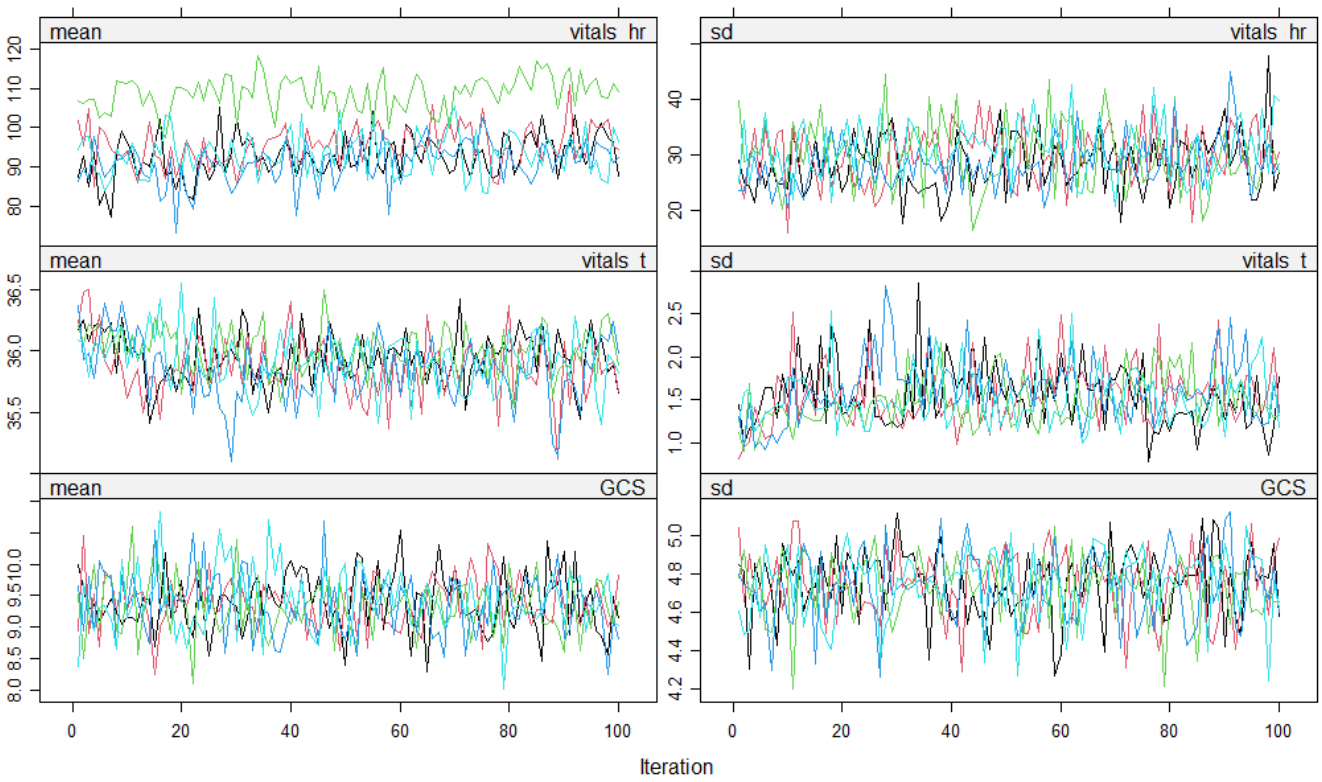


Figure 8: *Mean and standard deviation of the synthetic values plotted against iteration number for the imputed data.*

noticeable patterns and fit together smoothly. This indicates that the imputation process has successfully reached a stable point. When we tried using fewer iterations, the imputed values didn't work well and didn't show consistent patterns. This is why we decided to increase the number of iterations from 10 to 100. This change led to better imputation results, where the values became more reliable and consistent.

Model selection

The process of selecting the most effective model for explaining outcomes in each country commenced with the development of four distinct models (Appendix: 4. Model Selection). The initial model **Table 11** characterized by its simplicity, integrated patients' physiological attributes, encompassing age, gender, and comorbidities.

On the other hand, the second model **Table 12**, marked by its complexity, not only combined these physiological parameters but also incorporated laboratory results, all while accounting for age and gender variables. The comparison of these two models using the likelihood-ratio test yields a p-value less than 0.001 (**Table 7**). This outcome suggests that including the laboratory results in the analysis enhances the model's overall fit. Notably, the data seems to be derived from a model that is more intricate than one based solely on physiological parameters.

Moreover, as we proceed to delve into more intricate models, the analysis will incorporate the exposure variables. In Model 3, the scope broadens to encompass the physiological parameters from Block 1, along with age, gender, and comorbidities. Subsequently, in Block 3, we integrate the distinct exposure categories into our analysis ***Table 14***. When subjecting Model 1, our initial model, to a likelihood-ratio test, the resulting p-value registers at less than 0.001 (***Table 7)***. This pattern holds true for the comparison between Model 3 and Model 2, yielding p-values less than 0.001 once again . This compelling outcome underscores the augmentation of the model's fit through the inclusion of exposure variables in the analysis. Importantly, these results imply that the data aligns more closely with a model of greater complexity—one that not only hinges on physiological parameters or laboratory results but also incorporates the intricate interplay of exposure variables.

Table 7: *The comparisons of the models use to conclude in the variables selection that included in the "Current" final model. The p- value were retrieve after comparisons of the models fit with Likelihood Ratio Test.*

Model	Data	Fixed Effects	Random Effect	Average AIC	Comparisons	Anova Test
1#	Imputed	Block 1	Country	1797.162	Simpler Model	-
2#	Imputed	Block 1+ Block 2	Country	1726.652	Model 1	P-value< 0.001
3#	Imputed	Block 1+ Block 3	Country	1786.351	Model 1	P-value< 0.001
4#	Imputed	Block 1+ Block 2+ Block 3	Country	1728. 022	Model 1	P-value< 0.000

Additionally, as we proceed to make the most complex model, the Model 4, the analysis will incorporate the exposure categories (Block 3) along with the laboratory results (Block 2), and physiological parameters (Block 1) ***Table 13***. When subjecting Model 1, our initial model, to a likelihood-ratio test, the resulting p-value registers at less than 0.000 (***Table 7***). This compelling outcome strongly emphasizes the improvement in the model's accuracy when including all variables, making the model more complex. Similar trends are observed when comparing Model 4 to both Model 3 and Model 2, with p-values consistently below 0.001 and 0.02, respectively. These results reinforce the idea that the data is better explained by a more intricate model. Such a model not only considers physiological and laboratory aspects but also takes into account the intricate interactions of exposure variables.

Finally, our "current" final model, supported by a analysis incorporating imputed data and variable selection, offers a comprehensive view of the key factors influencing the outcome **Table 8**. Across the

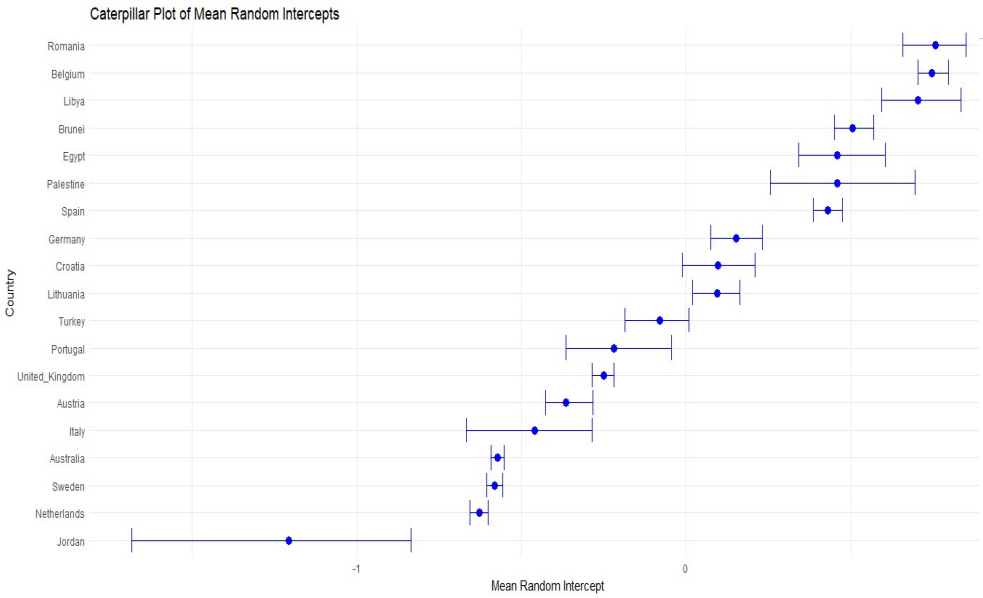
five imputed datasets, it is calculated AIC (Akaike Information Criterion) values as follows: 1698.059, 1707.825, 1720.260, 1714.664, 1714.138. By considering these multiple datasets, we enhance the reliability and generalizability of our conclusions. In **Table 8** showcases the pooled data, presenting the mean estimates of the 24 variables utilized in the model. This aggregation gives us a more stable and accurate representation of the relationships between these variables and the outcome. The synthesized results from the five imputed datasets provide a comprehensive view that accounts for potential variability.

Table 8: *Results of "Current"Final Model with selection of variables from the 3 different blocks made with the imputed datasets.*

	TERM	ESTIMATE	STD.ERROR	STATISTIC	DF	P.VALUE	2.50%	*97.50%*
1	(INTERCEPT)	214.492	4.938	1.087	16.889	0.292	0.006	7221874.780
2	AGE	1.018	0.004	4.251	1739.552	0.000	1.009	1.026
3	SEX_BINARY2	0.732	0.131	-2.373	1574.320	0.018	0.566	0.947
4	GCS	0.910	0.015	-6.160	923.712	0.000	0.884	0.938
5	COMORBIDITY_CATEGORYPSYCHIATRIC	1.176	0.181	0.895	1655.417	0.371	0.825	1.676
6	COMORBIDITY_CATEGORYSOMATIC	1.397	0.195	1.716	1687.034	0.086	0.953	2.046
7	LOW_SBPTRUE	1.678	0.153	3.376	1485.102	0.001	1.242	2.266
8	HIGH_HRTRUE	1.352	0.136	2.219	1658.572	0.027	1.036	1.766
9	BAD_TEMPTRUE	1.718	0.171	3.169	1322.791	0.002	1.229	2.401
10	BAD_POTASSIUMTRUE	2.167	0.198	3.899	158.688	0.000	1.465	3.207
11	CREAT_LN	1.633	0.130	3.779	224.559	0.000	1.265	2.109
12	ARTERIAL_PH	0.289	0.648	-1.916	16.726	0.073	0.073	1.136
13	ETHANOL	0.610	0.413	-1.196	1814.076	0.232	0.271	1.372
14	ALCOHOL	2.929	0.451	2.383	1137.938	0.017	1.209	7.096
15	ANALGESIC	1.159	0.498	0.296	1822.179	0.768	0.436	3.079
16	ANTI_DIABETIC	0.892	0.471	-0.242	1877.001	0.809	0.354	2.248
17	CHEMICALS	2.043	0.365	1.959	1720.952	0.050	0.999	4.178
18	ANTIDEPRESSANT	1.702	0.415	1.280	1769.305	0.201	0.754	3.842
19	CARDIOPULMONARY	1.295	0.450	0.574	1679.641	0.566	0.536	3.128
20	STREET_DRUG	1.233	0.366	0.571	1857.985	0.568	0.601	2.529
21	TNOS	2.225	0.289	2.763	1592.880	0.006	1.261	3.926
22	OPIOIDS	0.838	0.397	-0.445	1801.847	0.656	0.384	1.827
23	UNKNOWN	1.785	0.382	1.518	860.419	0.129	0.844	3.775
24	TWOORMORE	1.044	0.218	0.198	1646.944	0.843	0.681	1.602

A careful selection of the key variables that offer the most explanatory power for our data, accounting for both physiological parameters, laboratory results, and exposures. The variables used in the model based on imputed data are: from Block A, we incorporated crucial physiological indicators such as age, gender, GCS (Glasgow Coma Scale), low SBP (Systolic Blood Pressure), high heart rate, and abnormal temperature readings. In Block 2, we integrated laboratory results, including elevated potassium levels, creatinine levels, and arterial pH. Finally, Block 3's Exposures provided valuable insights through the inclusion of all the exposures found in the study.

