How asthma medication relates to mortality

Jason Li

Abstract

Background

Asthma remains a significant global health concern, leading to substantial morbidity and mortality despite advancements in treatment. Various pharmacological interventions, including inhaled corticosteroids (ICS), long-acting beta-agonists (LABA), and short-acting beta-agonists (SABA), are widely used to manage the condition. However, the effectiveness of these treatments in reducing asthma-related mortality continues to be an area of extensive research and debate. This study aims to assess the impact of asthma medications on mortality rates.

Method

This study analysed 2,240,628 patients who received asthma medications in New Zealand from 2008 to 2021, using data from the Ministry of Health. Focusing on fluticasone with salmeterol and budesonide with eformoterol, we employed Bayesian methods with the Weibull distribution. Python were used for statistical modelling, estimating parameters through posterior sampling and calculating highest density intervals to ensure robust and reliable findings.

Result

Patients on budesonide with eformoterol (B) had a higher shape parameter (alpha = 0.61) and lower scale parameter (beta = -1.33) compared to those on fluticasone with salmeterol (F), indicating different survival distributions. The Highest Density Intervals (HDIs) for these differences were [0.17, 1.06] for alpha and [-2.07, -0.59] for beta. Additionally, significant differences were observed in patients who transitioned medications, with a mean beta difference of 6.16 (HDI [5.54, 6.78]) for those moving from F to B. These findings highlight the varying impacts of asthma treatments on patient mortality, with transitions between medications also influencing survival outcomes.

Conclusion

Our study reveals significant differences in mortality outcomes based on asthma medication types and transitions, with budesonide with eformoterol showing different survival patterns compared to fluticasone with salmeterol. These findings underscore the need for tailored asthma management strategies to improve patient survival.

Acknowledgements

I would like to express my special thanks to my supervisors, Amy, Jeff, and Andreas for their guidance and support throughout my studies. This project wouldn't have been successful without their help and expertise.

I would like to thank my parents for their unconditional encouragement and support. Their trust in me has motivated me in completing this report.

I would also like to extend my appreciation to my girlfriend, Siki who had taken care of me in the hardest times and supported me in the makings of this project. This journey wouldn't have the smoothness without her.

Table of Contents

1. Introduction

Asthma is a chronic respiratory condition characterised by airway inflammation, bronchoconstriction, and mucus production, leading to recurrent episodes of wheezing, breathlessness, chest tightness, and coughing. The prevalence of asthma has been increasing globally, affecting individuals of all ages and contributing to significant morbidity and mortality. Despite advances in understanding the pathophysiology of asthma and the development of a range of effective therapeutic interventions, asthma-related mortality remains a public health concern (Wang et al., 2023) .

The management of asthma involves a combination of pharmacological and non-pharmacological strategies aimed at controlling symptoms, preventing exacerbations, and improving the quality of life for patients. Pharmacological treatments primarily include the use of short-acting beta-agonists (SABA), inhaled corticosteroids (ICS), long-acting beta-agonists (LABA), leukotriene receptor antagonists (LTRA), and biologics targeting specific pathways involved in asthma pathogenesis. These treatments have been shown to reduce symptoms and exacerbations, yet their impact on asthma-related mortality is an area of ongoing research and debate. Notably, the increased use of inhaled corticosteroids has been proposed as a key factor in the declining trend of asthma mortality rates in New Zealand following the epidemic of asthma deaths linked to high-dose inhaled fenoterol in the 1970s (Faisal & Yunus, 2019).

Asthma mortality is influenced by multiple factors, including the severity of the disease, adherence to treatment, access to healthcare, and socio-economic status. Moreover, the variability in response to treatment among individuals with asthma underscores the need for personalised approaches to management. For example, doubling inhaled corticosteroids for mild asthma exacerbations is a widely adopted strategy in Australia to reduce exacerbations, as supported by national and international guidelines (A. Douglass & Reddel, 2005).

The use of beta-2-agonists has been associated with asthma-related deaths, particularly noted in studies from New Zealand in the late 1970s, which led to a revaluation of therapeutic strategies. New therapies targeting specific pathways involved in asthma pathogenesis, such as omalizumab for patients with severe persistent asthma, have shown promise in improving outcomes and reducing mortality rates (Beasley, 2007).

Understanding how different asthma treatments influence mortality is crucial for optimising therapeutic strategies and improving patient outcomes. This thesis aims to investigate the relationship between asthma treatments and mortality. By examining clinical studies, population health data, and

Introduction

mechanistic insights, this research will explore the effectiveness of various treatment modalities in reducing asthma-related deaths. The findings will contribute to a more nuanced understanding of asthma management and provide evidence-based recommendations for clinical practice.

The scope of this research includes a comprehensive analysis of existing asthma treatments and their relationship with asthma-related mortality. The focus will be on pharmacological treatments such as inhaled corticosteroids (ICS), long-acting beta-agonists (LABA) and short-acting beta agonists (SABA). The research will include:

- Assessment of the impact of asthma medications on mortality rates, considering potential confounding factors such as transition from one medication to another.
- An analysis of clinical studies and population health data from New Zealand.
- Identification of knowledge gaps relating to asthma medication and mortality, including gaps in research methodology.
- Utilisation of New Zealand datasets and novel modelling techniques to address these gaps and provide robust answers to the research questions.

The research will exclude:

- Effects of asthma treatments on non-mortality related outcomes such as quality of life, mental health, and specific disease progressions.
- Case studies or anecdotal evidence that do not provide robust data on the relationship between asthma treatments and mortality.

The primary aim of this research is to investigate the relationship between asthma treatments and mortality. The mathematical modelling approach for this research will employ Bayesian survival analysis using the Weibull distribution to assess the impact of asthma medications on mortality rates. The objectives are:

- Assess the impact of asthma medications on mortality rates, controlling for potential confounders.
- Compare findings from New Zealand with international studies on asthma medication and mortality risks.

This report is organised into five sections:

- Introduction section provides the background information, the scope of the research, aims and objectives, and an outline of the report.
- Literature review section looks at the existing literature on polypharmacy and its effects on older adults, highlighting key findings and identifying gaps that this research aims to address.

- Methods and Material section details the research methodology, including data sources, study population, modelling approach, and analytical techniques used to achieve the research objectives.
- Presentation of findings section shows the results of the study, including statistical analysis and key findings.
- Discussion section discusses the findings and provides recommendations for future research
- Summary section summarises the findings and conclusion

2. Literature Review

Asthma is a significant public health issue, with a considerable impact on morbidity and mortality worldwide. Various pharmacological treatments have been developed to manage asthma symptoms and prevent severe exacerbations that can lead to death. Inhaled corticosteroids (ICS), long-acting beta-agonists (LABA), and short-acting beta-agonists (SABA) are among the most commonly used medications in New Zealand. The effectiveness of these treatments in reducing asthma-related mortality has been the subject of extensive research.

2.1 Impact of asthma treatments on mortality

A number of studies have investigated the impact of different asthma medications on mortality rates. Inhaled corticosteroids (ICS) are widely recognised for their role in reducing inflammation and preventing exacerbations. A study by Faisal and Yunus (2019) found that the increased use of ICS was associated with a significant decline in asthma mortality rates in New Zealand following an epidemic of asthma deaths linked to high-dose inhaled fenoterol in the 1970s (Faisal & Yunus, 2019).

Long-acting beta-agonists (LABA), when used in combination with ICS, have also been shown to improve asthma control and reduce mortality. However, the use of LABA alone has raised concerns due to potential risks of severe exacerbations, as highlighted in studies from the late 1990s and early 2000s. For instance, Wijesinghe et al. (2008) utilised meta-analysis and regression models to examine the risk associated with LABA use, concluding that combination therapy significantly mitigates this risk (Wijesinghe et al., 2008).

Short-acting beta-agonists (SABA) are typically used for immediate relief of asthma symptoms. While effective in managing acute symptoms, reliance on SABA without concurrent anti-inflammatory treatment (such as ICS) has been linked to increased mortality risk. This has led to recommendations for SABA to be used primarily as a rescue medication, with ICS as the mainstay of long-term management.

2.2 Overseas studies on asthma medication and mortality risks

Research from various countries provides insights into the impact of asthma medications on mortality. Studies from the United States, the United Kingdom, and Australia have examined the relationship between different asthma treatments and mortality rates.

Ulrik and Frederiksen (1995) conducted a cohort study in United Kingdom involving 1,075 outpatients, utilising logistic regression to assess the association between high-dose ICS and asthma mortality. The study found that high-dose ICS was associated with a 20% reduction in asthma mortality, emphasising the importance of adherence to maintenance therapy. The study used logistic regression, which is appropriate for assessing associations between a categorical outcome (mortality) and predictor variables (e.g. ICS usage). This method is suitable for this type of cohort data which allows for control of confounding variables. The study was limited to outpatient data, which may not capture the full story for asthma severity. (Ulrik & Frederiksen, 1995)

Tatham and Gellert conducted a meta-analysis in United States using pooled data from multiple randomised controlled trials (RCTs) and applied fixed-effects and random-effects models to evaluate the mortality risk associated with LABAs. The analysis highlighted a 30% increased mortality risk when LABAs were used without ICS, which led to revisions in treatment guidelines. The standardised mortality rate for asthma had not shown consistent improvement despite the advent of new treatments, indicating the need for better management strategies The study used Meta-analysis which is appropriate for synthesising results from multiple studies, increasing statistical power and generalisation ability. The use of both fixed and random effects models helps account for variability between studies. But on the other hand, Meta-analysis relies on the quality of the included studies. Any biases or limitations in the original studies can affect the overall findings. Additionally, heterogeneity in study designs and patient populations can complicate interpretation. There is a need for more standardised reporting in asthma studies to improve the quality of meta-analyses. (Tatham & Gellert, 1985).

Douglass and Bowes conducted a time-series analysis using national asthma mortality data and intervention timelines to evaluate the impact of changes in treatment guidelines. They applied interrupted time-series analysis to assess the effect of increasing ICS use on mortality rates. Their findings supported the increased use of small-particle inhaled corticosteroids as a beneficial strategy in managing uncontrolled asthma Interrupted time-series analysis is appropriate for evaluating the impact of policy changes or interventions over time. It allows researchers to assess trends before and after an intervention while controlling for underlying trends. Time-series analyses can be affected by confounding factors that change over time and are not accounted for in the model.

Additionally, while this method can suggest associations, it does not prove causation. Further studies are needed to confirm these findings and explore underlying mechanisms. (Douglass & Bowes, 1990).

2.3 New Zealand studies on asthma treatment and mortality

Studies have shown that the Māori population, in particular, experienced higher mortality rates compared to other ethnic groups, underscoring the need for targeted interventions and improved access to healthcare services.

Ellison-Loschmann et al. conducted a comprehensive analysis of asthma mortality rates among Māori and non-Māori from the 1960s to the early 2000s. They found that asthma mortality rates were disproportionately higher among Māori, with peak rates in 1979 being twice as high for Māori (7.4 per 100,000) compared to non-Māori (3.7 per 100,000). The study utilised longitudinal data and time-series analysis to examine trends and seasonal patterns over several decades (Ellison-Loschmann et al., 2008). The use of longitudinal data allowed for the observation of changes over time, which is appropriate for analysing trends. However, the study could have benefited from more advanced statistical methods to control for potential confounders.

Crengle et al. explored the rural-urban disparities in asthma mortality rates among Māori in New Zealand. Their study showed that rural Māori experience greater asthma mortality rates compared to their urban counterparts. This research employed population-level data and logistic regression models to assess the impact of rurality on asthma outcomes, highlighting significant health inequities and disparities in access to quality healthcare services (Crengle et al., 2022). Logistic regression is suitable for examining the relationship between a binary outcome (mortality) and multiple predictor variables (rural vs. urban residence). However, the study might have been improved by incorporating spatial analysis techniques to account for geographical variations more precisely.

Phillips et al. examined broader mortality trends among Māori and non-Māori in New Zealand, focusing on the life expectancy gaps between these groups from the early 1980s to the mid-2010s. The study found that the life expectancy gaps increased significantly during the 1980s and 1990s but have since decreased. In the most recent period analysed (2012-2014), the life expectancy gap between Māori and non-Māori was 7.3 years for males and 6.8 years for females. This study used decomposition analysis and Cox proportional hazards models to analyse cause-specific mortality, including asthma-related deaths(Phillips et al., 2017). Decomposition analysis helps to understand the contributions of different causes of death to overall mortality differences, while Cox proportional

hazards models are useful for assessing the impact of various covariates on survival time. Despite these strengths, the study could have been enhanced by incorporating more granular data to examine intra-group variations and using more recent data to capture ongoing trends.

2.4 Knowledge gaps in asthma medication and mortality

Despite the extensive research, several knowledge gaps remain regarding asthma medication and mortality. These gaps include:

- The long-term effects of transitioning from one medication to another on mortality outcomes.
- The need for more comprehensive studies that control for potential confounders such as environmental factors.

Additionally, there are methodological gaps in existing research, including variations in study design, data collection methods, and analytical approaches. Addressing these gaps is crucial for developing a more accurate understanding of the relationship between asthma treatments and mortality.