In addition, **Figure 9** provides a compelling visualization of the mean estimation of random intercepts for each country. This plot helps us understand how the model explains the variation in poor outcomes across different countries. By examining these random intercepts, we gain insights into the influence of country-specific factors on the outcome, contributing to a deeper understanding of the study's implications.



Country	MeanIntercept	LowerBound	UpperBound
Australia	-0.5726	-0.5805	-0.5643
Austria	-0.3648	-0.3839	-0.3265
Belgium	0.7448	0.7306	0.7679
Brunei	0.5046	0.485	0.5327
Croatia	0.0958	0.0538	0.1417
Egypt	0.4586	0.4202	0.5256
Germany	0.1514	0.1215	0.1841
Italy	-0.4606	-0.5538	-0.4008
Jordan	-1.206	-1.4292	-1.0903
Libya	0.7022	0.6655	0.761
Lithuania	0.0932	0.0616	0.1194
Netherlands	-0.6286	-0.64	-0.6182
Palestine	0.458	0.3864	0.5625
Portugal	-0.2206	-0.2684	-0.1405
Romania	0.7568	0.7141	0.791
Spain	0.4292	0.4109	0.4463
Sweden	-0.5828	-0.5918	-0.5726
Turkey	-0.0814	-0.1277	-0.05
United_Kingdom	-0.2506	-0.2655	-0.2387

Figure 9: *The caterpillar plot showcasing the random intercepts per country for Best Model. The plot offers insights into the variations in the intercepts across different countries within the context of the model. The data represent the mean values of the random intercepts from the five imputed datasets.*

The caterpillar plot for random effects provides a visual representation of how individual levels within a grouping variable affect the response variable in a mixed-effects model, allowing you to identify which levels have significant deviations from the overall trend. In Y-Axis we have each level of the grouping variable (countries). Each dot represents a unique level of the grouping variable. In X-Axis, the dots represent the estimated random effects (deviation from the fixed effect) for each level. This indicates how much each level's effect deviates from the overall average or fixed effect.

The caterpillar plot shows that between countries there is a variation on the random intercepts. The positive values indicated that a country has an effect grater than the fixed effects, while the negative values of intercept for the countries indicate that the country has a smaller effect. As an example, Netherlands has negative random intercept of -0.63, while Belgium has an intercept of 0.74, which shows a greater influence in the outcome. showing that including the countries as random effect was a good choice, because there is a difference in patients outcome between countries when the other variables are the same.

Levels with dots that are significantly different from zero (i.e., their error bars do not include zero) have random effects that significantly deviate from the fixed effect. This suggests that these levels have a substantial impact on the response variable beyond what would be expected based on the fixed effect alone. Levels with dots close to zero and overlapping error bars have random effects that are not significantly different from the fixed effect. This suggests that these levels do not contribute much additional variation beyond the fixed effect.

This we can also explain it if we compare a patient from Netherlands and a patient from Belgium (**Figure 10**). The #3 patient from Netherlands were female, 33 years old with both somatic&

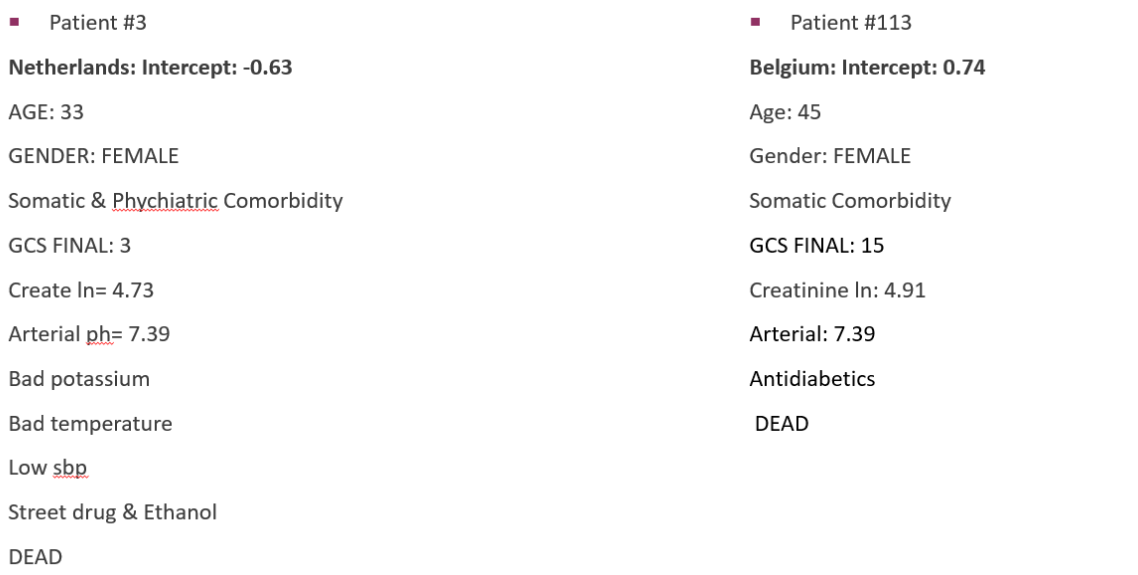


Figure 10: *Patients comparisons between Netherlands and Belgium. Both patients had poor outcome of death.*

psychiatric comorbidity. She shows bad results both in clinical assessment and lab results, after the intoxication of Street drugs and Ethanol, which conclude to her death. On the other hand, patient #113 from Belgium, also female age of 45 years old, with only Somatic comorbidity without any bad lab results, and exposure of Antidiabetics, conclude to her death. We can see that a patient with less severe situation comparing to the patient from Netherlands, in Belgium can died.

The Caterpillar Plots displayed in Appendix: 4.Model Selection offer a revealing insight into the variability of random intercepts among individual countries. These plots underscore the significance of considering the unique impact of each country, irrespective of the fixed effects, on the study's outcome. Notably, these analyses were conducted using real data, accounting for missing values without the need for imputation. This exploration highlights that each country's distinct characteristics and factors contribute to explaining the variability in patient outcomes observed in the study.

Generalized Linear Mixed Models VS Generalized Linear Model

In the initial analysis of the INTOXICATED data, we employed Generalized Linear Models (GLM) without considering the potential influence of different countries on patient outcomes. Recognizing the significance of this factor, we proceeded to compare our GLMM (Generalized Linear Mixed Model) against the GLM to determine the extent to which accounting for random effects associated with different countries enhances our understanding of the data. Both models utilized the original data without employing multiple imputations. The results of this comparison, conducted using the Likelihood Ratio Test (LRT) via the 'anova()' function and presented in ***Table 9***, are illuminating. They reveal that many of the laboratory and patient characteristic variables exhibit significant differences in the GLMM model. This underscores the importance of incorporating random effects, such as those associated with different countries, into our analysis.

Furthermore, when we compare the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and the Deviance between both models, it becomes evident that the GLMM

approach utilized in this study offers a superior fit for the data and provides a more comprehensive explanation of the outcomes compared to the previous GLM approach. As depicted in **Figure 11**, the AIC for GLM stands at 1121.017, whereas it's notably lower at 1102.747 for GLMM. Similarly, BIC is 1245.073 for GLM and 1231.765 for GLMM, while the Deviance is 1071.017 for GLM and considerably lower at 1025.847 for GLMM. These differences were evaluated using a Likelihood Ratio Test, revealing a significant improvement in model fit for GLMM with a p-value <0.05. Taken together, all three criteria (AIC, BIC, Deviance) strongly support the conclusion that the GLMM model significantly outperforms the GLM model when applied to the INTOXICATE data. This improvement in model fit and explanatory power is attributed to the inclusion of random effects, which complement the fixed effects

Table 9: *Results of comparison of the GLM (Generalized Linear Model) and the GLMM (Generalized Linear Mixed Model) with the variables selected after different simulations of GLMER models. The test use for the comparison is Likelihood Ratio Test (LRT). Variables with the Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1, means that are significant more explain in GLMER model than the GLM model.*

	DF	DEVIANCE	RESID. DF	RESID. DEV	PR(>CHI)	
NULL	NA	NA	1055	1269.517	NA	
age	1.00	16.72	1054.00	1252.80	0.00	***
sex_binary	1.00	14.46	1053.00	1238.34	0.00	***
comorbidity_category	2.00	1.77	1051.00	1236.57	0.41	
GCS_final	1.00	26.08	1050.00	1210.48	0.00	***
low_sbp	1.00	26.99	1049.00	1183.50	0.00	***
high_hr	1.00	5.94	1048.00	1177.55	0.01	*
bad_temp	1.00	7.88	1047.00	1169.68	0.005	**
bad_potassium	1.00	42.89	1046.00	1126.78	0.00	***
creat_ln	1.00	26.00	1045.00	1100.78	0.00	***
arterial_ph	1.00	7.01	1044.00	1093.78	0.008	**
Brandenburg_category_Ethanol	1.00	2.76	1043.00	1091.02	0.10	.
Brandenburg_category_Alcohol	1.00	4.40	1042.00	1086.61	0.04	*
Brandenburg_category_Analgesics	1.00	0.24	1041.00	1086.37	0.62	
Brandenburg_category_Antidiabetics	1.00	0.19	1040.00	1086.19	0.66	
Brandenburg_category_Chemicals	1.00	4.83	1039.00	1081.36	0.03	*
Brandenburg_category_Antidepressants	1.00	2.35	1038.00	1079.00	0.13	.
Brandenburg_category_Cardiopulmonary	1.00	0.72	1037.00	1078.28	0.40	
Brandenburg_category_Street_drug	1.00	0.00	1036.00	1078.28	0.99	
Brandenburg_category_Toxin_not_otherwise_specified	1.00	4.04	1035.00	1074.24	0.04	*
Brandenburg_category_Opioids	1.00	0.44	1034.00	1073.80	0.51	
Brandenburg_category_NA	1.00	1.43	1033.00	1072.37	0.23	
Brandenburg_category_2_or_more	1.00	0.59	1032.00	1071.78	0.44	

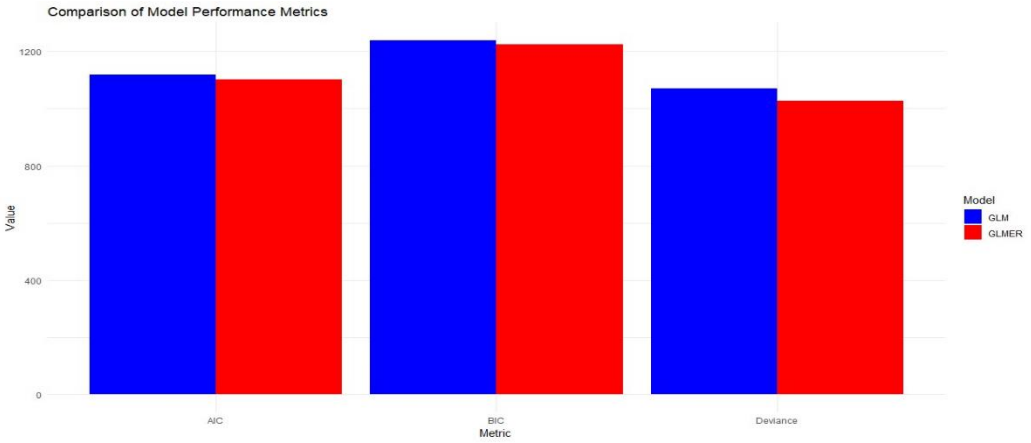


Figure 11: *The Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and he Deviance of the comparison of GLM and GLMER model. GLMER model has better fit of the data and explain better the poor outcome (p-value <0.05 for all three measurements).*

Discussion

In this study, we aimed to investigate the risk of poor outcomes in ICU patients with acute poisoning, with a specific focus on understanding the potential variations between countries. Additionally, we sought to assess whether the implementation of Generalized Linear Mixed Models (GLMM) could improve the predictive accuracy of poor outcomes in comparison to the previously used Generalized Linear Models (GLMs).