Our study aims to fill these gaps by employing a comprehensive and innovative approach to analysing the relationship between asthma medications and mortality. Here are the key aspects that make our methods unique and innovative:

- **Bayesian Survival Analysis with Weibull Distribution**: We utilise Bayesian survival analysis, leveraging the Weibull distribution to model time-to-event data. This approach allows us to incorporate prior information and provides a full probabilistic framework for parameter estimation, offering a more nuanced understanding of the impact of asthma medications on mortality rates.
- **Integration of Comprehensive Datasets**: By utilising extensive datasets from the New Zealand Ministry of Health (MoH), our study benefits from a rich source of longitudinal data, which enhances the robustness and reliability of our findings. This comprehensive dataset allows us to control for a wide range of potential confounders, including environmental factors and socio-economic variables.
- **Advanced Modelling Techniques**: Our approach includes advanced modelling techniques such as Bayesian hierarchical models and parametric Bayesian survival models. These methods are more flexible and robust in handling complex relationships and non-proportional hazards, addressing the limitations of traditional methods like the Cox proportional hazards model.

- **Methodological Innovation**: Our study addresses methodological gaps in existing research by standardising data collection methods and employing rigorous analytical approaches. The use of Bayesian methods allows for the integration of prior knowledge and the continuous updating of estimates as new data becomes available, enhancing the validity and reliability of our findings.

2.5 Utilisation of New Zealand datasets and Bayesian Analysis

This research seeks to address these knowledge gaps by utilising comprehensive datasets from New Zealand Ministry of Health (MoH). By applying novel modelling techniques, including Bayesian analysis, this study aims to provide robust answers to key research questions.

Cox proportional hazards model has a long history of successful application in medical research as seen in most of the literatures that were previously referenced in. For example, de la Cruz et al. (2021) highlighted the application of Cox models in predicting recidivism but found that Bayesian regression models and deep neural networks provided superior predictive performance. This illustrates the evolving landscape of survival analysis, where traditional methods like the Cox model are complemented by more advanced techniques. (de la Cruz et al., 2021) We have chosen the Bayesian methods because Bayesian methods offer greater flexibility, probabilistic interpretation, and robustness in complex or small-sample scenarios.

Bayesian modelling has become increasingly popular in medical research due to its ability to incorporate prior knowledge and provide probabilistic interpretations of model parameters. For instance, Kim et al. (2013) used Bayesian inference to improve the accuracy of multivariate meta-analysis in evaluating cholesterol-lowering drugs. This approach allows for the integration of complex data structures and provides a more nuanced understanding of treatment effects (Kim et al., 2013).

Historical data received from the MoH can be used to define the priors in our modelling process, which provides more reliable estimates in situations where we have a small sample size or highly censored data. The Cox model relies on the assumption of proportional hazards, meaning that the hazard ratios are constant over time. The model can produce biased or misleading results if this assumption is violated. Bayesian methods can be more flexible in handling non-proportional hazards. The Cox model provides hazard ratios and p-values but does not provide direct probabilistic statements about parameters. Bayesian methods offer a full probabilistic interpretation, and complete characterisation of uncertainty in parameter estimates through posterior distributions. This can give more insights than point estimates and confidence intervals.

We have also considered the Bayesian Cox modelling, which combines Bayesian with Cox Proportional hazards model. The Bayesian Cox model is specifically designed for survival analysis, focusing on time-to-event data and the effect of covariates on survival times. Like the Bayesian, Bayesian Cox models incorporate prior distributions, allowing for the integration of expert knowledge or previous study results into the survival analysis. But the reason why we did not choose this is because it does not allow us to explore further than survival analysis. Bayesian approaches are not confined to survival analysis. They can be used in a variety of fields and for different types of data, offering more general applicability compared to the specialised focus of Bayesian Cox models. If the proportional hazards assumption of the Cox model is violated, other Bayesian methods, such as parametric Bayesian survival models or Bayesian hierarchical models, might be more appropriate. These methods do not rely on the proportional hazards assumption and can model more complex relationships.(Sheidaei et al., 2022)

A simple way of thinking about what Bayesian Analysis is would be to imagine you have travelled to an unknown planet with a blue sun above. What is the probability of the sun rising the next day? For the purpose of simplicity, we know the sun rises every day on Earth, so the priors are stated, and we would assume the blue sun will 100% the next day. You observe the blue sun rise on the first day, which reinforces your prior belief. If the sun doesn't rise one day, you adjust your belief, now estimating there's a 50% chance it will rise the following day. This process continues, each day's observation updating your probability estimate, reinforcing your belief about the blue sun's behaviour based on accumulating evidence. Bayesian analysis allows you to incorporate new data iteratively and continuously improve your predictions about the sun's rising pattern. Now, consider using the Cox Proportional Hazards model instead. You still want to know how often the blue sun rises, but you start with no prior assumptions based on Earth. You begin by observing and recording each day whether the sun rises. Over time, you gather data on sunrise frequency and other relevant factors, such as planetary rotation speed or weather patterns. Using this data, you build a Cox Proportional Hazards model to estimate the likelihood of the sun rising, analysing the time intervals between sunrises. This model assumes the hazard, or risk, of the sun not rising is proportional over time. As you collect more data, you update your model to refine the hazard ratio, adjusting for any identified factors affecting the sun's rise. Bayesian analysis incorporates prior knowledge (e.g. sun rises every day on Earth) to inform initial estimates, while the Cox model starts with no such priors and relies entirely on observed data. Bayesian analysis updates probabilities iteratively with each new piece of evidence, continuously refining the belief about the sun's rise. The Cox model updates the hazard ratio based on collected data, focusing on proportional hazards over time. Bayesian analysis offers a full probabilistic framework, allowing for updates and handling for small sample

sizes or incomplete data effectively. Although simpler and computationally efficient, the Cox model requires the assumption of proportional hazards and may struggle with complex data structures.

The Bayesian t-test offers several advantages over traditional frequentist methods, particularly in medical research. Bayesian methods allow for the incorporation of prior information, which is particularly useful in medical research where previous studies and clinical expertise can inform the analysis. This approach can improve the robustness and reliability of the results, especially in studies with small sample sizes or highly censored data. For example, Bayesian t-tests provide a framework for integrating expert knowledge or previous study results, thus enhancing the analysis of complex datasets.(Kelter, 2020) Bayesian t-tests offer a full probabilistic interpretation of model parameters, providing a complete characterisation of uncertainty through posterior distributions. This contrasts with the frequentist approach, which relies on point estimates and confidence intervals. Bayesian methods can yield richer information about the effect sizes and group means, allowing for better conclusions (Gronau et al., 2020) . Bayesian t-tests are flexible and can handle non-standard data situations, such as outliers or non-normal distributions. They are also well-suited for dealing with the Behrens-Fisher problem, which involves comparing means from two populations with different variances (Kelter, 2022) . Bayesian methods can improve the sensitivity and specificity of statistical tests by controlling for false discovery rates and detecting more truly significant effects. This is particularly important in fields like proteomics and genomics, where the detection of subtle differences is crucial (Millikin et al., 2020).

In summary, Bayesian methods allow for the incorporation of prior information and provide a probabilistic framework for making inferences about the parameters of interest. This approach is particularly useful in medical research, where prior studies and clinical expertise can inform the analysis. Bayesian models can account of uncertainty and variability in the data, which then gives credible estimates of the impact of asthma treatments on mortality.

3. Methods and material

This study consists of all individuals who has received asthma related medications in New Zealand between 1st January 2008 and 31 December 2021 (2,240,628). The data we used are taken from New Zealand Ministry of Health consisting of sex, age, medication name, date of death, and dosage the patient had taken over the time period. The permanent table generated is called Cohort. We used Microsoft SQL Server Studio (SMSS) from a Virtual Machine (VM) based in the University of

Auckland (UoA) to access and analyse data. After forming permanent tables on the SMSS, we are able to code in Jupyter Notebooks from another Virtual Machine to perform statistical analysis from tables in SMSS. The reason for looking at this is that there is evidence that LABA may be associated with mortality as seen in the literature review. salmeterol and eformoterol are two different types of LABA salmeterol has been around longer and has a slower onset of action compared with eformoterol which has a faster onset and is now the first option recommended in our guidelines. The data is useful for us to see how these two LABA compare in terms of mortality. There is also due to limiting raw computing power, we had to narrow it down to two medications: fluticasone with salmeterol and budesonide with eformoterol.

3.1 Use of artificial intelligence

I have used ChatGPT on SMSS Querying and forming skeleton codes in Jupyter notebooks. Particularly, due to the limiting compute power, the style of which I have written queries in SMSS is unoptimized which would have taken a long time to run or taken more RAM which the VM cannot run at all. It can help me form temporary and permanent tables as well as making comments in sections of the query to refresh memories in the future. In which I can store temporary data into further query on instead of saving all these into the RAM. Running the queries with the least amount of common table expression (CTE)'s also allows for the queries to run in parallel instead of in series which lowers the compute power and wait times for the queries to complete.

In Jupyter notebooks where I need to write Python codes to perform statistical analysis, I used ChatGPT to help me better improve and structure my codes in areas such as connecting to SMSS or giving me the skeleton codes for different distributions. In instances where I needed quick information about distributions for example, I would ask ChatGPT similar questions, and it could break up the answer to an easier to understand manner, if it fails to do so, I can ask it to explain the answer in more details. After enough understanding, I could go to Wikipedia to validate if what ChatGPT produced were correct and consistent with what Statistics By Jim states.

3.2 Documentation and Reproducibility

For the purpose of reproducible research, I try to use literal programming wherever possible. Specifically, naming variables and documents more easily understandable to the human eye. I named the documents in Jupyter Notebooks to be what methods it is using or the purpose of the document with its corresponding dates and which notebook it was cloned from. I have also written an introduction at the start of every document to explain the purpose of the notebook in what it wants to achieve and what it has achieved in the past where the document was cloned from. The

variables in my queries and codes are named as what they are such as medName in my SQL queries or meds_of_interest in my codes in Jupyter notebooks. Comments are made wherever there could be confusion in the future which could be reminded to the reader what the code or query is trying to achieve. Conclusions are also made to the results in Jupyter Notebooks to help the ease of reading in the future.

3.3 Formation of Cohort Table

Data about the patients prescribed with the medications in interest were first collected through SMSS. This process includes the formulation of the Cohort table with the SQL 1 query. Patient's information such as age and National Health Index (NHI) are included in this table. After initial cleaning and experimenting with the original cohort table, we realised we need to reduce the sample size due to our limiting compute power. We then selected Fluticasone with salmeterol (F) and Budesonide with eformoterol (B) as our medication of interest. SQL 2 query shows how this was done in SMSS to update the Cohort table. The last column of the Cohort table displays this information as either F, B, F_B, B_F or others. The group types are determined if the patient had been prescribed with such medications in history. F and B refers to if the patient had stayed on the one specified medication over the course of treatment. F_B and B_F refers to if the patient had made a transition from F to B or B to F in history. Other refers to patients who are not in these 4 groups.

After formulation of the data, we started estimating of priors for the Weibull Distribution. We have experimented with Wald (Inverse Gaussian) and Normal distribution and saw that it did not fit our data as well as what the Weibull distribution did as seen in the estimating of age distribution plot shown in Fig 1 constructed from Code 1. A Weibull distribution is assumed to be more suitable for our problem, mainly because Weibull distribution is defined for non-negative data, making it suitable for modelling time-to-failure and life data, where negative values are not meaningful. The Normal distribution, on the other hand, is symmetric and can take on negative values, which may not be appropriate for these our study as we do not allow negative age of death. Wald Distribution is similar to the Weibull Distribution as they are both right skewed. But Wald primarily models hazard functions that have an initial high hazard rate that decreases over time. This is reflected in Fig 2 as well as Wald is leaning more to the left than the other two distributions. It is less flexible than the Weibull distribution in modelling different hazard shapes.
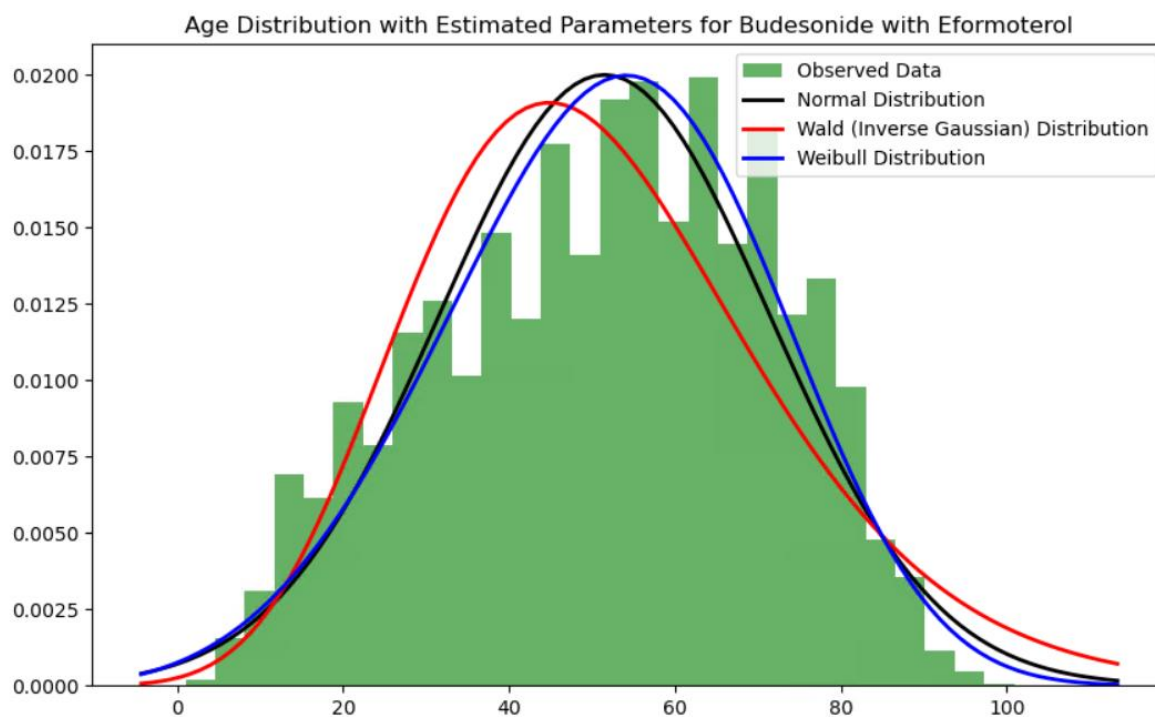
3.4 Choice of distribution for modelling age of death

Figure 1 Age Distribution with Estimated Parameters for Budesonide with Eformeterol

The reason behind the choosing of Weibull Distributions is because when estimating the age of death of patients, we can simply assume the patient cannot die before birth (age of 0). So, we looked for distributions that can be positively (right) skewed with the correct parameters. The four distributions that we experimented are Wald (Inverse Gaussian), Logit-normal, Rice, and Weibull. We first estimated the parameters and evaluated how they compare with current data. We fitted these distributions to our data and found that Weibull gave the best fitting (see in Figure 1). In the Logit-Normal and Rice distribution, we experimented with different estimated parameters but the HDI's (Highest Density Interval) included 0 which makes the range of credible values for the parameter include zero. The Logit-Normal distribution mostly looks at the scale and spread of the data rather than specific pinpoints, so we eventually turned to other distributions. We also experimented with the Gaussian Mixture model which similarly to the Weibull Distribution, provides a wide range in terms of skewness when set with the adequate parameters. This was eventually abandoned as Weibull Distribution was already giving promising results. This distribution could be experimented further and compared against Weibull in the future.