Firstly, it is evident that patient outcomes do indeed vary by country, even after adjusting for an array of demographic, clinical, and exposure-related factors. Our study revealed that countries act as random effects, exerting a significant influence on patient outcomes. These variations may be attributed to differences in healthcare infrastructure, treatment protocols, access to medical resources, or even cultural and socioeconomic factors. Furthermore, we explored the utility of Generalized Linear Mixed Models (GLMM) in enhancing our understanding of these variations. Our findings clearly support the use of GLMM as an advance analytical tool compared to traditional Generalized Linear Models (GLMs). Incorporating random effects at the country level in GLMM provides a more refined perspective on the factors contributing to patient outcomes.

When comparing the initial model, which solely considers patients' characteristics, with the second model that incorporates laboratory results, a noteworthy enhancement in poor outcome discrimination becomes evident. The Likelihood Ratio Test underlines the significance of this improvement (p-value<0.001). Moving on to Model 3, which introduces exposure types alongside patients' characteristics, we observe that while exposures contribute significantly to improving poor outcome discrimination, their impact appears to be less influential than laboratory results. This trend is highlighted by the higher AIC values of model 3 compare to model 2 across the five imputed datasets presented (**Table 7**).

In contrast, when we introduced exposures in alongside physiological parameters, we didn't observe significant improvements in outcome interpretation. However, upon transitioning to Model 4, which encompasses all available parameters—physiological parameters, laboratory results, and exposures—we witnessed a noteworthy enhancement in the discrimination of poor outcomes. This improvement consistently received strong support from high levels of significance in the Likelihood Ratio Test (LRT) and when comparing the AIC values across the five imputed datasets. While there was a slight AIC penalty due to the model's increased complexity compared to Model 2, which solely considers lab results, Model 4 significantly (p-value=0.01) outperformed in explaining poor outcomes.

The results from the likelihood-ratio tests robustly point towards a higher level of complexity in the data than what can be effectively captured by a model containing solely physiological and laboratory parameters. This implies that these parameters alone are inadequate for comprehensively delineating the distinctions between poor and good outcomes. In the clinical context, this signifies that when evaluating a patient's condition, factors pertaining to exposure types should be taken into account. Furthermore, it emphasizes that physiological and laboratory parameters do not entirely account for the impact of exposure types. This conclusion gains further credence from their inclusion in the final model. This aligns with the findings from a prior study by Brandenburg et al. (2014), which similarly reported a notable impact of the type of intoxication on both in-hospital mortality and mortality within one month after ICU admission.

The intricate interplay and the associated AIC penalty in the model comparisons underscore the necessity for refining the approach to exposure categories in our study. To provide a more comprehensive explanation of the outcomes and to establish a more resilient relationship with other

parameters, we must reevaluate the categorization of exposures. Our study's classification of exposure types faced challenges due to inconsistencies and variations in categorization methods. Exposure types can be grouped differently, causing differences in toxicity and overlaps. Testing each substance individually is impractical due to numerous exposures. To address this, we added categories like chemicals, ethanol, and opiates to improve accuracy. Future research can refine exposure categorization approaches.

A significant concern is the high prevalence of patients exposed to multiple substances in our dataset. Future research might combine certain exposures into distinct categories or exclude low ethanol doses from the "two or more types" category. This is prompted by patients consuming substances with low ethanol doses, which may not be relevant for our research focused on significant outcomes and ICU admissions. Moreover, enhancing our models' precision in discriminating patient outcomes could be achieved by considering the inclusion of additional exposure-related data, specifically exposure dosages. However, it's imperative to approach this cautiously due to the potential biases stemming from missing dosage information during patient inclusion in the study.

Another avenue for improving the models' discrimination ability between patient outcomes involves the transformation of physiological and laboratory parameters. Several laboratory parameters, such as creatinine levels and arterial pH, exhibit right-skewed distributions (as illustrated in ***Figure 4***), which can pose challenges in establishing a linear relationship with the outcome variable. Additionally, the presence of outliers in categorical variables significantly associated with poor outcomes can lead to misinterpretations within the model (***Figure 3***). These issues, in turn, disrupt the assumption of homoscedasticity, yielding potentially erroneous results (**Figure 5**).

To enhance the models, applying techniques like splines and log transformations to these variables can facilitate the establishment of more linear relationships with poor outcomes, leading to stronger associations between the variables and adverse outcomes (Gurka, Edwards, Muller, & Kupper, 2006). However, this approach comes with a trade-off, as it can lead to decreased interpretability when working with data that is no longer on its original scale (Grilli & Rampichini, 2015; Houle, Pélabon, Wagner, & Hansen, 2011; Jacqmin-Gadda, Sibillot, Proust, & Molina, 2007). This study did not manipulate the value variables, as its primary purpose was to investigate the relationship between country and outcomes, as well as to assess the suitability of GLMM models for this analysis.

On the contrary, the random effects of the countries consistently displayed a significant influence on the differentiation of outcomes in all four models, even when other parameters were not considered. This is evident from the caterpillar plots (**Appendix: 4.Model Selection**) for each model. The random intercept for each country significantly deviates from 0, indicating a substantial random effect on the model's outcome. It's important to note that the distribution of patients across countries is not uniform, with the Netherlands having the highest number of included patients, which could potentially introduce bias into the estimations.

In crafting the final model, we meticulously selected variables that demonstrated significant explanatory power regarding the outcome. The criteria employed for variable selection, as derived from the full model (Model 4: Table 13), involved considering a p-value threshold of 0.2 for patient characteristics, clinical assessment criteria, and laboratory results. These p-value criteria were determined following established guidelines (Steyerberg, 2019). Based on these p-values, we made the decision to exclude psychiatric comorbidity (p-value 0.399) and bad_sodium (p-value 0.439) from the "current final model." This decision aligns with previous research on intoxication, which has often integrated a range of laboratory parameters and patient characteristics, including age, gender,

somatic comorbidities, temperature, heart rates, GCS, systolic blood pressure, and more (Han, et al., 2021; O'Brien, Murphy, Conrick-Martin, & Marsh, 2009; Liisanantti, Ohtonen, Kiviniemi, Laurila, & Ala-Kokko, 2011; Lionte, Sorodoc, Jaba, & Botezat, 2017). However, it's worth noting that all exposure categories were retained in the model due to their crucial relevance to the analysis.

Nonetheless, it's important to note that GLMM models are particularly well-suited for managing non-normally distributed variables, making them a fitting choice for research of this nature. In a robustness study involving mixed models, it was observed that violations of distributional assumptions, whether pertaining to random effect variances or residual variances, had surprisingly minimal impact on the estimates of interest. However, it's worth highlighting that these violations did affect the precision of the estimates. Specifically, when dealing with severely skewed distributions and heteroskedastic residuals, the precision of the estimates decreased (Schielzeth, et al., 2020).

In our study, although GLMM models are well-equipped to handle missing data, we adopted a more robust approach by incorporating Multiple Imputation (MI). Prior to MI, a meticulous data selection process was executed, specifically targeting patients with at least half of their data complete. Patients with over 50% missing data were deliberately excluded from the study to uphold the integrity of the multiple imputation process and to avert potential bias. This practice aligns with studies that have indicated that a substantial amount of missing data can introduce bias into MI estimations (Marino, Lucas, Latour, & Heintzman, 2021; Lee & Carlin, 2012; Haji-Maghsoudi, Haghdoost , Rastegari, & Baneshi, 2013; Wulff & Jeppesen, 2017).

In contrast, Madley-Dowda, Hughesa, Tilling, & Heron, (2019) study emphasizes the paramount importance of prioritizing imputation quality over merely focusing on the proportion of missing data. Their research underscores that robust imputation methods effectively mitigate bias and enhance analysis efficiency. Recent findings also suggest that increasing the number of imputed datasets from the conventional 5 to 10 to as many as 40 significantly bolsters statistical power, particularly when dealing with varying degrees of missing data (Graham, Olchowski, & Gilreath , 2007; AZUR, STUART, FRANGAKIS, & LEAF, 2011). To ensure reliability, the imputation model must encompass all relationships explored in subsequent analyses, including both dependent variables and potential interactions. Additionally, the inclusion of "auxiliary" variables in imputation, even if unused in the analysis, enhances imputation quality (Schafer, 2003; Collins, Schafer, & Kam, 2001). This comprehensive approach substantially elevates imputed dataset accuracy and dependability (Austin, White, Lee, & van Buuren, 2021).

Lastly, the comparative analysis of GLM and GLMM models highlights the superior resilience of GLMM models in handling complex data like this. The Likelihood Ratio Test, considering AIC and BIC, along with the examination of Deviance between the two models, clearly demonstrates that the inclusion of random effects significantly enhances the model's capacity to explain poor outcomes (p-value <0.05). This improvement is particularly notable when considering the unbalance in the distribution of the data among exposure types and non-transformed data. Likewise, mixed models have been employed in two distinct studies focusing on population heterogeneity concerning childhood stunting in Ethiopia and Indonesia. These studies utilized mixed models as a means to explore variations within populations characterized by diverse socioeconomic backgrounds (Mulyaningsih, et al., 2021; Takele, Zewotir, & Ndanguza, 2019).

Similarly, the incorporation of various countries into our study enriches our comprehension of patient outcomes in the aftermath of intoxication incidents. This strategy acknowledges the varying socioeconomic factors and healthcare disparities existing among different nations, thereby

contributing to a more comprehensive analysis of the factors shaping patient results. In a similar vein, to gain a deeper insight into healthcare disparities, we can consider including each hospital as a random effect in our models, while treating countries as fixed effects. This approach signifies our intent to estimate random effects for both "country" and "hospital," recognizing that hospitals within the same country may share certain common characteristics or exhibit similar behavior due to their geographical location. It accommodates the possibility that different hospitals within the same country may exert distinct effects on the response variable.