The analytic form of the Weibull distribution is as follows:

$$f(x; \lambda, k) = \begin{cases} \frac{k}{\lambda} \left(\frac{x}{\lambda} \right)^{k-1} e^{-(x/\lambda)^k}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

Figure 2 Probability density function of Weibull distribution (Pymc, n.d.)

Weibull, this distribution is characterised by its flexibility in modelling different types of hazard functions, which describe the likelihood of an event occurring over time. The probability density function (PDF) of the Weibull distribution is defined by two parameters: the scale parameter (λ) and the shape parameter (k). The scale parameter λ stretches or compresses the distribution along the x-axis, affecting the spread of the data. The shape parameter k determines the shape of the distribution; when k=1, it simplifies to an exponential distribution, k<1 indicates a decreasing hazard function, and k>1 indicates an increasing hazard function.

The packages I have used for the modelling of distributions in my Python environment are included in the Appendix as Package Code. A detailed description of the packages and parameters used are below.

From table 1, I have extracted the relevant group types using Code 2. The code uses several key Python packages, including SQLAlchemy, SciPy, PyMC, ArviZ, Pandas, NumPy, and Matplotlib, to connect to the database, extract and filter data, and perform probabilistic modeling.

To start, SQLAlchemy is utilised to establish a connection to the SQL Server database using Windows Authentication. The connection string, defined as `connection_str`, specifies the database details and the driver required for the connection. An engine is created with `create_engine(connection_str)`, allowing for database interactions. A SQL query is executed to extract data from the Cohort table, filtering out records with unspecified or invalid gender, non-positive study differences, and non-positive ages. The result is loaded into a Pandas DataFrame for further processing.

The dataset is then filtered to include only deceased patients who have been treated with either "Fluticasone with salmeterol" or "Budesonide with eformoterol". This subset of data is crucial for analysing the effects of these medications on different patient groups. The unique group types within this subset are identified, and the analysis proceeds by modelling the age distribution of deceased patients within each group using a Bayesian framework.

For each group type, a Bayesian model is defined using PyMC. The priors for the Weibull distribution parameters, alpha and beta, are specified as half-normal distributions with given standard deviations. The likelihood function, representing the observed age data, is modelled using the

Weibull distribution with the defined priors. Posterior sampling is performed to estimate the parameters, with 1000 samples and 500 tuning steps to ensure convergence. In machine learning sense, this could be understood as learning rates and learning steps. The higher the number, the bigger the iterations, which is why I have chosen those specific numbers. The results, including the model and trace data, are stored for each group type.

To visualize the results, trace plots of the posterior distributions are generated using ArviZ and saved as PNG files. Summaries of the Weibull parameters, including mean and highest density intervals (HDI), are printed for each group type, providing insights into the parameter estimates. Additionally, differences in alpha and beta parameters between group types are calculated. The mean, standard deviation, and HDIs for these differences are summarised in DataFrames and saved as CSV files for further analysis. These summaries facilitate a comparative analysis of the parameter estimates across different patient groups, highlighting any significant variations in the effects of the medications.

In table 2, the medication group types are from the cohort table generated as shown in table 1. Main difference is everyone in this table is deceased after receiving such medications. Similar to table 1, F and B refers to if the patient had only taken these two specific medications over the course of treatment and had not made any transition from one to another. B_to_F refers to if the patient has made the switching of medication B to F at least once during the study time, they are in this group type and will never be able to return back to B or F only. The same idea applies for group F_to_B. Alpha refers to the shape parameter k difference. This could be interpreted as an indicator in estimated mean standard deviation of age of death between the specified groups. Beta refers to the scale parameter λ difference. This could also be interpreted as an indicator in estimated mean age of death between the two groups.

4. Results

In total 2,240,628 people has received asthma related medications in this study. Their characteristics at the time of inclusion are reported in Table 1.

Results

Table 1. Population characteristics at time of inclusion.

	Population Characteristics	
	N	**%**
All	2,240,628	100%
Sex		
Male	1,058,117	47.22%
Female	1,182,511	52.78%
Age		
0-9	492,075	15.83%
10-19	316,433	10.18%
20-29	286,731	9.22%
30-39	328,398	10.57%
40-49	380,690	12.25%
50-59	405,969	13.06%
60-69	386,690	12.44%
70-79	304,052	9.78%
80-89	170,267	5.48%
90+	37,023	1.19%
Medication Group Type		
Budesonide with eformoterol (B)	17,342	0.77%
Fluticasone with salmeterol (F)	6,531	0.29%
B_to_F	10,652	0.48%
F_to_B	31,194	1.39%
Other	2,174,909	97.07%

Results

The population is nearly evenly split between males (47.22%) and females (52.78%),
indicating a balanced representation of genders in the dataset. The age distribution shows
that the majority of the population falls within the age range of 0-9 years (15.83%), with a
noticeable decline in the proportion of older age groups. This distribution is typical for a
general population where younger age groups tend to be more numerous. A small
proportion of the population is treated with budesonide with eformoterol (0.77%) or
fluticasone with salmeterol (0.29%). Additionally, there is a significant movement between
medications, with 0.48% of patients switching from budesonide to Fluticasone (B_to_F) and
1.39% switching from fluticasone to budesonide (F_to_B). The vast majority (97.07%) of the
population falls into the "Other" category, which indicates that these patients are either not
using the specified medications or are using different treatments. This suggests that the
focus on Budesonide and Fluticasone represents a small subset of the overall medication
usage.

Table 2. Summary of difference in age at death for different medication group types

Parameter Difference	Mean	SD	HDI 3%	HDI 97%
alpha (B - F)	0.61	0.24	0.17	1.06
beta (B - F)	-1.33	0.39	-2.07	-0.59
alpha (B - B_to_F)	0.89	0.17	0.58	1.22
beta (B - B_to_F)	2.59	0.25	2.13	3.05
alpha (B - F_to_B)	1.92	0.17	1.61	2.25
beta (B - F_to_B)	4.83	0.26	4.34	5.33
alpha (B - Other)	1.11	0.16	0.81	1.42
beta (B - Other)	1.97	0.23	1.54	2.41
alpha (F - B_to_F)	0.29	0.19	-0.08	0.64
beta (F - B_to_F)	3.91	0.32	3.31	4.53
alpha (F - F_to_B)	1.32	0.19	0.94	1.67

Results

beta (F - F_to_B)	6.16	0.33	5.54	6.78
alpha (F - Other)	0.50	0.18	0.14	0.84
beta (F - Other)	3.30	0.31	2.72	3.88
alpha (B_to_F - F_to_B)	1.03	0.07	0.89	1.17
beta (B_to_F - F_to_B)	2.25	0.14	1.99	2.52
alpha (B_to_F - Other)	0.22	0.05	0.12	0.31
beta (B_to_F - Other)	-0.61	0.08	-0.77	-0.45
alpha (F_to_B - Other)	-0.81	0.06	-0.92	-0.71
beta (F_to_B - Other)	-2.86	0.12	-3.10	-2.63

The table above presents the differences in the shape (alpha) and scale (beta) parameters between various medication groups. The mean, standard deviation (SD), and Highest Density Interval (HDI) for the parameter differences are provided. Here, F refers to fluticasone with salmeterol and B refers to budesonide with eformoterol. The differences in the shape and scale parameters (alpha and beta) between budesonide with eformoterol (B) and fluticasone with salmeterol (F) indicate notable variations in the estimated age of death. For instance, budesonide with eformoterol shows a higher shape parameter (alpha) and a lower scale parameter (beta) compared to fluticasone with salmeterol. Significant differences are also observed between other groups, such as those transitioning from one medication to another (B_to_F and F_to_B). These differences highlight variations in survival times and risk profiles across different treatment groups. The Highest Density Intervals (HDIs) provide a credible range for the parameter differences, helping to understand the uncertainty around these estimates. For some comparisons, the HDI ranges do not include zero, suggesting significant differences between the groups.

The following trace plots illustrate the posterior distributions and the sampling chains for the shape (alpha) and scale (beta) parameters of the Weibull distribution for different medication groups.

Figure 3. Trace plot of age at death for Group B in Table 2

The posterior distribution of the shape parameter shows a peak around 8.4, indicating a specific tendency in the age of death distribution for the B group. The posterior distribution of the scale parameter peaks around 80, showing the spread of the age distribution

Figure 4. Trace plot of age at death for Group F in Table 2

The posterior distribution of the shape parameter shows a peak around 7.9, which is lower than the B group, suggesting different age of death distribution characteristics. The posterior distribution of the scale parameter peaks around 81.5, indicating a different spread compared to the B group.
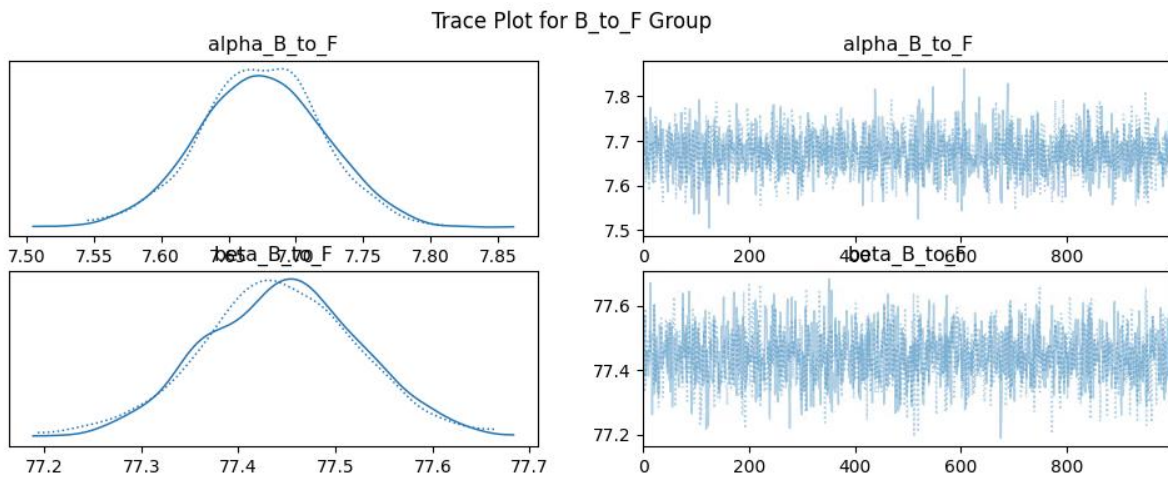
Figure 5. Trace plot of age at death for Group B_to_F in Table 2

The posterior distribution shows a peak around 7.7, indicating the shape of the age of death distribution for patients who switched from budesonide with eformoterol to fluticasone with salmeterol.
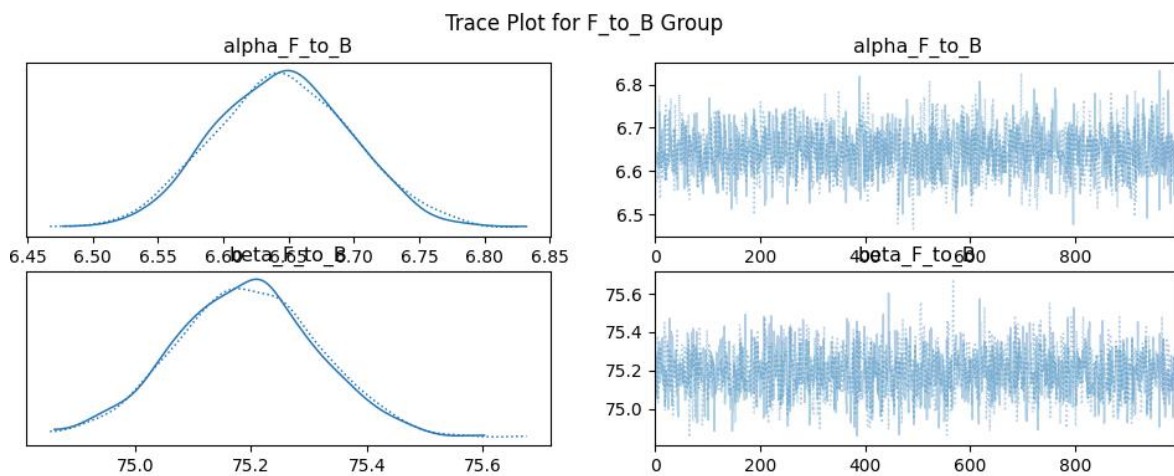
Figure 6. Trace plot of age at death for Group F_to_B in Table 2

The shape parameter peaks around 6.7, which is lower than the B_to_F group, indicating different age of death distribution characteristics. The scale parameter peaks around 75.4, showing a different spread compared to the B_to_F group.
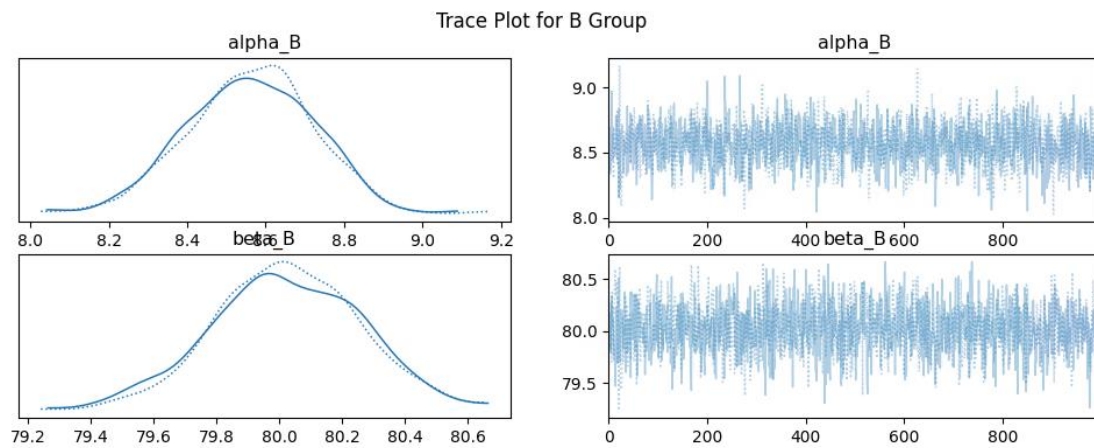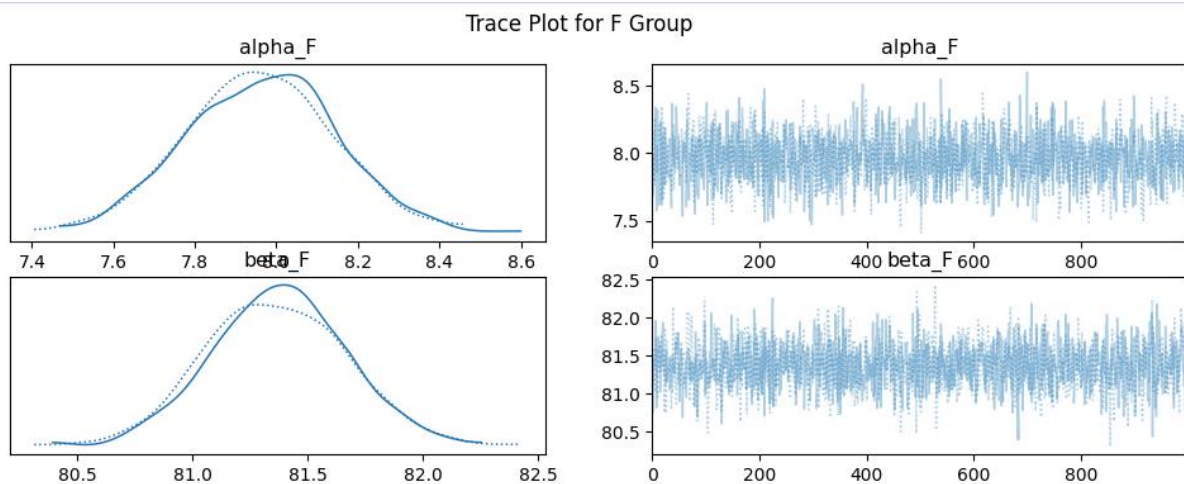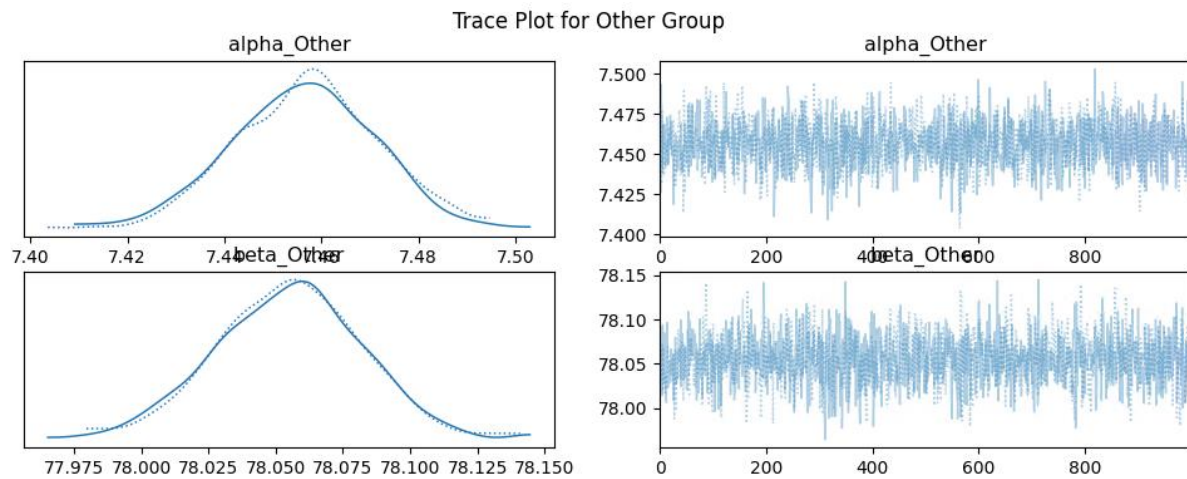
Figure 7. Trace plot of age at death for Group Other in Table 2

The posterior distribution of the shape parameter peaks around 7.45, indicating the shape of the age of death distribution for the "Other" group. The posterior distribution peaks around 78.1, showing the spread of the age distribution for this group.

The B_to_F and F_to_B groups show distinct parameter distributions compared to the groups that did not switch medications, indicating the impact of medication transitions on the age of death distributions.

Table 3 Medication taken in all of cohort

Medication Taken		
Beclomethasone dipropionate	331,296	7.26%
Budesonide	279,208	6.12%
Budesonide with eformoterol	216,728	4.75%
Dexamethasone	148,673	3.26%
Eformoterol fumarate	33,371	0.73%
Eformoterol fumarate dihydrate	345	0.01%
Fluticasone	553,893	12.14%

Fluticasone furoate with vilanterol	74,846	1.64%
Fluticasone with salmeterol	220,574	4.84%
Indacaterol	2,141	0.05%
Methylprednisolone	2,472	0.05%
Prednisolone	292,765	6.42%
Prednisone	1,052,627	23.08%
Salbutamol	1,190,613	26.10%
Salmeterol	78,256	1.72%
Terbutaline sulphate	83,836	1.84%

The most frequently taken medications are Salbutamol (26.10%) and Prednisone (23.08%), indicating their widespread use in the population. Fluticasone (12.14%) and Budesonide (6.12%) are commonly used inhaled corticosteroids. Budesonide with eformoterol (4.75%) and Fluticasone with salmeterol (4.84%) are notable combination therapies used by a significant portion of the population. Indacaterol and Methylprednisolone are among the least frequently taken medications, each accounting for just 0.05% of the population

5. Discussions

5.1 Key findings

The analysis presented in Tables 1 and 2 offers significant insights into the population characteristics, medication usage, and the differences in age at death among various medication groups. Here we summarise and interpret the key findings from these tables to provide a comprehensive understanding of the results.

Table 1 details the demographics and medication usage of the 2,240,628 individuals included in the study. The key observations are as follows:

Discussions

- The population is almost evenly divided by gender, with 47.22% male and 52.78% female.

- The age distribution shows a higher proportion of younger individuals, with 15.83% aged 0-9 years, decreasing steadily across older age groups.

- Regarding medication usage, a small fraction of the population is treated with budesonide with eformoterol (0.77%) or fluticasone with salmeterol (0.29%). Notably, there is a significant transition between medications, with 0.48% of patients switching from budesonide with eformoterol to fluticasone with salmeterol (B_to_F) and 1.39% switching from fluticasone to budesonide (F_to_B). The majority (97.07%) fall into the "Other" category, suggesting diverse medication usage outside the primary focus of budesonide and fluticasone.

Table 2 presents the differences in the shape (alpha) and scale (beta) parameters between various medication groups, indicating variations in estimated age at death. The key findings are:

- Budesonide with eformoterol (B) shows a higher shape parameter (alpha) and a lower scale parameter (beta) compared to fluticasone with salmeterol (F), suggesting that patients on budesonide with eformoterol have a different distribution of age at death compared to those on fluticasone with salmeterol.

- Patients who transitioned from fluticasone with salmeterol to budesonide with eformoterol (F_to_B) have significant differences in both alpha and beta parameters compared to those on fluticasone (F). Specifically, the mean difference in the scale parameter (beta) is -6.16 years, indicating that transitioning from fluticasone to budesonide is associated with a significantly lower age at death.

- Conversely, those who transitioned from Budesonide with eformoterol to fluticasone with salmeterol (B_to_F) show a mean difference of 3.91 years in beta compared to fluticasone, suggesting a higher age at death after the transition.

- Compared to the "Other" medication group, both budesonide and fluticasone groups show significant differences in the age at death distributions. Budesonide with eformoterol has a mean difference of 1.97 years (beta) compared to the "Other" group, while fluticasone with salmeterol has a difference of 3.30 years (beta).

Table 3 lists the different medications taken by the cohort. The most frequently used medications are:

- Salbutamol (26.10%) and Prednisone (23.08%), indicating their widespread use for immediate relief and long-term management.

22

Discussions

- Fluticasone (12.14%) and Budesonide (6.12%) are commonly used inhaled corticosteroids.
- Combination therapies such as Budesonide with eformoterol (4.75%) and Fluticasone with salmeterol (4.84%) are also notable.

In summary, in Table 2, we highlighted three significant comparisons between different medication groups based on the differences in the shape (alpha) and scale (beta) parameters of the Weibull distribution, which model the age of death.

Alpha (B - F): The difference in the shape parameter between Budesonide with eformoterol (B) and Fluticasone with salmeterol (F) is 0.61, with a mean difference and an HDI ranging from 0.17 to 1.06. This indicates a notable variation in the age of death distribution's shape, suggesting that B leads to a slightly more spread-out distribution of ages at death compared to F.

Beta (B - F): The scale parameter difference is -1.33, with an HDI from -2.07 to -0.59, showing that the age of death distribution for B is less than F. Patients on B have a shorter mean age at death compared to those on F.

Alpha (F - F_to_B): The shape parameter difference between F and F_to_B is 1.32, with an HDI from 0.94 to 1.67. This substantial difference indicates that switching from F to B significantly alters the age of death distribution's shape, making it more spread out for the F group.

Beta (F - F_to_B): The scale parameter difference is 6.16, with an HDI from 5.54 to 6.78. This suggests that, on average, patients on F live approximately 6.16 years longer than those who transitioned from F to B. This is the largest observed difference in our study, indicating a significant impact of the medication switch on survival. This box is also highlighted to indicate its importance.

Beta (B_to_F - Other): The scale parameter difference between patients from B_to_F and the "Other" group is -0.61, with an HDI from -0.77 to -0.45. This indicates that B_to_F group has a slightly lower mean age at death compared to the "Other" group, highlighting the impact of this medication switch on survival.

Significant differences are observed in both alpha and beta parameters for patients who transitioned between medications (B_to_F and F_to_B). These differences underscore the potential impact of medication switches on survival times. The transitions from one medication to another appear to affect the age of death distribution, with notable variations in the shape and scale parameters.

The trace plots (Figures 3) provide visual insights into these parameter estimates. For instance, the trace plot for the B group shows a peak around 8.4 for the alpha parameter and 80 for the beta

parameter, indicating specific tendencies in the age of death distribution for this group. Similar interpretations can be made for the other groups, showing the distinct parameter distributions and the impact of medication transitions.

5.2 Comparison with the existing literature.

Previous studies have extensively investigated the role of different asthma medications in influencing mortality rates. Inhaled corticosteroids (ICS) are widely recognised for their effectiveness in reducing inflammation and preventing exacerbations. For instance, Faisal and Yunus (2019) demonstrated a significant decline in asthma mortality rates in New Zealand with increased use of ICS, following the asthma deaths epidemic linked to high-dose inhaled fenoterol in the 1970s. Similarly, combination therapy of ICS with LABA has been shown to improve asthma control and reduce mortality, while the use of LABA alone has raised concerns due to potential risks of severe exacerbations (Wijesinghe et al., 2008). Short-acting beta-agonists (SABA), though effective for immediate relief, have been linked to increased mortality risk when used without concurrent anti-inflammatory treatment, underscoring the importance of ICS as the gold standard of long-term management.