Conclusion

In conclusion, our study not only validates the presence of significant variations in patient outcomes across different countries but also highlights the pivotal role of advanced statistical methodologies such as Generalized Linear Mixed Models (GLMM) in unravelling the intricate nuances of these disparities. Our findings underscore the importance of addressing critical considerations brought to the forefront in this analysis. These include the refinement of exposure categorization, the enhancement of explanatory variable characterization, and the thoughtful handling of missing data. By embracing these refinements and applying them to our modelling approach, we can aspire to not only improve the accuracy of our predictions but also to foster a deeper understanding of the multifaceted factors influencing patient outcomes.

References

Abayomi, K., Gelman, A., & Levy, M. (2007). Diagnostics for multivariate imputations. *Journal of the Royal Statistical Society Series C: Applied Statistics, 57*(3), 273-291. Retrieved from https://doi.org/10.1111/j.1467-9876.2007.00613.x

Austin, P. C., White, I. R., Lee, D. S., & van Buuren, S. (2021). Missing Data in Clinical Research: A Tutorial on Multiple Imputation. *Canadian Journal of Cardiology, 37*, 1322-1331. Retrieved from https://doi.org/10.1016/j.cjca.2020.11.010

AZUR, M. J., STUART, E. A., FRANGAKIS, C., & LEAF, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? *International Journal of Methods in Psychiatric Research, 20*(1), 40-49. doi:10.1002/mpr.329

Bevans, R. (2023). Akaike Information Criterion | When & How to Use It. Scribbr. Retrieved from https://www.scribbr.com/statistics/akaike-information-criterion/

Bevans, R. (2023, June 22). *Scribbr.* Retrieved August 21, 2023, from Akaike Information Criterion | When & How to Use It (Example).: https://www.scribbr.com/statistics/akaike-information-criterion/

Brandenburg, R., Brinkman, S., de Keizer, N. F., Kesecioglu, J., Meulenbelt, J., & de Lange, D. W. (2016). The need for ICU admission in intoxicated patients: a prediction model. *Clinical Research, 55*(1), 4-11. doi: 10.1080/15563650.2016.1222616

Brandenburg, R., Brinkman, S., de Keizer, N. F., Meulenbelt, J., & de Lange, D. W. (2014). In-Hospital Mortality and Long-Term Survival of Patients With Acute Intoxication Admitted to the ICU. *Critical Care Medicine, 42*(6), 1471-1479. doi:10.1097/CCM.0000000000000245

Brien, B., Murphy, D., Conrick-Martin, I., & Marsh, B. (2009). The functional outcome and recovery of patients admitted to an intensive care unit following drug overdose: a follow-up study. *Anaesth Intensive Care, 37*(5), 802-806.

Brown, V. A. (2021). An Introduction to Linear Mixed-Effects Modeling in R. *Advances in Methods and Practices in Psychological Science, 4*(1), 1-19.

Cavanaugh, J. E., & Neath, A. A. (2019). The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *WIREs COMPUTATIONAL STATISTICS, 11*(3). Retrieved from https://doi.org/10.1002/wics.1460

Chakrabarti, A., & Ghosh, J. (2011). AIC, BIC and Recent Advances in Model Selection. *Philosophy of Statistics*, 583–605. doi:10.1016/b978-0-444-51862-0.50018-6

Clark, B. J., Binswanger, I. A., & Moss, M. (2014). The Intoxicated ICU Patient: Another Opportunity to Improve Long-Term Outcomes. *Critical Care Medicin, 42*(6), 1563–1564. doi:doi:10.1097/CCM.0000000000000274

Collins, L., Schafer, J., & Kam, C. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychol Methods, 6*(4), 330-351. Retrieved from https://pubmed.ncbi.nlm.nih.gov/11778676/

Cretikos, M., & Parr, M. (2003). Drug Related Admissions to Intensive Care: The Role of Illicit Drugs and Self Poisoning. *Critical Care and Resuscitation, 5*, 253-257.

de Lange, D., Hunault, C. C., & Zwaag, S. M. (2021, September 09). *Protocol title: Intoxicated patients outcome in Europe and other continents.* Retrieved August 17, 2023, from toxicstudy.org.: https://toxicstudy.org/wp-content/uploads/2021/09/20210909.-C1.-INTOXICATE-Europe-protocol.pdf

Fathelrahman, A. I., Ab Rahman, A., & Zain, Z. M. (2008). Self-poisoning by drugs and chemicals: variations in demographics, associated factors and final outcomes. *General Hospital Psychiatry, 30*, 467-470. doi:10.1016/j.genhosppsych.2008.04.004

Glen, S. (2023). *Likelihood-Ratio Tests (Probability and Mathematical Statistics).* Retrieved from StatisticsHowTo.com: Elementary Statistics for the rest of us!: https://www.statisticshowto.com/likelihood-ratio-tests/

Graham, J. W., Olchowski, A. E., & Gilreath , T. D. (2007). How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory. *Prevention Science, 8*, 206-213. doi:10.1007/s11121-007-0070-9

Grilli, L., & Rampichini, C. (2015). Specification of random effects in multilevel models: a review. *Quality & Quantity, 49*, 967–976. Retrieved from https://doi.org/10.1007/s11135-014-0060-5

Gurka, M. J., Edwards, L. J., Muller, K. E., & Kupper, L. L. (2006). Extending the Box–Cox Transformation to the Linear Mixed Model. *Journal of the Royal Statistical Society Series A: Statistics in Society, 169*(2), 273-288. Retrieved from https://doi.org/10.1111/j.1467-985X.2005.00391.x

Haji-Maghsoudi, S., Haghdoost , A.-A., Rastegari, A., & Baneshi, M. R. (2013). Influence of pattern of missing data on performance of imputation methods: an example using national data on drug injection in prisons. *Int J Health Policy Manag, 1*(1), 69-77. doi:10.15171/ijhpm.2013.11.

Han, K. S., Kim, S. J., Lee, E. J., Shin, J. H., Lee, J. S., & Lee, S. W. (2021). Development and validation of new poisoning mortality score system for patients with acute poisoning at the emergency department. *Critical Care, 25*(29). Retrieved from https://doi.org/10.1186/s13054-020-03408-1

Hegde, H., Shimpi, N., Panny, A., Glurich, I., Christie, P., & Acharya, A. (2019). MICE vs PPCA: Missing data imputation in healthcare. *Informatics in Medicine Unlocked, 17*(100275). Retrieved from https://doi.org/10.1016/j.imu.2019.100275

Henderson, A., Wright, M., & Pond, S. M. (1993). Experience wtih 732 acute overdose patients admitted to an intensive care unit over six years. *The Medical Journal of Australia, 158*, 28-30.

Hondebrink, L., Rietjens, S. J., Donker, D. W., Hunault, C. C., van den Hengel-Koot, I., Verputten, P. M., . . . de Lange, D. W. (2021). A quarter of admitted poisoned patients have a mild poisoning and require no treatment: An observational study. *European Journal of Internal Medicine, 66*, 41-47. Retrieved from https://doi.org/10.1016/j.ejim.2019.05.012

Houle, D., Pélabon, C., Wagner, G. P., & Hansen, T. F. (2011). Measurement and Meaning in Biology. *The Quarterly Review of Biology, 86*(1). Retrieved from https://doi.org/10.1086/658408

Jacqmin-Gadda, H., Sibillot, S., Proust, C., & Molina, J.-M. (2007). Robustness of the linear mixed model to misspecified error distribution. *Computational Statistics & Data Analysis, 51*(10), 5142-5154.

Lam, S.-M., Lau, A. C.-W., & Yan, W.-W. (2010). Over 8 years experience on severe acute poisoning requiring intensive care in Hong Kong, China. *Human and Experimental Toxicology, 29*(9), 757-765. doi:10.1177/0960327110361753

Lee, K. J., & Carlin, J. B. (2012). Recovery of information from multiple imputation: a simulation study. *Emerg Themes Epidemiol, 9*(3). doi:10.1186/1742-7622-9-3

Lee, Y., & Nelder, J. (2004). Conditional and marginal models: Another view. Statistical Science . *19*, 219-228. doi:10.1214/088342304000000305

Li, P., Stuart, E. A., & Allison, D. B. (2015). Multiple Imputation: A Flexible Tool for Handling Missing Data. *JAMA Guide to Statistics and Methods, 314*(18), 1966-1967. doi:10.1001/jama.2015.15281

Li, T., Hutfless, S., Scharfstein, D. O., Daniels, M. J., Hogan, J. W., Little, R. J., . . . Dickersina, K. (2014). Standards should be applied in the prevention and handling of missing data for patient-centered outcomes research: a systematic review and expert consensus. *Journal of Clinical Epidemiology, 67*, 15-32. Retrieved from http://dx.doi.org/10.1016/j.jclinepi.2013.08.013

Liisanantti, J. H., Ohtonen, P., Kiviniemi, O., Laurila, J. J., & Ala-Kokko, T. I. (2011). Risk factors for prolonged intensive care unit stay and hospital mortality in acute drug-poisoned patients: An evaluation of the physiologic and laboratory parameter on admission. *Jouranl of Critical Care, 26*, 160-165. doi:10.1016/j.jcrc.2010.08.009

Lionte, C., Sorodoc, V., Jaba, E., & Botezat, A. (2017). Development and validation of a risk-prediction nomogram for in-hospital mortality in adults poisoned with drugs and nonpharmaceutical agents. *Medicine, 96*(12). Retrieved from http://dx.doi.org/10.1097/MD.0000000000006404

LITTLE, R. J., & RUBIN, D. B. (1989). The Analysis of Social Science Data with Missing Values. *Sociological Methods & Research, 18*(2-3), 292-326. Retrieved from https://doi.org/10.1177/00491241890180020

Madley-Dowda, P., Hughesa, R., Tilling, K., & Heron, J. (2019). The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of Clinical Epidemiology, 110*, 63-73. Retrieved from https://doi.org/10.1016/j.jclinepi.2019.02.016

Manly, C. A., & Wells , R. S. (2014). Reporting the Use of Multiple Imputation for Missing Data in Higher Education Research. *56*, 497-409. Retrieved from https://doi.org/10.1007/s11162-014-9344-9

Marino, M., Lucas, J., Latour, E., & Heintzman, J. D. (2021). Missing data in primary care research: importance, implications and approaches. *Family Practice, 38*(2), 199-202. Retrieved from https://doi.org/10.1093/fampra/cmaa134

Masconi, K. L., Matsha, T. E., Erasmus, R. T., & Kengne , A. P. (2015). Effects of Different Missing Data Imputation Techniques on the Performance of Undiagnosed Diabetes Risk Prediction Models in a Mixed-Ancestry Population of South Africa. *PLOS ONE, 10*(9). Retrieved from https://doi.org/10.1371/journal.pone.0139210

Mulyaningsih, T., Mohanty, I., Widyaningsih, V., Gebremedhin, T. A., Miranti, R., & Wiyono, V. H. (2021). Beyond personal factors: Multilevel determinants of childhood stunting in Indonesia. *PLOS ONE, 16*(11). Retrieved from https://doi.org/10.1371/journal.

O'Brien, B., Murphy, D., Conrick-Martin, I., & Marsh, B. (2009). The functional outcome and recovery of patients admitted to an intensive care unit following drug overdose: a follow-up study. *Anaesth Intensive Care, 37*, 802-806.