Research from countries like the United States, the United Kingdom, and Australia also supports the beneficial impact of ICS on reducing mortality. For example, Ulrik and Frederiksen (1995) found a 20% reduction in asthma mortality associated with high-dose ICS use in the UK, while Tatham and Gellert (1985) highlighted a 30% increased mortality risk with LABA use without ICS in the US. Douglass and Bowes (1990) further emphasised the positive impact of increased ICS use through interrupted time-series analysis, which showed improved outcomes in asthma management.

Studies focusing on the New Zealand context have highlighted disparities in asthma mortality rates among different population groups. Ellison-Loschmann et al. (2008) reported disproportionately higher asthma mortality rates among Māori compared to non-Māori, with significant peaks in the late 1970s. Crengle et al. (2022) identified rural-urban disparities in asthma mortality rates among Māori, with rural Māori experiencing greater mortality rates. These findings underscore the need for targeted interventions and improved access to healthcare services for these populations.

Previous studies used various methodologies to analyse the impact of asthma medications on mortality. Faisal and Yunus (2019) utilised epidemiological analysis to examine trends in asthma mortality rates over time. Wijesinghe et al. (2008) conducted a meta-analysis and regression models

to assess the risks associated with LABA use, focusing on combination therapy with ICS. Ulrik and Frederiksen (1995) employed logistic regression to assess the association between high-dose ICS and asthma mortality. Tatham and Gellert (1985) used meta-analysis with fixed and random-effects models to evaluate mortality risk associated with LABA use without ICS. Douglass and Bowes (1990) applied interrupted time-series analysis to assess the impact of increased ICS use on mortality rates. Ellison-Loschmann et al. (2008) conducted longitudinal data and time-series analysis to examine trends in asthma mortality rates among different ethnic groups. Crengle et al. (2022) utilised logistic regression models to explore rural-urban disparities in asthma mortality rates among Māori.

Our study uses Bayesian survival analysis with the Weibull distribution to model time-to-event data, unlike the traditional regression or meta-analysis approaches used in previous studies. This method allows for the incorporation of prior information and provides a probabilistic framework for parameter estimation, offering a better understanding of the impact of asthma medications on mortality rates. By leveraging extensive datasets from the New Zealand Ministry of Health (MoH), our study integrates comprehensive longitudinal data. This enhances the robustness and reliability of our findings.

Our Bayesian approach provides a full probabilistic interpretation of model parameters, allowing for a more comprehensive understanding of uncertainty and variability in the data. This is a significant improvement over the point estimates and confidence intervals provided by traditional methods. The use of parametric Bayesian survival models allows us to model non-proportional hazards, addressing a key limitation of the Cox proportional hazards model, which assumes proportional hazards. The Bayesian hierarchical models we use are more flexible and robust in handling complex relationships and small sample sizes, providing more reliable estimates in situations where traditional methods may struggle.

5.3 Strengths and weaknesses of my methods and data

One of the primary strengths of our Bayesian approach is its ability to provide a full probabilistic interpretation of model parameters. This allows for a more comprehensive understanding of uncertainty and variability in the data, which is a significant improvement over the point estimates and confidence intervals provided by traditional methods. Traditional methods like logistic regression and Cox proportional hazards models typically offer point estimates and confidence intervals, which do not fully capture the uncertainty and variability in parameter estimates.

Additionally, the use of parametric Bayesian survival models in our study allows for the modelling of non-proportional hazards. This flexibility is crucial because it addresses a key limitation of the Cox

Discussions

proportional hazards model, which assumes proportional hazards. This assumption, if violated, can lead to biased or misleading results in traditional models. Our Bayesian methods enable more accurate modelling of complex survival data, providing a clear advantage over conventional approaches.

Our methods are more adaptive and can provide more reliable estimates in situations where traditional methods may struggle. Traditional methods like logistic regression and fixed/random-effects models used in meta-analyses may not be as flexible or robust in handling complex data structures or small sample sizes.

However, our Bayesian methods are not without weaknesses. One significant drawback is their computational complexity. Bayesian methods, particularly those involving hierarchical models and complex distributions like the Weibull distribution, can be computationally intensive and require significant computational resources and expertise. In contrast, traditional methods like logistic regression and Cox proportional hazards models are generally less computationally demanding and easier to implement, making them more accessible for researchers with limited computational resources.

The complexity of interpreting the probabilistic outputs of Bayesian models is another potential weakness. These outputs, such as posterior distributions, can be more complex to interpret and communicate to stakeholders who are more familiar with traditional point estimates and confidence intervals. Traditional methods provide more straightforward results (e.g., hazard ratios, odds ratios) that are easier to interpret and communicate to a broad audience, including non-specialists.

Bayesian models can be sensitive to the specification of the model structure and priors. Incorrect model specification can lead to biased results, requiring rigorous model checking and validation. While traditional methods also require appropriate model specification, they are generally perceived as less sensitive to the exact form of the model compared to Bayesian methods.

We are also only comparing two medications within the SABA group. F and B do not necessarily have a high usage when compared to the whole asthma patient's population as seen in Table 3. Our conclusions could have been more reliable and important if the sample size was a lot higher.

We did not control for other covariates such as the severity of the patient's conditions which may eventually determine the likelihood of a patient's survival after treatment. The patient's ability to access to healthcare services were also not considered. Because the data were collected from MoH, we were unable to determine if the patient had taken the prescribed medication according to instructions or even taken the medication at all.

Lastly, we could not conclude with an integer that specifies the age of death difference between patients in different medication groups. Alpha and Beta are only parameters of the Weibull Distribution which means the numbers from table 2 only indicate that there is a substantial difference. We could not conclude with actual digits that specifies how much longer the patient could have lived if the patient did not make the transition for example.

5.4 Where to next

In terms of next steps, we could consider the severity of the patients through the ranking of medication groups that the patient is taking. The table could be called Severity steps and based on what medication the patient is taking or has taken in the specified date, we could rank the severity of the patient's conditions based on the medication group he/she lies in. Medication groups could be classified as SABA, LABA, or ICS. An exemplar query is shown in SQL query 3 of the appendix.

we could also calculate the yearly dosage of the patient before death which in conjunction with the severity steps, we will have a better understanding of the patient's condition. An exemplar query is shown in SQL query 4 of the appendix. There are three queries in the box. We should be mindful however, when in SMSS, we shouldn't allow for null values in between column names which is why I have used dynamic PIVOT query to put "_" in between medication names to fill in blank spaces in the column names. This is calculated by putting an identifier on the life status of the patient. We could then calculate the patterns in the dosage for the living. We could also calculate the yearly dosage of specific medications before death which eventually lead to a pattern of dosage which may result in probable cause of patient's death.

```
# For Budesonide with eformoterol
budesonide_weibull_mean, budesonide_weibull_std = weibull_stats(7.346, 77.140)
```

Figure 8. Utilisation of the weibull_stats function

In order for better conclusions about the parameter estimations. We could also utilise code 2 in the Appendix which converts the Weibull parameter estimates into actual numbers through reversion in the Weibull formula. The PyMC package does not have such function which could convert estimated parameters back to estimated numbers. Figure 8 shows the usage of the Weibull_stats function. The input to the function is simply the two numbers from estimated parameters which were calculated beforehand.

6. Summary

In conclusion, asthma is a chronic respiratory condition that significantly impacts global health, contributing to substantial morbidity and mortality. Various pharmacological treatments, including inhaled corticosteroids (ICS), long-acting beta-agonists (LABA), and short-acting beta-agonists (SABA), are widely used to manage asthma symptoms and prevent severe exacerbations. The effectiveness of these treatments in reducing asthma-related mortality has been extensively studied, with findings indicating that ICS, particularly when used in combination with LABA, can significantly reduce mortality rates. However, the use of LABA alone has been associated with increased mortality risks, and reliance on SABA without concurrent anti-inflammatory treatment has been linked to higher mortality rates. Specifically, this study has found that there is strong indication that patients who have transitioned from fluticasone with salmeterol to budesonide with eformoterol have significant differences in mortality rate, their mean difference in scale parameter (beta) is -6.16 which indicates transitioning from fluticasone with salmeterol to budesonide with eformoterol is associated with a significantly lower age at death.

This study uses Bayesian methods with the Weibull distribution to model mortality rate. By using extensive datasets from the New Zealand Ministry of Health, this study integrates comprehensive longitudinal data, allowing for the control of a wide range of potential confounders. The findings indicate significant differences in mortality outcomes based on medication types and transitions, with Bayesian methods offering greater flexibility and robustness compared to traditional statistical approaches. The study highlights that patients on Fluticasone with salmeterol tend to have a higher mean age of death compared to those on Budesonide with eformoterol or those who transitioned between these medications. This research provides new insights into the survival times associated with different asthma treatments, informing clinical decisions and guiding further research on the impact of medication transitions.

To advance this work, future research should focus on expanding the scope of analysis to include additional asthma medications, such as prednisone and salbutamol, which are more frequently used. This would provide a more comprehensive understanding of the impact of various asthma treatments on mortality. Further studies should also explore the long-term effects of transitioning between different medications, controlling for a broader range of confounders, such as environmental factors and genetic predispositions. Additionally, methodological improvements, including the use of more advanced Bayesian hierarchical models and parametric survival models, can enhance the robustness of findings. Collaborative efforts involving larger and more diverse datasets, both nationally and internationally, would help validate the results and provide a more generalized understanding of

Summary

asthma treatment outcomes. Finally, targeted interventions and policy recommendations should be developed to address the identified disparities in asthma mortality, particularly among vulnerable populations such as the Māori in New Zealand.

References

A. Douglass, J., & Reddel, H. K. (2005). Double or nothing? Treating exacerbations of asthma. *Internal Medicine Journal*, *35*(12), 689–691. https://doi.org/https://doi.org/10.1111/j.1445-5994.2005.00995.x

BEASLEY, R. (2007). Unmet need in inadequately controlled asthma. *Respirology*, *12*(s3), S18–S21. https://doi.org/https://doi.org/10.1111/j.1440-1843.2007.01046.x

Douglass, J. A., & Bowes, G. (1990). Inhaled corticosteroids in asthma. *Medical Journal of Australia*, *152*(9), 475–479. https://doi.org/https://doi.org/10.5694/j.1326-5377.1990.tb125307.x

Ellison-Loschmann, L., Gray, M., Cheng, S., & Pearce, N. (2008). Follow-up Study of Asthma Severity in Maori Adolescents. *Australasian Epidemiologist*, *15*(2), 4–10. https://search-informit-org.ezproxy.auckland.ac.nz/doi/10.3316/informit.317089625895078

Faisal, H. K. P., & Yunus, F. (2019). Asthma death. *Pneumologia*, *68*(4), 162–168. https://doi.org/doi:10.2478/pneum-2019-0030

Gronau, Q. F., Ly, A., & Wagenmakers, E.-J. (2020). Informed Bayesian t-Tests. *The American Statistician*, *74*(2), 137–143. https://doi.org/10.1080/00031305.2018.1562983

Kelter, R. (2020). Analysis of Bayesian posterior significance and effect size indices for the two-sample t-test to support reproducible medical research. *BMC Medical Research Methodology*, *20*(1), 88. https://doi.org/10.1186/s12874-020-00968-2

Kelter, R. (2022). A New Bayesian Two-Sample t Test and Solution to the Behrens–Fisher Problem Based on Gaussian Mixture Modelling with Known Allocations. *Statistics in Biosciences*, *14*(3), 380–412. https://doi.org/10.1007/s12561-021-09326-2

Kim, S., Chen, M.-H., Ibrahim, J. G., Shah, A. K., & Lin, J. (2013). Bayesian inference for multivariate meta-analysis Box–Cox transformation models for individual patient data with applications to evaluation of cholesterol-lowering drugs. *Statistics in Medicine*, *32*(23), 3972–3990. https://doi.org/https://doi.org/10.1002/sim.5814

Millikin, R. J., Shortreed, M. R., Scalf, M., & Smith, L. M. (2020). A Bayesian Null Interval Hypothesis Test Controls False Discovery Rates and Improves Sensitivity in Label-Free

References

Quantitative Proteomics. *Journal of Proteome Research*, *19*(5), 1975–1981. https://doi.org/10.1021/acs.jproteome.9b00796

Phillips, B., Daniels, J., Woodward, A., Blakely, T., Taylor, R., & Morrell, S. (2017). Mortality trends in Australian Aboriginal peoples and New Zealand Māori. *Population Health Metrics*, *15*(1), 25. https://doi.org/10.1186/s12963-017-0140-6

Sheidaei, A., Foroushani, A. R., Gohari, K., & Zeraati, H. (2022). A novel dynamic Bayesian network approach for data mining and survival data analysis. *BMC Medical Informatics and Decision Making*, *22*(1), 251. https://doi.org/10.1186/s12911-022-02000-7

Tatham, M. E., & Gellert, A. R. (1985). The management of acute severe asthma. *Postgraduate Medical Journal*, *61*(717), 599–606. https://doi.org/10.1136/pgmj.61.717.599

Ulrik, C. S., & Frederiksen, J. (1995). Mortality and Markers of Risk of Asthma Death Among 1,075 Outpatients With Asthma. *Chest*, *108*(1), 10. https://link.gale.com/apps/doc/A54099410/AONE?u=learn&sid=bookmark-AONE&xid=866ead2e

Wang, Z., Li, Y., Gao, Y., Fu, Y., Lin, J., Lei, X., Zheng, J., & Jiang, M. (2023). Global, regional, and national burden of asthma and its attributable risk factors from 1990 to 2019: a systematic analysis for the Global Burden of Disease Study 2019. *Respiratory Research*, *24*(1), 169. https://doi.org/10.1186/s12931-023-02475-6

Wijesinghe, M., Perrin, K., Harwood, M., Weatherall, M., & Beasley, R. (2008). The risk of asthma mortality with inhaled long acting β-agonists. *Postgraduate Medical Journal*, *84*(995), 467–472. https://doi.org/10.1136/pgmj.2007.067165

Crengle, S., Davie, G., Whitehead, J., de Graaf, B., Lawrenson, R., & Nixon, G. (2022). Mortality outcomes and inequities experienced by rural Māori in Aotearoa New Zealand. The Lancet regional health. Western Pacific, 28, 100570. https://doi.org/10.1016/j.lanwpc.2022.100570

PyMC. (n.d.). `pymc.Weibull`. PyMC documentation. Retrieved June 23, 2024, from https://www.pymc.io/projects/docs/en/stable/api/distributions/generated/pymc.Weibull.html

de la Cruz, R.; Padilla, O.; Valle, M.A.; Ruz, G.A. Modeling Recidivism through Bayesian Regression Models and Deep Neural Networks. *Mathematics* **2021**, *9*, 639. https://doi.org/10.3390/math9060639

References

de la Cruz, R.; Padilla, O.; Valle, M.A.; Ruz, G.A. Modeling Recidivism through Bayesian Regression

Models and Deep Neural Networks. Mathematics 2021, 9, 639.

https://doi.org/10.3390/math9060639

OpenAI. (2024). *ChatGPTo* (May 14 version) [Large language

model]. https://chat.openai.com/chat

Appendix

SQL 1 Generating cohort table

```
USE moh_2000_2021

go

--identified whichever drug we need
WITH dim
    AS (SELECT DISTINCT moh.ref_pharms_codes.dim_form_pack_subsid
y_key AS
                        DIM_FORM_PACK_SUBSIDY_KEY,
                        moh.ref_pharms_codes.chemical_id
     AS
                        chemID,
                        chemical_name
        FROM    moh.ref_pharms_codes
        WHERE   moh.ref_pharms_codes.chemical_id IN
                ( '1083', '4112', '4042', '1066',
                  '1108', '1168', '3758', '1065',
                  '4056', '3858', '2096', '2404',
                  '2038', '2034', '1811', '1383' ))
--
picked all the people dispensed with our drugs and all their perso
nal information from the pharms table
SELECT moh.pharmaceuticals.moh_encmasternhi      AS NHI,
       moh.pharmaceuticals.ageatdispensing       AS age,
       moh.pharmaceuticals.gender                AS gender,
       dim.chemid                                AS chemID,
       dim.chemical_name                         AS chemName,
       moh.pharmaceuticals.datedispensed         AS DateDispensed,
       pharmaceuticals.quantitydispensed,
       pharmaceuticals.dim_form_pack_subsidy_key AS subsidyKey
INTO    jason_jli404.dbo.pharm
```

```sql
FROM    moh.pharmaceuticals
        INNER JOIN dim
                ON moh.pharmaceuticals.dim_form_pack_subsidy_key =
                    dim.dim_form_pack_subsidy_key
WHERE   moh.pharmaceuticals.date_dispensed >= '2008-01-01'


go


SELECT  nhi,
        chemid,
        chemname,
        Min(age)            AS age,
        gender,
        Min(datedispensed) AS indexDate,
        Max(datedispensed) AS endDate,
        Cast(NULL AS BIT)  AS Flag,
        Cast(NULL AS DATE) AS exitDate,
        Cast(NULL AS INT)  AS dayDiff,
        Cast(NULL AS INT)  AS studyDiff,
        Cast(NULL AS INT)  AS dose
INTO    jason_jli404.dbo.cohort
FROM    jason_jli404.dbo.pharm
GROUP   BY nhi,
            chemid,
            chemname,
            age,
            gender;


SELECT DISTINCT nhi
INTO    ##cohort
FROM    jason_jli404.dbo.cohort


SELECT  moh_encmasternhi AS NHI,
        dateofdeath
INTO    jason_jli404.dbo.death
FROM    moh_2000_2021.moh.mortality_unified
WHERE   moh_encmasternhi IN (SELECT nhi
```

```sql
                              FROM    ##cohort);


SELECT  DISTINCT nhi
INTO    ##events
FROM    jason_jli404.dbo.death


UPDATE  jason_jli404.dbo.cohort
SET     flag = CASE
                  WHEN nhi IN (SELECT nhi
                                 FROM    ##events) THEN 1
                  ELSE 0
               END;


UPDATE  jason_jli404.dbo.cohort
SET     exitdate = jason_jli404.dbo.death.dateofdeath
FROM    jason_jli404.dbo.death
WHERE   jason_jli404.dbo.cohort.nhi = jason_jli404.dbo.death.nhi;


UPDATE  jason_jli404.dbo.cohort
SET     exitdate = '2021-12-31'
WHERE   flag = 0;


UPDATE  jason_jli404.dbo.cohort
SET     daydiff = Datediff(dd, indexdate, enddate);


UPDATE  jason_jli404.dbo.cohort
SET     studydiff = Datediff(dd, indexdate, exitdate);


SELECT  p.nhi,
        Sum(quantitydispensed) AS totalDose
INTO    ##totaldose
FROM    jason_jli404.dbo.pharm AS p
        INNER JOIN jason_jli404.dbo.cohort AS c
                ON p.nhi = c.nhi
WHERE   datedispensed BETWEEN c.indexdate AND c.exitdate
GROUP   BY p.nhi
```

```sql
UPDATE jason_jli404.dbo.cohort
SET    dose = totaldose
FROM   ##totaldose
WHERE  jason_jli404.dbo.cohort.nhi = ##totaldose.nhi;
```

SQL 2 Incorporating group type into Cohort table

```sql
USE jason_jli404

go


-- Drop the GroupType column if it exists
IF EXISTS(SELECT *
          FROM   information_schema.columns
          WHERE  table_name = 'Cohort'
                 AND column_name = 'GroupType')
  ALTER TABLE dbo.cohort
    DROP COLUMN grouptype;


go


-- Add the GroupType column again
ALTER TABLE dbo.cohort
  ADD grouptype VARCHAR(50);


go


-
- Query to determine medication transitions, and label the group t
ypes
WITH prescriptionhistory
     AS (SELECT nhi,
                chemid,
                Min(datedispensed) AS FirstDate,
                Max(datedispensed) AS LastDate
         FROM    dbo.pharm
```

```sql
        GROUP  BY nhi,
                  chemid),
    transitions
    AS (SELECT p1.nhi,
                CASE
                   WHEN p1.chemid = '3858'
                         AND p2.chemid = '3758' THEN 'F_to_B'
                   WHEN p1.chemid = '3758'
                         AND p2.chemid = '3858' THEN 'B_to_F'
                END AS GroupType
         FROM    prescriptionhistory p1
                JOIN prescriptionhistory p2
                   ON p1.nhi = p2.nhi
                      AND p1.lastdate < p2.firstdate
         WHERE  p1.chemid <> p2.chemid),
    grouptypes
    AS (SELECT nhi,
                Max(grouptype) AS GroupType
         FROM    transitions
         GROUP  BY nhi),
    singlemed
    AS (SELECT nhi,
                CASE
                   WHEN Count(DISTINCT chemid) = 1 THEN Max(CASE
                   WHEN chemid = '3858' THEN 'F'
                   WHEN chemid = '3758' THEN 'B'

                                                        END)
                END AS GroupType
         FROM    prescriptionhistory
         GROUP  BY nhi
         HAVING Count(DISTINCT chemid) = 1
          -
- Checks that all records for the NHI are for the same medication
         )
-- Update the Cohort table with the determined group types
UPDATE dbo.cohort
SET    grouptype = COALESCE(g.grouptype, s.grouptype, 'Other')
```

```sql
-- Fallback to 'Other' if somehow no valid group is found
FROM    dbo.cohort c
        LEFT JOIN grouptypes g
              ON c.nhi = g.nhi
        LEFT JOIN singlemed s
              ON c.nhi = s.nhi;
```

SQL query 3 Severity steps table

```sql
USE jason_jli404;


-
- Drop the existing tables if they exist to prevent duplicate data
IF Object_id('dbo.PatientMedicationFlags', 'U') IS NOT NULL
  DROP TABLE dbo.patientmedicationflags;


IF Object_id('dbo.AsthmaSeveritySummary', 'U') IS NOT NULL
  DROP TABLE dbo.asthmaseveritysummary;


go


-- Create a new table for patient medication flags
CREATE TABLE dbo.patientmedicationflags
  (
     nhi                VARCHAR(20) NOT NULL,
     saba               BIT,
     ics                BIT,
     laba               BIT,
     ltra               BIT,
     biologics          BIT,
     asthmaseveritystep INT
  );


go


-- Insert the medication flags for each patient
```

```sql
INSERT INTO dbo.patientmedicationflags
            (nhi,
             saba,
             ics,
             laba,
             ltra,
             biologics,
             asthmaseveritystep)
SELECT nhi,
       Max(saba)      AS SABA,
       Max(ics)       AS ICS,
       Max(laba)      AS LABA,
       Max(ltra)      AS LTRA,
       Max(biologics) AS Biologics,
       Max(severity)  AS AsthmaSeverityStep
FROM   (SELECT nhi,
               CASE
                 WHEN chemname IN ( 'Salbutamol', 'Terbutaline sul
phate' ) THEN
                   1
                 ELSE 0
               END AS SABA,
               CASE
                 WHEN chemname IN ( 'Fluticasone furoate with vila
nterol',
                                    'Budesonide'
                                    ,
                                    'Fluticasone'
                                    ,
                                    'Budesonide with eformoterol',
                                    'Beclomethasone dipropionate',
                                    'Fluticasone with salmeterol'
) THEN 1
                   ELSE 0
                 END AS ICS,
               CASE
                 WHEN chemname IN ( 'Eformoterol fumarate', 'Indac
```

```
aterol',
                                'Eformoterol fumarate dihydrat
e',
                                                'Salmeterol
',
                                'Fluticasone with salmeterol',
                                    'Fluticasone furoate wi
th vilanterol'
                                ) THEN
            1
            ELSE 0
        END AS LABA,
        CASE
            WHEN chemname IN ( 'LTRA1', 'LTRA2' ) THEN 1
            ELSE 0
        END AS LTRA,
        CASE
            WHEN chemname IN ( 'Biologics', 'Biologics2' ) TH
EN 1
            ELSE 0
        END AS Biologics,
        CASE
            WHEN chemname IN ( 'Biologics1', 'Biologics2' ) T
HEN 4
            WHEN chemname IN ( 'Eformoterol fumarate', 'Indac
aterol',
                                'Eformoterol fumarate dihydrat
e',
                                                'Salmeterol
',
                                'Fluticasone with salmeterol',
                                    'Fluticasone furoate wi
th vilanterol'
                                ) THEN
            3
            WHEN chemname IN ( 'Fluticasone furoate with vila
nterol',
```

```sql
                                         'Budesonide'
                                         ,
                                         'Fluticasone'
                                         ,
                                         'Budesonide with eformoterol',
                                         'Beclomethasone dipropionate',
                                         'Fluticasone with salmeterol'
) THEN 2
                WHEN chemname IN ( 'Salbutamol', 'Terbutaline sul
phate' ) THEN
                1
                ELSE 0
            END AS Severity
     FROM    dbo.pharm
     WHERE   nhi IS NOT NULL) AS MedicationData
GROUP   BY nhi;


go


-- Create a new AsthmaSeveritySummary table
CREATE TABLE dbo.asthmaseveritysummary
  (
     asthmaseveritystep INT,
     saba               INT,
     ics                INT,
     laba               INT,
     ltra               INT,
     biologics          INT,
     numberofpatients   INT
  );


-
- Populate the AsthmaSeveritySummary table with the summarized fla
g data
INSERT INTO dbo.asthmaseveritysummary
            (asthmaseveritystep,
             saba,
```

```sql
            ics,
            laba,
            ltra,
            biologics,
            numberofpatients)
SELECT asthmaseveritystep,
       -
- Ensure that the flags are being treated as integers for the SUM
function
       Sum(Cast(saba AS INT))      AS SABA,
       Sum(Cast(ics AS INT))       AS ICS,
       Sum(Cast(laba AS INT))      AS LABA,
       Sum(Cast(ltra AS INT))      AS LTRA,
       Sum(Cast(biologics AS INT)) AS Biologics,
       Count(*)                    AS NumberOfPatients
FROM   (SELECT asthmaseveritystep,
               -
- indicating whether the patient is on each medication type
               saba,
               ics,
               laba,
               ltra,
               biologics
        FROM   dbo.patientmedicationflags) AS PatientFlags
GROUP  BY asthmaseveritystep,
          saba,
          ics,
          laba,
          ltra,
          biologics
ORDER  BY asthmaseveritystep,
          saba DESC,
          ics DESC,
          laba DESC,
          ltra DESC,
          biologics DESC;
```

Appendix

```sql
-- view the results
SELECT *
FROM    dbo.asthmaseveritysummary
ORDER  BY asthmaseveritystep ASC;
```