Raghunathan, T., Lepkowski, J., Van Hoewyk, J., & Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology, 27*, 85-95.

Rezar, R., Jung, C., Mamandipoor, B., Seelmaier, C., Felder, T. K., Lichtenauer, M., . . . Osmani, V. (2022). Management of intoxicated patients – a descriptive outcome analysis of 4,267 ICU patients. *BMC Emergency Medicine, 22*(38), 1-11. Retrieved from https://doi.org/10.1186/s12873-022-00602-y

RUBIN, D. B. (1976). Inference and missing data. *Biometrika, 63*(3), 581-592.

Schafer, J. L. (2003). Multiple Imputation in Multivariate Problems When the Imputation and Analysis Models Differ. *Statistica Neerlandica, 57*(1), 19-35. Retrieved from https://doi.org/10.1111/1467-9574.00218

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychol Methods, 7*(2), 147-177.

Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., . . . Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution, 11*, 1141–1152. doi:10.1111/2041-210X.13434

Sterne, J. A., Carlin, J. B., Royston, P., Kenward, M. G., White, I. R., Spratt, M., . . . Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ , 338*. Retrieved from https://doi.org/10.1136/bmj.b2393

Steyerberg, E. W. (2019). 11.2 Stepwise Selection. In E. W. Steyerberg, *Clinical Prediction Models: A Practical Approach to Development,* (pp. 210-12). Leiden: Springer.

Takele, K., Zewotir, T., & Ndanguza, D. (2019). Understanding correlates of child stunting in Ethiopia using generalized linear mixed models. *BMC Puplic Health, 19*(626). Retrieved from https://doi.org/10.1186/s12889-019-6984-x

Takele, K., Zewotir, T., & Ndanguza, D. (2019). Understanding correlates of child stunting in Ethiopia using generalized linear mixed models. *BMC Public Health, 19*(626). Retrieved from https://doi.org/10.1186/s12889-019-6984-x

(n.d.). *University of Bristol.* Retrieved from Centre for Multilevel Modelling: https://www.bristol.ac.uk/cmm/learning/videos/random-slopes.html

Winter, B., & Wieling, M. (2016). How to analyze linguistic change using mixed models, Growth Curve Analysis and Generalize Additive Modeling. *Journal of Language Evolution, 1*(1), 7-18. doi:10.1093/jole/lzv003

Wulff, J. N., & Jeppesen, L. E. (2017). Multiple imputation by chained equations in praxis. *Electronic Journal of Business Research Methods, 15*(1), 41-56.

Zhang, Z. (2015). Multiple imputation with multivariate imputation by chained equation (MICE) package. *Big-data Clinical Trial Column, 4*(2). doi: 10.3978/j.issn.2305-5839.2015.12.63

Appendix

1. Distribution of variables

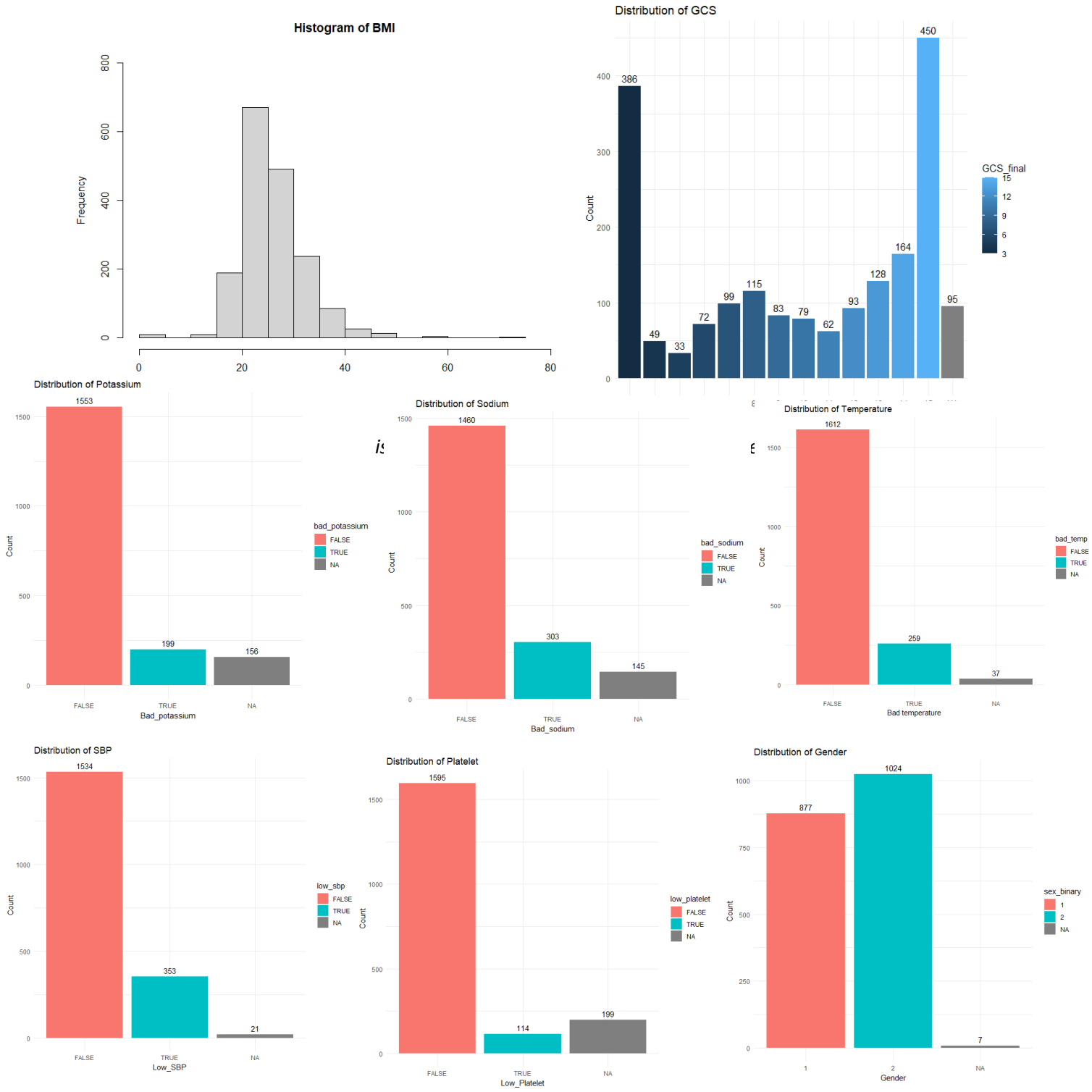


Figure 13: *Distribution of binary variables. Bad potassium, Bad sodium, Low Temperature, Low SBP, Low Platelet and Gender.*

2. Distribution of Exposure Variables in different countries

Table 10: *The distribution of patients exposures across five participated countries in the study. The results are presented as percentages of patients in each country for each exposure occurrence.*

Variables	All Patients (N=1907)	Belgium (N=114)	Turkey (N=119)	Spain (N=214)	Sweden (N=200)	Netherlands (N= 873)
Alcohol (TRUE)	2.53 (N=47)	0.88 (N=1)	18.49 (N=22)	1.87 (N=4)	0.50 (N=1)	0.34 (N=3)
Ethanol (TRUE)	3.35 (N=64)	2.63 (N=3)	3.36 (N=4)	3.74 (N=8	4.50 (N=9)	2.86 (N=25)
Analgesics (TRUE)	2.30 (N=44)	10.53 (N=12)	2.52 (N=3)	3.74 (N=8)	1 (N=2)	0.92 (N=8)
Antidiabetics (TRUE)	2.57 (N=49)	0.88 (N=1)	0	3.27 (N=7)	2.5 (N=5)	3.55 (N=31)
Antipsychotics (TRUE)	10.59 (N=202)	3.51 (N=4)	9.24 (N=11	19.16 (N=41)	5.5 (N=11)	10.54 (N=92)
Antidepressants (TRUE)	2.93 (N=56)	3.51 (N=4)	5.04 (N=6)	1.40 (N=3)	2.50 (N=5)	3.20 (N=28)
Cardiopulmonary (TRUE)	2.73 (N=52)	2.63 (N=3)	8.40 (N=10)	0.47 (N=1)	2 (N=4)	1.83 (N=16)
Chemicals (TRUE)	3.30 (N=63)	6.14 (N=7)	5.88 (N=7)	7 (N=15)	0.50 (N=1)	0.8 (N=7)
Opioids (TRUE)	3.46 (N=66)	4.39 (N=5)	0.84 (N=1)	1.87 (N=4)	4 (N=8)	3.78 (N=33)
Street_drug (TRUE)	3.98 (N=76)	0	0	5.14 (N=11)	1 (N=2)	5.04 (N=44)
2_or_more (TRUE)	51.31 (N=979)	56.14 (N=64)	21 (N=25)	45.79 (N=98)	66 (N=132)	60.72 (N=530)
Toxin_not_otherwise_specified (TRUE)	7.34 (N=140)	7.89 (N=9)	11.76 (N=14)	4.67 (N=10)	9 (N=18)	4.24 (N=37)
NA (TRUE)	3.67 (N=70)	0.87 (N=1)	13.45 (N=16)	1.87 (N=4)	1 (N=2)	2.18 (N=19)