SQL Query 4 Dosage formation

```sql
USE jason_jli404

go

IF Object_id('tempdb..#Dim') IS NOT NULL
  DROP TABLE #dim;

IF Object_id('tempdb..#Years') IS NOT NULL
  DROP TABLE #years;

IF Object_id('tempdb..#PatientMedicationYears') IS NOT NULL
  DROP TABLE #patientmedicationyears;

IF Object_id('tempdb..#PatientYearlyDosage') IS NOT NULL
  DROP TABLE #patientyearlydosage;

IF Object_id('tempdb..#PatientDosage') IS NOT NULL
  DROP TABLE #patientdosage;

IF Object_id('dbo.PatientDosage') IS NOT NULL
  DROP TABLE dbo.patientdosage;

go

-- Create a list of all years for the dataset
SELECT DISTINCT Year(datedispensed) AS YearDispensed
INTO    #years
FROM    jason_jli404.dbo.pharm;
```

Appendix

```sql
go

SELECT c.nhi,
       c.chemname,
       y.yeardispensed
INTO   #patientmedicationyears
FROM   (SELECT DISTINCT nhi,
                        chemname,
                        indexdate,
                        exitdate
        FROM   jason_jli404.dbo.cohort) c
       JOIN #years y
         ON y.yeardispensed BETWEEN Year(c.indexdate) AND Year(Dat
eadd(year, -1,
                                                            c.exitda
te));


go


---- Generate all combinations of NHI, chemName, and Year
SELECT pm.nhi,
       pm.chemname,
       pm.yeardispensed,
       Avg(CONVERT(FLOAT, p.age))         AS Age,
       p.gender                           AS Gender,
       Isnull(Sum(p.quantitydispensed), 0) AS TotalDosage
INTO   #patientyearlydosage
FROM   #patientmedicationyears pm
       LEFT JOIN jason_jli404.dbo.pharm p
              ON pm.nhi = p.nhi
                 AND Replace(pm.chemname, ' ', '') =
                     Replace(p.chemname, ' ', '')
                 AND pm.yeardispensed = Year(p.datedispensed)
GROUP  BY pm.nhi,
          pm.chemname,
          pm.yeardispensed,
          p.gender;
```

Appendix

```sql
go


-
- Generate the dynamic pivot query to create the final PatientDosa
ge table
DECLARE @DynamicPivotQuery AS NVARCHAR(max);
DECLARE @ColumnName AS NVARCHAR(max) = '';
DECLARE @SelectList AS NVARCHAR(max) = '';


-- Prepare column names for the pivot table
SELECT @ColumnName += CASE WHEN @ColumnName = '' THEN '' ELSE ','
END +
                    Quotename(
                            Replace(chemname,
                            ' ', '_'))
FROM    (SELECT DISTINCT chemname
        FROM    #patientyearlydosage) AS ChemNames;


-
- Create select list for the pivot to replace NULL with 0 in dosag
e columns
SELECT @SelectList += CASE WHEN @SelectList = '' THEN '' ELSE ','
END +
                    'ISNULL('
                    + Quotename(Replace(chemname, ' ', '_'))
                    + ', 0) AS '
                    + Quotename(Replace(chemname, ' ', '_'))
FROM    (SELECT DISTINCT chemname
        FROM    #patientyearlydosage) AS ChemNames;


-- Construct the dynamic PIVOT query
SET @DynamicPivotQuery = 'SELECT NHI, YearDispensed, Age, Gender,
'
                        + @SelectList
                        +
' INTO dbo.PatientDosage    FROM (SELECT NHI, REPLACE(chemName, ''
```

```sql
 '', ''_'') AS chemName, Age, Gender, TotalDosage, YearDispensed F
ROM #PatientYearlyDosage) x '
                        + ' PIVOT (SUM(TotalDosage) FOR chemName
IN ('
                        + @ColumnName + ')) AS PivotTable ';


-- Execute the dynamic PIVOT query
EXEC Sp_executesql
   @DynamicPivotQuery;


go


-- Check if the table was created and contains data
IF Object_id('dbo.PatientDosage') IS NOT NULL
   PRINT 'PatientDosage successful!';
ELSE
   PRINT 'Failed to create PatientDosage';
--SELECT * FROM dbo.PatientDosage;
```

```sql
USE jason_jli404

go

IF Object_id('dbo.PatientDeath') IS NOT NULL
   DROP TABLE dbo.patientdeath;

go

SELECT cohort.nhi,
       Year(pharm.datedispensed) AS YearDispensed,
       CASE
         WHEN Year(pharm.datedispensed) = Year(cohort.exitdate) -
1 THEN 1
         ELSE 0
       END                            AS DeathFlag
INTO   patientdeath
```

```sql
FROM    jason_jli404.dbo.pharm pharm
        JOIN jason_jli404.dbo.cohort cohort
          ON pharm.nhi = cohort.nhi
WHERE   pharm.datedispensed BETWEEN cohort.indexdate AND cohort.exi
tdate
GROUP   BY cohort.nhi,
           Year(pharm.datedispensed),
           Year(cohort.exitdate)
ORDER   BY cohort.nhi,
           yeardispensed;
```

```sql
USE jason_jli404;

go

-- Ensure no conflicts with existing objects
IF Object_id('dbo.PatientSummary') IS NOT NULL
  DROP TABLE dbo.patientsummary;

go

-
- Combine PatientDosage and PatientDeath into one table with condi
tions based on death flag
SELECT dosage.nhi                             AS NHI,
       dosage.yeardispensed                   AS Year,
       dosage.age                             AS Age,
       dosage.gender                          AS Gender,
       Isnull(death.deathflag, 0)             AS DeathFlag,
       dosage.eformoterol_fumarate            AS Eformoterol_f
umarate,
       dosage.fluticasone_furoate_with_vilanterol AS
       Fluticasone_furoate_with_vilanterol,
       dosage.salbutamol                      AS Salbutamol,
       dosage.indacaterol                     AS Indacaterol,
       dosage.fluticasone                     AS Fluticasone,
```

```sql
        dosage.budesonide                          AS Budesonide,
        dosage.eformoterol_fumarate_dihydrate      AS
        Eformoterol_fumarate_dihydrate,
        dosage.dexamethasone                       AS Dexamethasone
,
        dosage.salmeterol                          AS Salmeterol,
        dosage.prednisone                          AS Prednisone,
        dosage.prednisolone                        AS Prednisolone,
        dosage.budesonide_with_eformoterol         AS Budesonide_wi
th_eformoterol
        ,
        dosage.terbutaline_sulphate                AS
        Terbutaline_sulphate,
        dosage.beclomethasone_dipropionate         AS Beclomethason
e_dipropionate
        ,
        dosage.fluticasone_with_salmeterol         AS
        Fluticasone_with_salmeterol,
        dosage.methylprednisolone                  AS Methylprednis
olone
INTO    dbo.patientsummary
FROM    patientdosage dosage
        LEFT JOIN dbo.patientdeath death
              ON dosage.nhi = death.nhi
                 AND dosage.yeardispensed = death.yeardispensed
WHERE   NOT EXISTS (
                  -- Exclude records where a death flag exists
                  SELECT 1
                   FROM    dbo.patientdeath deathCheck
                   WHERE   deathCheck.nhi = dosage.nhi
                          AND deathCheck.yeardispensed < dosage.ye
ardispensed
                          AND deathCheck.deathflag = 1)
ORDER   BY dosage.nhi,
        dosage.yeardispensed;


go
```

Appendix

```sql
SELECT TOP 500 *
FROM    dbo.patientsummary
ORDER  BY nhi,
          year ASC;


go
```

Code 1

```python
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from scipy.stats import norm, invgauss, weibull_min

from sqlalchemy import create_engine


# Define connection string for Windows Authentication

connection_str =

"mssql+pyodbc://FMHSPHARMDBPRD1/Jason_jli404?driver=ODBC+Driver+17+for+SQL+Server&trusted_connection=yes"


# Create an engine

engine = create_engine(connection_str)


query = "select * from dbo.Cohort where (gender NOT like 'U' AND gender not like 'O' AND studyDiff >= 0 AND age > 0)"

df = pd.read_sql(query, engine)


meds_of_interest = ['Fluticasone with salmeterol', 'Budesonide with eformoterol']

df_filtered = df[df['chemName'].isin(meds_of_interest)]


# Separate the data for each medication

fluticasone_salmeterol_data = df_filtered[
```

```
    df_filtered["chemName"] == "Fluticasone with salmeterol"
]["age"]
budesonide_eformoterol_data = df_filtered[
    df_filtered["chemName"] == "Budesonide with eformoterol"
]["age"]


# Function to estimate parameters and plot for a given dataset and medication name
def estimate_and_plot(data, medication_name):
    # Estimate parameters for Normal, Wald (Inverse Gaussian), and Weibull distributions
    norm_params = norm.fit(data)
    wald_params = invgauss.fit(data)
    weibull_params = weibull_min.fit(data)


    # Plotting
    plt.figure(figsize=(10, 6))
    plt.hist(data, bins=30, alpha=0.6, color="g", density=True, label="Observed Data")


    # Range for plotting distributions
    xmin, xmax = plt.xlim()
    x = np.linspace(xmin, xmax, 100)


    # Normal Distribution
    p_norm = norm.pdf(x, *norm_params)
    plt.plot(x, p_norm, "k", linewidth=2, label="Normal Distribution")


    # Wald Distribution
    p_wald = invgauss.pdf(x, *wald_params)
    plt.plot(x, p_wald, "r", linewidth=2, label="Wald (Inverse Gaussian) Distribution")


    # Weibull Distribution
    p_weibull = weibull_min.pdf(x, *weibull_params)
    plt.plot(x, p_weibull, "b", linewidth=2, label="Weibull Distribution")


    plt.title(f"Age Distribution with Estimated Parameters for {medication_name}")
```

```
    plt.legend()

    plt.show()


    # Print estimated parameters

    print(f"Estimated parameters for {medication_name}:")

    print(f"Normal: {norm_params}")

    print(f"Wald (Inverse Gaussian): {wald_params}")

    print(f"Weibull: {weibull_params}")



# Estimate parameters and plot for Fluticasone with salmeterol

estimate_and_plot(fluticasone_salmeterol_data, "Fluticasone with Salmeterol")


# Estimate parameters and plot for Budesonide with eformoterol

estimate_and_plot(budesonide_eformoterol_data, "Budesonide with Eformoterol")
```

Code 2

```
from sqlalchemy import create_engine

from scipy.stats import norm, wald, weibull_min

import pymc as pm

import arviz as az

import pandas as pd

import numpy as np


# Define connection string for Windows Authentication

connection_str =

"mssql+pyodbc://FMHSPHARMDBPRD1/Jason_jli404?driver=ODBC+Driver+17+for+SQL+Server&tr

usted_connection=yes"


# Create an engine

engine = create_engine(connection_str)
```

Appendix

```
query = "select * from dbo.Cohort where (gender NOT like 'U' AND gender not like 'O' AND
studyDiff >= 0 AND age > 0)"
df = pd.read_sql(query, engine)


meds_of_interest = ['Fluticasone with salmeterol', 'Budesonide with eformoterol']


# Filter to include only deceased patients
deceased_patients = df[(df['Flag'] == 1) & df['chemName'].isin(meds_of_interest)]


group_types = deceased_patients['GroupType'].unique()


import pymc as pm
import arviz as az
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt


# Data preparation
group_types = deceased_patients['GroupType'].unique()
results = {}
trace_data = {}


# Modeling for each group type
for group in group_types:
    data = deceased_patients[deceased_patients['GroupType'] == group]['age'].values
    if len(data) > 0:
        with pm.Model() as model:
            # Define priors for Weibull parameters
            alpha = pm.HalfNormal(f'alpha_{group}', sigma=5)
            beta = pm.HalfNormal(f'beta_{group}', sigma=10)

            # Weibull likelihood
            Y_obs = pm.Weibull(f'Y_obs_{group}', alpha=alpha, beta=beta, observed=data)
```

Appendix

```python
        # Posterior sampling
        trace = pm.sample(1000, tune=500, return_inferencedata=True)


    results[group] = {'model': model, 'trace': trace}
    trace_data[group] = {'alpha': trace.posterior[f'alpha_{group}'].values.flatten(),
                'beta': trace.posterior[f'beta_{group}'].values.flatten()}


    # Plot trace and save with title as filename
    trace_plot = az.plot_trace(trace)
    plot_title = f'Trace Plot for {group} Group'
    plt.suptitle(plot_title)
    plt.savefig(f'C:\\Users\\jli404\\Desktop\\{plot_title}.png')
    plt.show()


    summary = az.summary(trace, var_names=[f'alpha_{group}', f'beta_{group}'])
    print(f"{group} Group - Weibull Parameters")
    print(summary)
  else:
    print(f"No data available for group {group}")


# Compare alpha and beta parameters across groups and save summaries
for group1 in group_types:
  for group2 in group_types:
    if group1 != group2:
      alpha_diff = trace_data[group1]['alpha'] - trace_data[group2]['alpha']
      beta_diff = trace_data[group1]['beta'] - trace_data[group2]['beta']


      # Summarize differences
      alpha_diff_summary = {
        'mean': np.mean(alpha_diff),
        'sd': np.std(alpha_diff),
        'hdi_3%': np.quantile(alpha_diff, 0.03),
        'hdi_97%': np.quantile(alpha_diff, 0.97),
      }
```

```
        beta_diff_summary = {

          'mean': np.mean(beta_diff),

          'sd': np.std(beta_diff),

          'hdi_3%': np.quantile(beta_diff, 0.03),

          'hdi_97%': np.quantile(beta_diff, 0.97),

        }


        summary_df = pd.DataFrame({

          'Parameter Difference': [f'alpha ({group1} - {group2})', f'beta ({group1} - {group2})'],

          'Mean': [alpha_diff_summary['mean'], beta_diff_summary['mean']],

          'SD': [alpha_diff_summary['sd'], beta_diff_summary['sd']],

          'HDI 3%': [alpha_diff_summary['hdi_3%'], beta_diff_summary['hdi_3%']],

          'HDI 97%': [alpha_diff_summary['hdi_97%'], beta_diff_summary['hdi_97%']],

        })


        # Save summary DataFrame to CSV

        summary_filename = f'summary_{group1}_vs_{group2}.csv'

        summary_df.to_csv(f'C:\\Users\\jli404\\Desktop\\{summary_filename}', index=False)


        print(summary_df)
```