3. Multiple Imputations

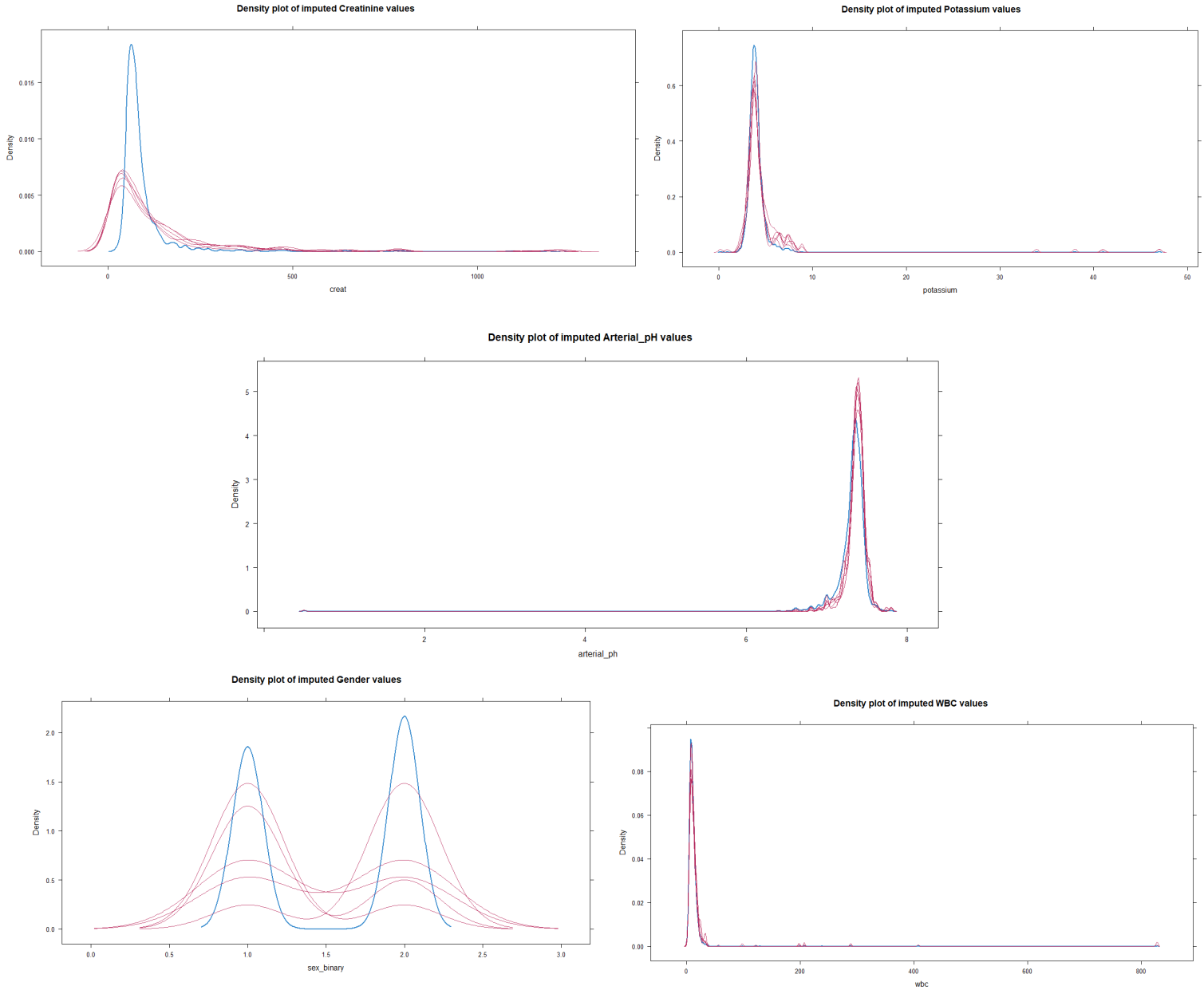
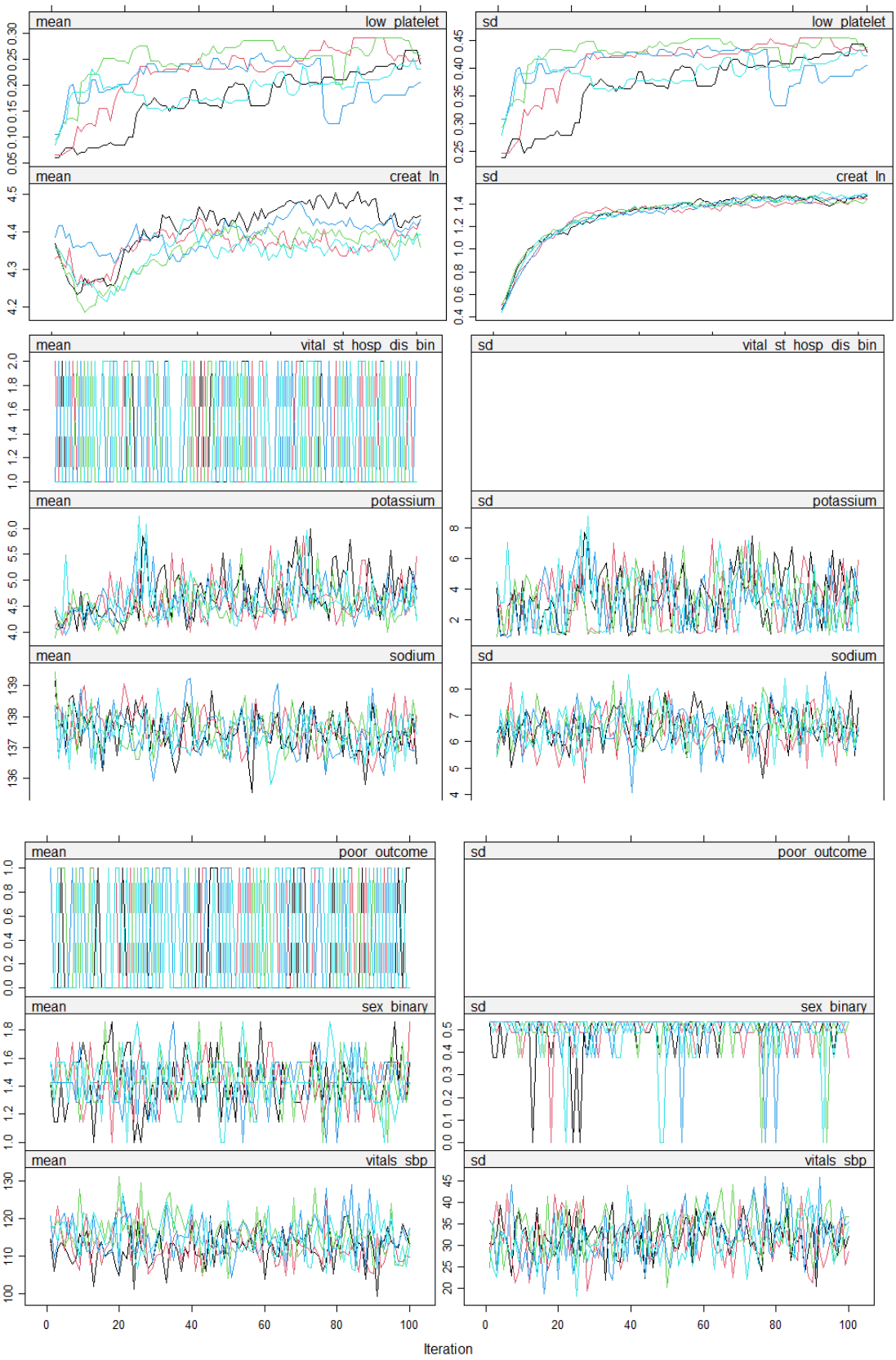


Figure 14:*The density plots of comparison of imputed data (depicted in red) with real data (shown in blue) pertaining to patients' characteristics (sex_binary) and laboratory results (creatinine, potassium arterial pH and white blood cells counts). The graphs demonstrate a consistent alignment between the imputed and real data trends. While most imputed data points closely track the curve of the actual data, certain imputed values deviate slightly below the real data curve such as creatinine.*



*F**igure 15**:Mean and standard deviation of the synthetic values plotted against iteration number for the imputed data.*

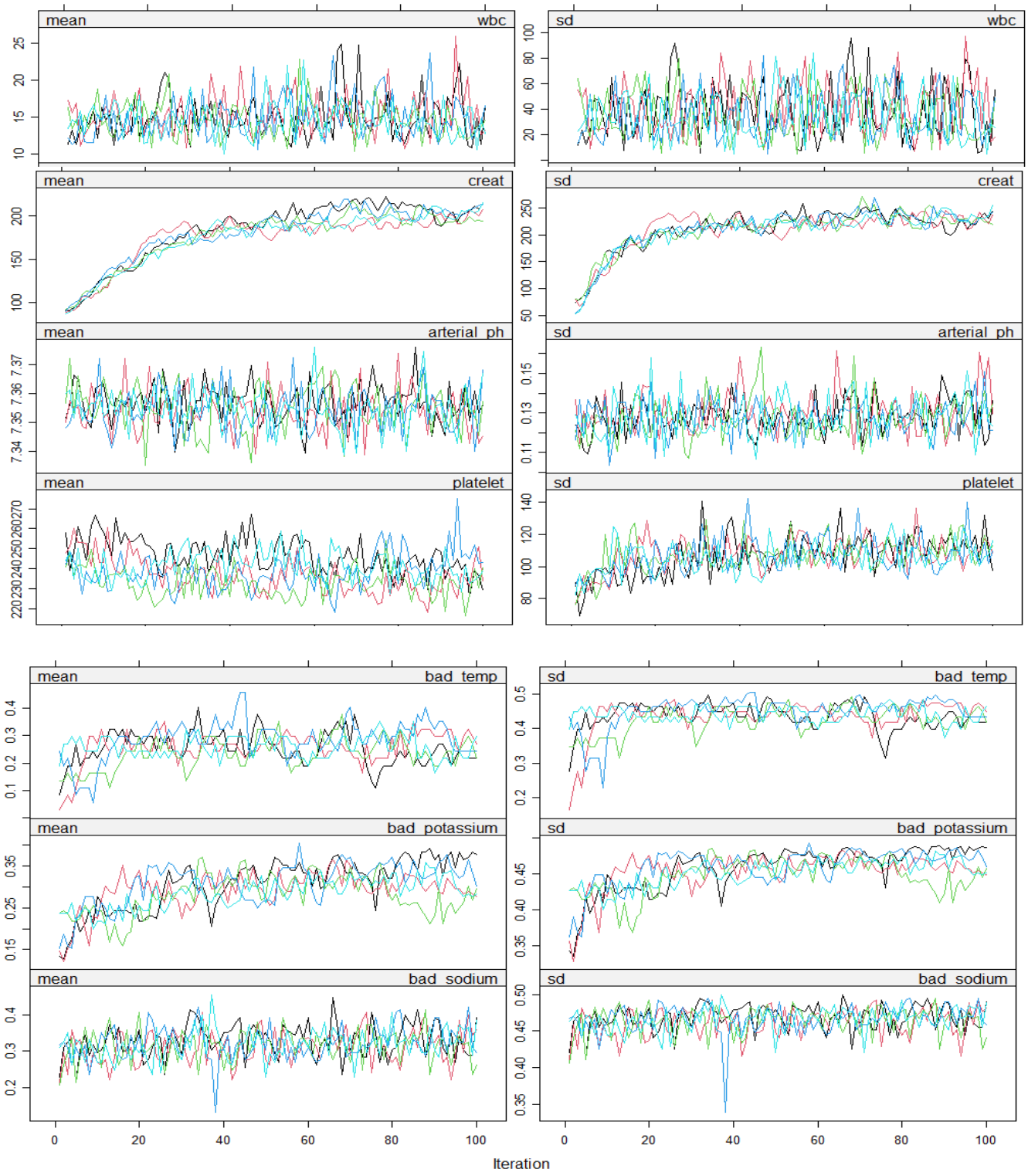


Figure 16:*Mean and standard deviation of the synthetic values plotted against iteration number for the imputed data.*

4. Model Selection

Table 11: *Results of model 1 with Block 1 made with the imputed datasets. The Block includes the age and gender of the patients, the physiological parameters and the somatic comorbidities of the patients.*

	TERM	ESTIMATE	STD.ERROR	STATISTIC	DF	P.VALUE	2.50%	*97.50%*
1	(Intercept)	0.333	0.293	-3.743	1811.037	0.000	0.188	0.593
2	age	1.022	0.004	5.609	1888.253	0.000	1.014	1.030
3	sex_binary2	0.655	0.123	-3.436	1506.088	0.001	0.515	0.834
4	comorbidity_categoryPsychiatric	0.919	0.166	-0.510	1890.464	0.610	0.663	1.273
5	comorbidity_categorySomatic	1.307	0.181	1.479	1895.010	0.139	0.917	1.862
6	GCS	0.901	0.014	-7.616	1089.998	0.000	0.877	0.926
7	low_sbpTRUE	2.269	0.139	5.902	1885.813	0.000	1.728	2.978
8	high_hrTRUE	1.444	0.130	2.823	1755.807	0.005	1.119	1.863

Table 12*: Results of model 2 with Block 1& Block 2 made with the imputed datasets. The Block 1 includes the age and gender of the patients, the physiological parameters and the somatic comorbidities of the patients. The Block 2 includes the laboratory results.*

	TERM	ESTIMATE	STD.ERROR	STATISTIC	DF	P.VALUE	2.50%	*97.50%*
1	(Intercept)	6357.665	4.878	1.795	13.313	0.095	0.173	233835174.331
2	age	1.020	0.004	4.842	1762.089	0.000	1.012	1.028
3	sex_binary2	0.703	0.128	-2.761	1340.461	0.006	0.547	0.903
4	comorbidity_categoryPsychiatric	1.045	0.175	0.253	1505.791	0.800	0.742	1.473
5	comorbidity_categorySomatic	1.276	0.189	1.287	1677.956	0.198	0.880	1.849
6	GCS	0.919	0.015	-5.459	181.720	0.000	0.892	0.948
7	low_sbpTRUE	1.777	0.149	3.866	1832.125	0.000	1.327	2.378
8	high_hrTRUE	1.438	0.134	2.714	1873.960	0.007	1.106	1.870
9	bad_sodiumTRUE	1.175	0.155	1.038	1142.309	0.300	0.867	1.592
10	bad_tempTRUE	1.715	0.167	3.235	1843.669	0.001	1.237	2.378
11	bad_potassiumTRUE	2.310	0.189	4.427	223.305	0.000	1.591	3.353
12	creat_ln	1.238	0.121	1.756	50.761	0.085	0.970	1.580
13	arterial_ph	0.217	0.635	-2.404	14.080	0.031	0.056	0.847

Table 14: *Results of model 3 with Block 1& Block 3 made with the imputed datasets. The Block 1 includes the age and gender of the patients, the physiological parameters and the somatic comorbidities of the patients. The Block 3 includes the different exposure categories.*

	TERM	ESTIMATE	STD.ERROR	STATISTIC	DF	P.VALUE	2.50%	*97.50%*
1	(Intercept)	0.251	0.354	-3.905	1700.907	0.000	0.125	0.502
2	age	1.021	0.004	5.318	1876.161	0.000	1.013	1.029
3	sex_binary2	0.675	0.127	-3.100	1574.930	0.002	0.526	0.865
4	comorbidity_categoryPsychiatric	1.040	0.173	0.228	1870.857	0.819	0.741	1.461
5	comorbidity_categorySomatic	1.366	0.187	1.666	1882.684	0.096	0.946	1.972
6	GCS	0.895	0.015	-7.622	775.286	0.000	0.870	0.921
7	low_sbpTRUE	2.260	0.142	5.744	1867.987	0.000	1.711	2.985
8	high_hrTRUE	1.417	0.132	2.644	1672.988	0.008	1.094	1.836
9	Ethanol	0.577	0.405	-1.358	1819.316	0.175	0.261	1.277
10	Alcohol	4.917	0.421	3.780	1875.703	0.000	2.152	11.236
11	Analgesic	1.353	0.484	0.624	1880.558	0.533	0.524	3.495
12	Anti_diabetic	2.132	0.418	1.811	1882.680	0.070	0.939	4.841
13	Chemicals	1.832	0.355	1.708	1868.146	0.088	0.914	3.673
14	Antidepressant	1.473	0.409	0.947	1882.366	0.344	0.660	3.288
15	Cardiopulmonary	1.058	0.440	0.128	1883.554	0.898	0.446	2.509
16	Street_drug	1.352	0.353	0.855	1844.843	0.393	0.677	2.702
17	TNOS	2.158	0.280	2.746	1881.932	0.006	1.246	3.739
18	Opioids	0.931	0.380	-0.190	1881.734	0.850	0.442	1.959
19	twoormore	1.065	0.211	0.297	1855.471	0.767	0.703	1.612
20	Unknown	1.814	0.363	1.641	1860.340	0.101	0.890	3.699

Table 13*: Results of model 4 with all the variables from the 3 blocks made with the imputed datasets. The Block 1 includes the age and gender of the patients, the physiological parameters and the somatic comorbidities of the patients. The Block 2 includes the laboratory results and the Block 3 includes the different exposure categories*

	TERM	ESTIMATE	STD.ERROR	STATISTIC	DF	P.VALUE	2.50%	*97.50%*
1	(Intercept)	4425.022	5.294	1.586	12.735	0.137	0.047	420141294.087
2	age	1.019	0.004	4.581	1737.749	0.000	1.011	1.027
3	sex_binary2	0.711	0.131	-2.598	1441.620	0.009	0.549	0.920
4	comorbidity_categoryPsychiatric	1.166	0.182	0.844	1540.533	0.399	0.816	1.665
5	comorbidity_categorySomatic	1.373	0.195	1.624	1760.402	0.105	0.936	2.014
6	GCS	0.911	0.016	-5.800	215.298	0.000	0.883	0.941
7	low_sbpTRUE	1.772	0.151	3.780	1806.247	0.000	1.317	2.385
8	high_hrTRUE	1.398	0.135	2.477	1854.489	0.013	1.072	1.824
9	bad_sodiumTRUE	1.131	0.159	0.774	977.697	0.439	0.828	1.545
10	bad_tempTRUE	1.733	0.169	3.247	1793.736	0.001	1.243	2.415
11	bad_potassiumTRUE	2.175	0.197	3.949	177.402	0.000	1.475	3.208
12	creat_ln	1.272	0.125	1.917	52.872	0.061	0.989	1.635
13	arterial_ph	0.218	0.686	-2.218	13.445	0.044	0.050	0.956
14	Ethanol	0.575	0.419	-1.321	1458.854	0.187	0.253	1.308
15	Alcohol	2.979	0.448	2.437	1454.190	0.015	1.237	7.174
16	Analgesic	1.152	0.496	0.285	1747.731	0.776	0.435	3.048
17	Anti_diabetic	1.116	0.461	0.238	1857.736	0.812	0.452	2.753
18	Chemicals	2.027	0.365	1.936	1865.793	0.053	0.991	4.148
19	Antidepressant	1.673	0.415	1.240	1860.811	0.215	0.741	3.779
20	Cardiopulmonary	1.281	0.454	0.546	1667.124	0.585	0.526	3.119
21	Street_drug	1.273	0.369	0.655	1585.040	0.513	0.617	2.626
22	TNOS	2.224	0.288	2.774	1866.744	0.006	1.264	3.913
23	Opioids	0.869	0.398	-0.353	1457.206	0.724	0.398	1.897
24	Unknown	1.710	0.386	1.390	657.146	0.165	0.801	3.647
25	twoormore	1.035	0.220	0.158	1368.772	0.875	0.673	1.592

```
anova(model_3,model_2)
  test statistic df1      df2 dfcom      p.value         riv
1 ~~ 2  2.807484  12 1883.189  1887 0.0008141668 0.004534292
anova(model_4,model_2)
  test statistic df1     df2 dfcom    p.value       riv
1 ~~ 2  2.012415  12 1856.21  1882 0.01997616 0.0168381
anova(model_4,model_3)
  test statistic df1      df2 dfcom      p.value        riv
1 ~~ 2  12.10401   5 264.0818  1882     0.00      0.2674293
```

Figure 18: *Comparisons of model 2 and model 3 with model 4 using anova() function using Likelihood Ratio test and the AIC criteria. Model 4 is explain better the data in the study from Model 2 and Model 3.*

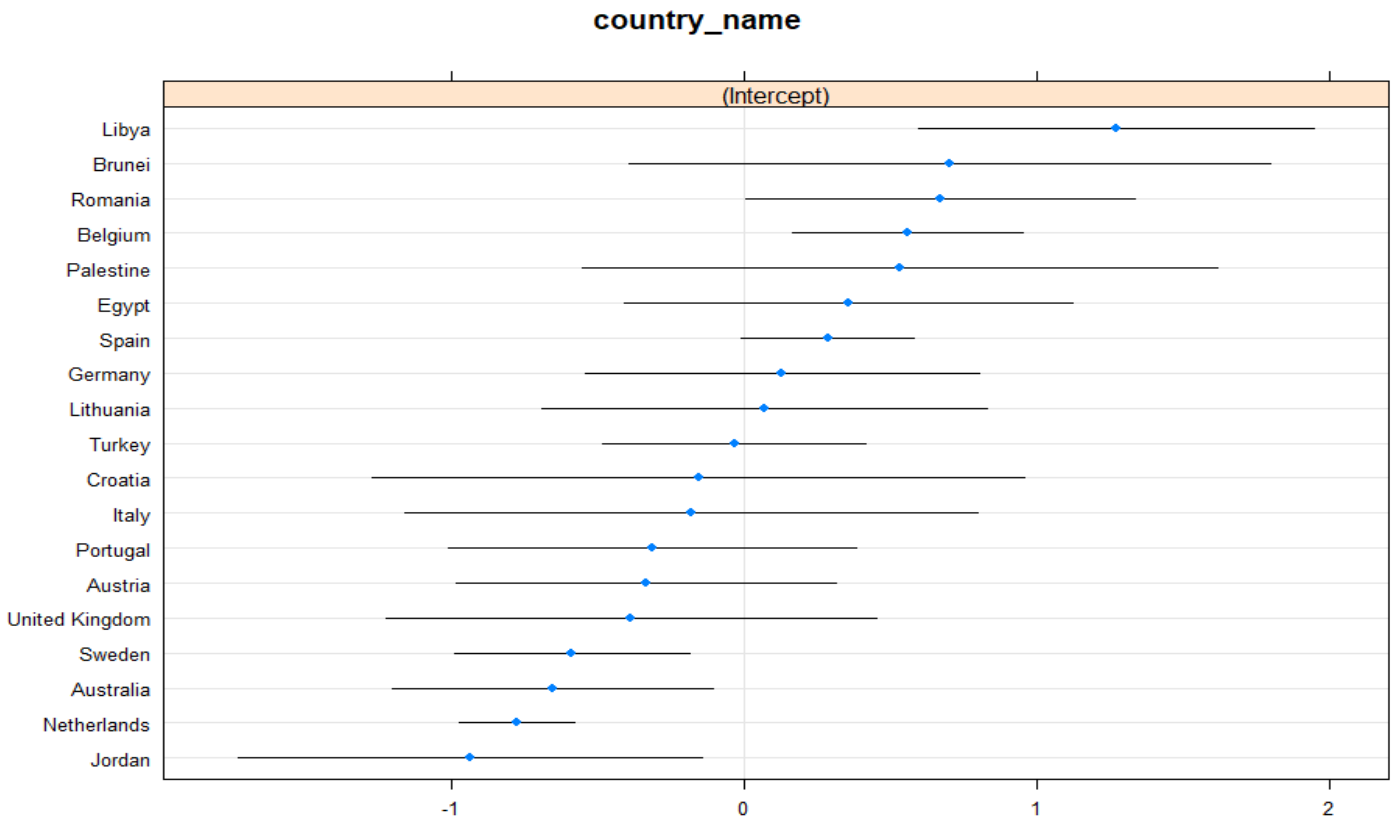


Figure 17*: The caterpillar plot showcasing the random intercepts per country for Model 1. The plot offers insights into the variations in the intercepts across different countries within the context of the model.*