Code 3 Function to calculate mean and standard deviation for Weibull distribution

```
# Function to calculate mean and std for Weibull distribution
def weibull_stats(shape, scale):
    mean = scale * gamma(1 + 1/shape)
    variance = (scale**2) * (gamma(1 + 2/shape) - (gamma(1 + 1/shape))**2)
    std = np.sqrt(variance)
    return mean, std
```

Appendix

Package Code

anyio @ file:///C:/ci/anyio_1644481921011/work/dist

argon2-cffi @ file:///opt/conda/conda-bld/argon2-cffi_1645000214183/work

argon2-cffi-bindings @ file:///C:/ci/argon2-cffi-bindings_1644551690056/work

arviz @ file:///home/conda/feedstock_root/build_artifacts/arviz_1666645025910/work

asttokens @ file:///opt/conda/conda-bld/asttokens_1646925590279/work

async-lru @ file:///C:/b/abs_e0hjkvwwb5/croot/async-lru_1699554572212/work

attrs @ file:///C:/b/abs_35n0jusce8/croot/attrs_1695717880170/work

Babel @ file:///C:/b/abs_a2shv_3tqi/croot/babel_1671782804377/work

backcall @ file:///home/ktietz/src/ci/backcall_1611930011877/work

beautifulsoup4 @ file:///C:/b/abs_0agyz1wsr4/croot/beautifulsoup4-split_1681493048687/work

bleach @ file:///opt/conda/conda-bld/bleach_1641577558959/work

Brotli @ file:///C:/Windows/Temp/abs_63l7912z0e/croots/recipe/brotli-
split_1659616056886/work

cachetools==5.3.3

certifi @ file:///C:/b/abs_91u83siphd/croot/certifi_1700501720658/work/certifi

cffi @ file:///C:/b/abs_924gv1kxzj/croot/cffi_1700254355075/work

cftime @ file:///D:/bld/cftime_1666833911368/work

charset-normalizer @ file:///tmp/build/80754af9/charset-normalizer_1630003229654/work

cloudpickle==3.0.0

colorama @ file:///C:/b/abs_a9ozq0l032/croot/colorama_1672387194846/work

comm @ file:///C:/b/abs_1419earm7u/croot/comm_1671231131638/work

cons==0.4.6

contourpy==1.2.0

cryptography @ file:///C:/b/abs_f4do8t8jfs/croot/cryptography_1694444424531/work

cycler==0.12.1

debugpy @ file:///C:/b/abs_c0y1fjipt2/croot/debugpy_1690906864587/work

decorator @ file:///opt/conda/conda-bld/decorator_1643638310831/work

defusedxml @ file:///tmp/build/80754af9/defusedxml_1615228127516/work

deprecat @ file:///home/conda/feedstock_root/build_artifacts/deprecat_1653044502293/work

dill @ file:///home/conda/feedstock_root/build_artifacts/dill_1690101045195/work

docopt==0.6.2

etuples==0.3.9

Appendix

```
exceptiongroup @ file:///C:/b/abs_25wqfvkf25/croot/exceptiongroup_1668714345637/work

executing @ file:///opt/conda/conda-bld/executing_1646925071911/work

fastjsonschema @

file:///C:/Users/BUILDE~1/AppData/Local/Temp/abs_ebruxzvd08/croots/recipe/python-

fastjsonschema_1661376484940/work

fastprogress==1.0.3

filelock==3.13.1

fonttools==4.50.0

greenlet==3.0.1

h5py==3.10.0

idna @ file:///C:/b/abs_bdhbebrioa/croot/idna_1666125572046/work

importlib-metadata @ file:///home/conda/feedstock_root/build_artifacts/importlib-

metadata_1701632192416/work

importlib_resources==6.3.1

ipykernel @ file:///C:/b/abs_07rkft_vaz/croot/ipykernel_1691121700587/work

ipython==8.12.3

ipywidgets @ file:///C:/b/abs_5awapknmz_/croot/ipywidgets_1679394824767/work

jedi @ file:///C:/ci/jedi_1644315428289/work

Jinja2 @ file:///C:/b/abs_7cdis66kl9/croot/jinja2_1666908141852/work

joblib==1.3.2

json5 @ file:///tmp/build/80754af9/json5_1624432770122/work

jsonschema @ file:///C:/b/abs_d1c4sm8drk/croot/jsonschema_1699041668863/work

jsonschema-specifications @ file:///C:/b/abs_0brvm6vryw/croot/jsonschema-

specifications_1699032417323/work

jupyter @ file:///C:/Windows/TEMP/abs_56xfdi__li/croots/recipe/jupyter_1659349053177/work

jupyter-console @ file:///C:/b/abs_82xaa6i2y4/croot/jupyter_console_1680000189372/work

jupyter-events @ file:///C:/b/abs_17ajfqnlz0/croot/jupyter_events_1699282519713/work

jupyter-lsp @ file:///C:/b/abs_ecle3em9d4/croot/jupyter-lsp-meta_1699978291372/work

jupyter_client @ file:///C:/b/abs_a6h3c8hfdq/croot/jupyter_client_1699455939372/work

jupyter_core @ file:///C:/b/abs_c769pbqg9b/croot/jupyter_core_1698937367513/work

jupyter_server @ file:///C:/b/abs_7esjvdakg9/croot/jupyter_server_1699466495151/work

jupyter_server_terminals @

file:///C:/b/abs_ec0dq4b50j/croot/jupyter_server_terminals_1686870763512/work

jupyterlab @ file:///C:/b/abs_aergn8zopq/croot/jupyterlab_1700518316761/work
```

Appendix

jupyterlab-pygments @ file:///tmp/build/80754af9/jupyterlab_pygments_1601490720602/work

jupyterlab-widgets @
file:///C:/b/abs_adrrqr26no/croot/jupyterlab_widgets_1700169018974/work

jupyterlab_server @ file:///C:/b/abs_e08i7qn9m8/croot/jupyterlab_server_1699555481806/work

kiwisolver==1.4.5

logical-unification==0.4.6

Mako @ file:///home/conda/feedstock_root/build_artifacts/mako_1699482234420/work

MarkupSafe @ file:///D:/bld/markupsafe_1695367558436/work

matplotlib==3.8.3

matplotlib-inline @ file:///C:/ci/matplotlib-inline_1661915841596/work

miniKanren==1.0.3

mistune @
file:///C:/Users/BUILDE~1/AppData/Local/Temp/abs_081kimkskf/croots/recipe/mistune_166149
6225923/work

multipledispatch==1.0.0

munkres==1.1.4

nbclient @ file:///C:/b/abs_cal0q5fyju/croot/nbclient_1698934263135/work

nbconvert==7.16.4

nbformat @ file:///C:/b/abs_5a2nea1iu2/croot/nbformat_1694616866197/work

nest-asyncio @ file:///C:/b/abs_3a_4jsjlqu/croot/nest-asyncio_1672387322800/work

netCDF4 @ file:///D:/bld/netcdf4_1687961316745/work

notebook @ file:///C:/b/abs_26737osg4x/croot/notebook_1700582146311/work

notebook_shim @ file:///C:/b/abs_a5xysln3lb/croot/notebook-shim_1699455926920/work

numpy==1.26.4

overrides @ file:///C:/b/abs_cfh89c8yf4/croot/overrides_1699371165349/work

packaging==24.0

pandas==2.2.1

pandocfilters @ file:///opt/conda/conda-bld/pandocfilters_1643405455980/work

parso @ file:///opt/conda/conda-bld/parso_1641458642106/work

patsy @ file:///home/conda/feedstock_root/build_artifacts/patsy_1701443970942/work

pickleshare @ file:///tmp/build/80754af9/pickleshare_1606932040724/work

pillow==10.2.0

pipreqs==0.5.0

platformdirs @ file:///C:/b/abs_b6z_yqw_ii/croot/platformdirs_1692205479426/work

Appendix

ply==3.11

prometheus-client @

file:///C:/Windows/TEMP/abs_ab9nx8qb08/croots/recipe/prometheus_client_1659455104602/w

ork

prompt-toolkit @ file:///C:/b/abs_6coz5_9f2s/croot/prompt-toolkit_1672387908312/work

psutil @ file:///C:/Windows/Temp/abs_b2c2fd7f-9fd5-4756-95ea-

8aed74d0039flsd9qufz/croots/recipe/psutil_1656431277748/work

pure-eval @ file:///opt/conda/conda-bld/pure_eval_1646925070566/work

pycparser @ file:///tmp/build/80754af9/pycparser_1636541352034/work

Pygments @ file:///C:/b/abs_fay9dpq4n_/croot/pygments_1684279990574/work

pygpu==0.7.6

pymc==5.11.0

pyodbc==5.0.1

pyOpenSSL @ file:///C:/b/abs_08f38zyck4/croot/pyopenssl_1690225407403/work

pyparsing==3.1.2

PyQt5==5.15.10

PyQt5-sip @ file:///C:/b/abs_c0pi2mimq3/croot/pyqt-split_1698769125270/work/pyqt_sip

PySocks @ file:///C:/ci/pysocks_1605307512533/work

pytensor==2.18.6

python-dateutil==2.9.0.post0

python-json-logger @ file:///C:/b/abs_cblnsm6puj/croot/python-json-

logger_1683824130469/work

pytz==2024.1

pywin32==305.1

pywinpty @

file:///C:/b/abs_73vshmevwq/croot/pywinpty_1677609966356/work/target/wheels/pywinpty-

2.0.10-cp39-none-win_amd64.whl

PyYAML @ file:///C:/b/abs_782o3mbw7z/croot/pyyaml_1698096085010/work

pyzmq @ file:///C:/b/abs_655zk4a3s8/croot/pyzmq_1686601465034/work

qtconsole @ file:///C:/b/abs_4awqjtg1ug/croot/qtconsole_1700160696631/work

QtPy @ file:///C:/b/abs_derqu__3p8/croot/qtpy_1700144907661/work

referencing @ file:///C:/b/abs_09f4hj6adf/croot/referencing_1699012097448/work

requests @ file:///C:/b/abs_316c2inijk/croot/requests_1690400295842/work

rfc3339-validator @ file:///C:/b/abs_ddfmseb_vm/croot/rfc3339-validator_1683077054906/work

Appendix

```
rfc3986-validator @ file:///C:/b/abs_6e9azihr8o/croot/rfc3986-validator_1683059049737/work

rpds-py @ file:///C:/b/abs_76j4g4la23/croot/rpds-py_1698947348047/work

scikit-learn==1.4.1.post1

scipy==1.12.0

seaborn==0.13.2

semver @ file:///home/conda/feedstock_root/build_artifacts/semver_1696861993140/work

Send2Trash @ file:///C:/b/abs_08dh49ew26/croot/send2trash_1699371173324/work

sip @ file:///C:/b/abs_edevan3fce/croot/sip_1698675983372/work

six==1.16.0

sniffio @ file:///C:/ci/sniffio_1614030527509/work

soupsieve @ file:///C:/b/abs_bbsvy9t4pl/croot/soupsieve_1696347611357/work

SQLAlchemy==2.0.23

stack-data @ file:///opt/conda/conda-bld/stack_data_1646927590127/work

terminado @ file:///C:/b/abs_25nakickad/croot/terminado_1671751845491/work

Theano-PyMC @ file:///D:/bld/theano-pymc_1611363584953/work

threadpoolctl==3.3.0

tinycss2 @ file:///C:/b/abs_52w5vfuaax/croot/tinycss2_1668168823131/work

tomli @ file:///C:/Windows/TEMP/abs_ac109f85-a7b3-4b4d-bcfd-
52622eceddf0hy332ojo/croots/recipe/tomli_1657175513137/work

toolz==0.12.1

tornado @ file:///C:/b/abs_0cbrstidzg/croot/tornado_1696937003724/work

traitlets @ file:///C:/b/abs_e5m_xjjl94/croot/traitlets_1671143896266/work

typing_extensions==4.10.0

tzdata==2024.1

unicodedata2 @ file:///D:/bld/unicodedata2_1695847967701/work

urllib3 @ file:///C:/b/abs_9cmlsrm3ys/croot/urllib3_1698257595508/work

wcwidth @ file:///Users/ktietz/demo/mc3/conda-bld/wcwidth_1629357192024/work

webencodings==0.5.1

websocket-client @ file:///C:/ci/websocket-client_1614804375980/work

widgetsnbextension @
file:///C:/b/abs_882k4_4kdf/croot/widgetsnbextension_1679313880295/work

win-inet-pton @ file:///C:/ci/win_inet_pton_1605306162074/work

wrapt @ file:///D:/bld/wrapt_1699532935905/work

xarray @ file:///home/conda/feedstock_root/build_artifacts/xarray_1689599939832/work
```

Appendix

xarray-einstats @ file:///home/conda/feedstock_root/build_artifacts/xarray-

einstats_1689089835984/work

yarg==0.1.9

zipp==3.18.1