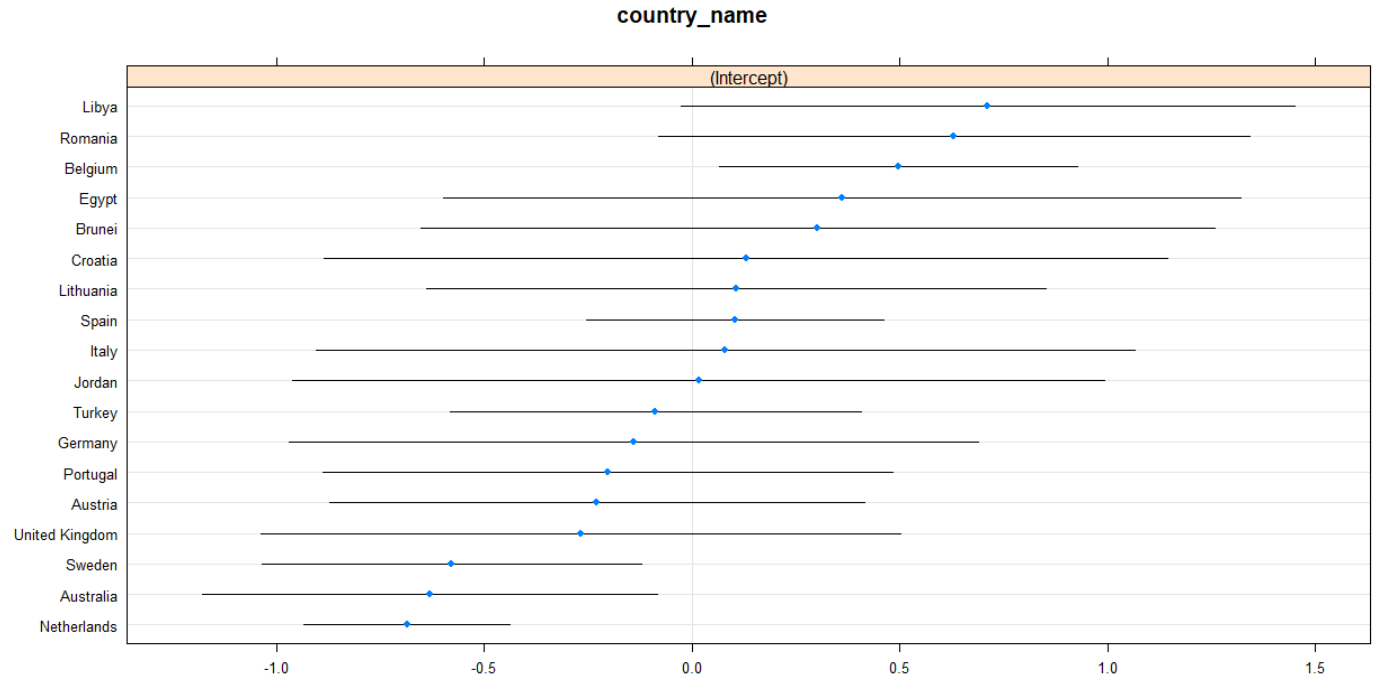


Figure 19: *The caterpillar plot showcasing the random intercepts per country for Model 2. The plot offers insights into the variations in the intercepts across different countries within the context of the model.*

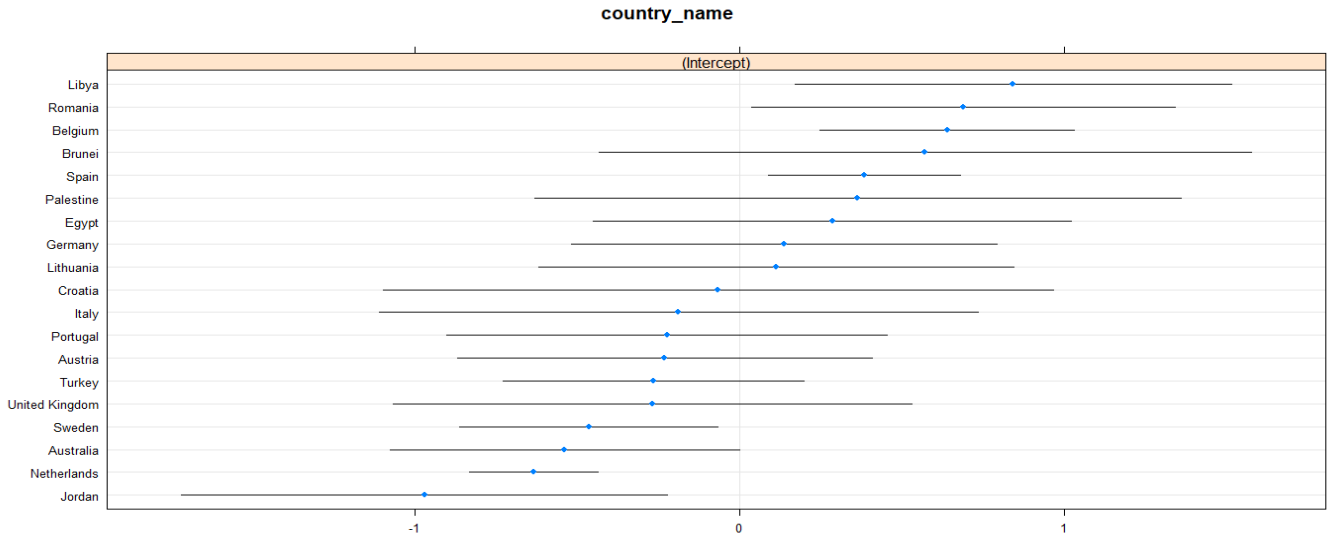


Figure 20:*The caterpillar plot showcasing the random intercepts per country for Model 4. The plot offers insights into the variations in the intercepts across different countries within the context of the model.*

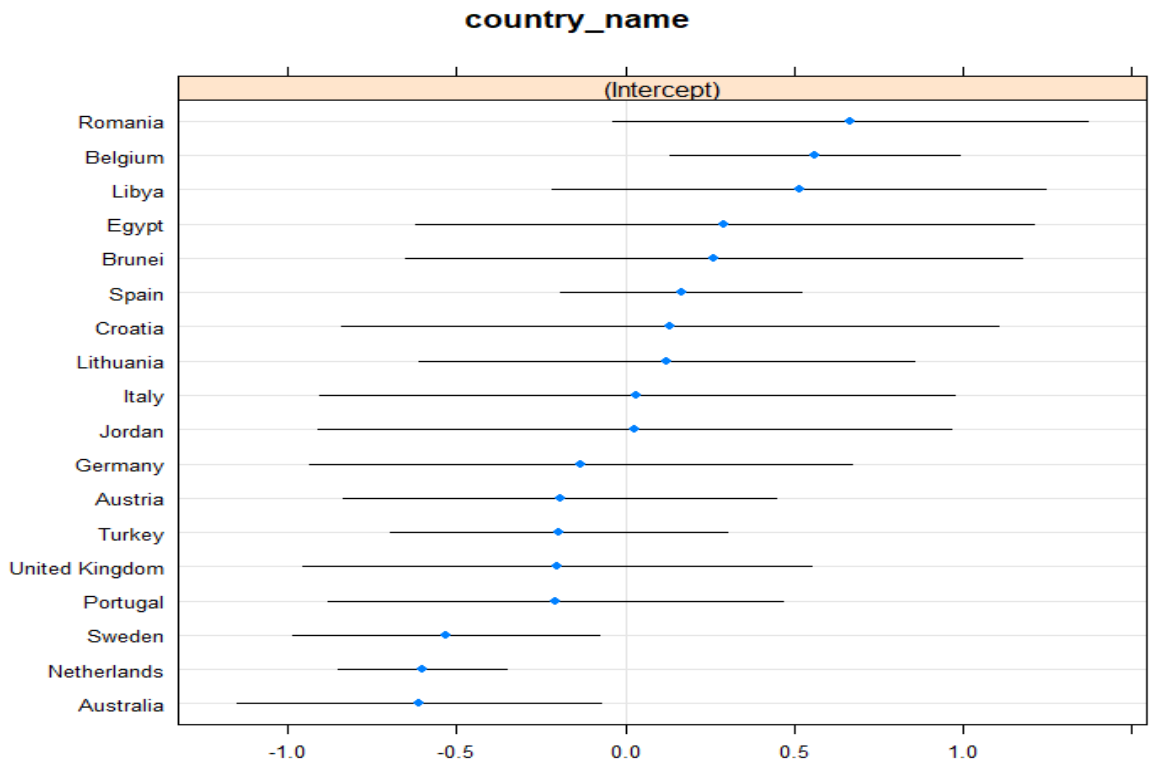
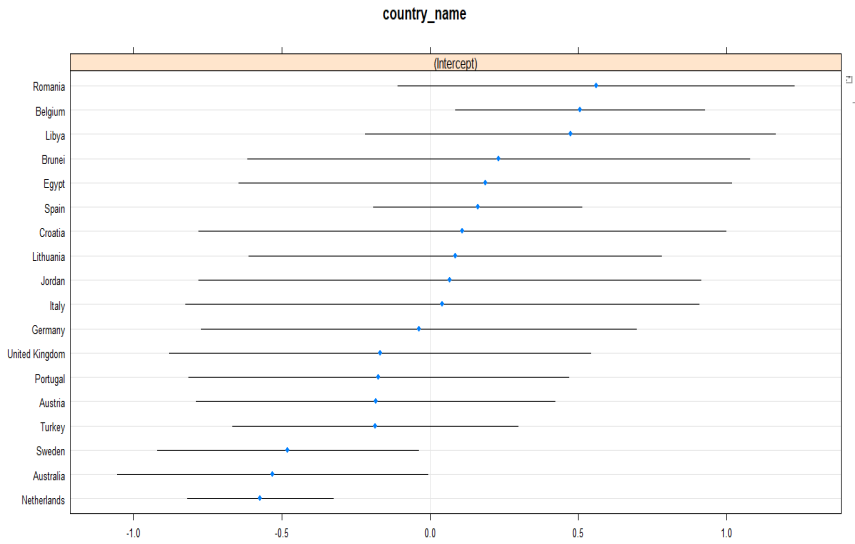


Figure 20: *The caterpillar plot showcasing the random intercepts per country for Model 4. The plot offers insights into the variations in the intercepts across different countries within the context of the model.*



LEVEL	ESTIMATE	STD.ERROR	CONF.LOW	CONF.HIGH
Australia	-0.53103	0.266896	-1.05414	-0.00792
Austria	-0.18293	0.308621	-0.78781	0.42196
Belgium	0.504987	0.21424	0.085085	0.924889
Brunei	0.231119	0.43192	-0.61543	1.077668
Croatia	0.108741	0.453095	-0.77931	0.996792
Egypt	0.18576	0.424154	-0.64557	1.017086
Germany	-0.03787	0.374235	-0.77136	0.695618
Italy	0.040914	0.441271	-0.82396	0.905789
Jordan	0.065996	0.431746	-0.78021	0.912203
Libya	0.472896	0.352403	-0.2178	1.163595
Lithuania	0.08516	0.354952	-0.61053	0.780852
Netherlands	-0.5723	0.125378	-0.81804	-0.32656
Portugal	-0.17343	0.326981	-0.8143	0.467442
Romania	0.559497	0.340627	-0.10812	1.227114
Spain	0.160669	0.179218	-0.19059	0.51193
Sweden	-0.47962	0.224696	-0.92002	-0.03922
Turkey	-0.18466	0.245419	-0.66568	0.296347
United Kingdom	-0.1683	0.362364	-0.87852	0.541918

Figure 19: *The caterpillar plot showcasing the random intercepts per country for "Current" Final Model. The plot offers insights into the variations in the intercepts across different countries within the context of the model. The data represent the real values of the variables without any imputations.*