## 3. Literature Review

A thorough literature review revealed three methods that used data modelling techniques to estimate global maternal mortality ratios. Section 3.1 discusses these approaches in detail. My proposed models’ MMR estimates will compared against estimates from these existing models later in this thesis. Section 3.2 introduces an alternative algorithmic modelling approach, with Section 3.3 providing examples of this approach in public and maternal health research.

### 3.1 Existing Methods for Estimating the MMR Using Data Modelling

#### 3.11 United Nations Maternal Mortality Estimation Inter-Agency Group Estimates

The Maternal Mortality Estimation Inter-Agency Group (MMEIG) is a collaboration between United Nations (UN) Member States, the WHO, the World Bank Group, and various UN agencies. This collaboration models regional and country-specific MMRs between 2000 and 2023 for women 15 to 49 years old in 195 countries and territories [1]. The consistent methodology used to determine these estimates enables global monitoring of maternal mortality trends and progress towards the UN’s Sustainable Development Goal. The group’s modelling approach has evolved over this time frame, with each report updating the model used to estimate the MMR [1].

The MMEIG uses a combination of two models to estimate MMR [1,3]. First, the Bayesian maternal mortality misclassification (BMis) model calculates adjustment factors for the provided civil registration and vital statistics (CRVS) data to account for underreporting and misclassification of maternal deaths, as discussed in the Background Information [1,3]. This adjustment was applied specifically to CRVS data because it was the largest input to the MMEIG estimates [3]. Global and country-specific adjustment factors were calculated using high-quality specialised studies, which provided a benchmark for the accuracy of CRVS data [3]. Adjustment factors for countries with no specialised studies were calculated using global estimates [3]. Data from non-CRVS or specialised study sources was increased by 10% to account for underestimation of maternal mortality [3]. This explicit adjustment for poor-quality data was a strength of the UN MMEIG model, especially given the widespread underreporting and misclassification of maternal deaths reported in the literature [cite].

After using the BMis model to correct CRVS data errors, the MMEIG estimated the MMR per country per year using the Bayesian maternal mortality estimation (BMat) model [1]. The BMat calculates MMR as the sum of non-HIV-related maternal deaths and HIV-related maternal deaths, where death was due to pregnancy aggravating an existing HIV/AIDs condition [1]. The BMat estimates non-HIV-related MMR using a Bayesian hierarchical regression model [3]. Briefly, a hierarchical model determines a general trend and individual-specific deviations from the trend, which are referred to as random effects [15]. In BMat model, non-HIV related maternal deaths were estimated using global covariate trends with region and country specific effects, reflecting the belief that countries in the same geographic region have similar MMR trends [2]. This enables the model to estimate MMR for countries with sparse data, as it can use information from countries in a similar geographic region, which is an advantage of the modelling approach. The model parameterises the general trend and individual effects using prior knowledge, with each of the parameters initially drawn from a prior probability distribution [15]. The parameters were updated via Bayes’ rule upon observation of data, with less data causing the final parameter values to be closer to the parameters drawn from the prior distribution [15].

The covariates used to calculate non-HIV MMR were gross-domestic product (GDP) per capita, general fertility rate, and presence of a skilled birth attendant [1]. An autoregressive integrated moving average (ARIMA) model was then used to determine whether empirical country-maternal mortality tracked with the covariates [3]. For example, if the data indicated that non-HIV MMR decreased more slowly than predicted by the covariates, the non-HIV MMR estimate would be reduced [1]. This incorporation of a data-driven factor is a strength of the BMat model, as it enables the model to reflect local conditions and be more responsive to sudden shocks that cause the country’s MMR to deviate from the trend predicted by the covariates [2017 paper]. The less data available for a specific country-year, the less the associated non-HIV MMR estimate could be adjusted, making it more strongly influenced by the covariates [1].

HIV-related MMR was estimated separately because evidence indicates that HIV/AIDS is a prominent cause of maternal mortality in countries with ongoing HIV/AID epidemics, with studies showing that women infected with HIV have approximately 8 times higher risk of pregnancy-related death [1, 19]. Calculation of HIV-related MMR involves a constant that defines the relative risk of dying from HIV/AIDS for a pregnant versus non-pregnant women, which is estimated in conjunction with experts [1,2]. The small subjectivity in the value of this constant is demonstrated through the change in its value between the 2012 and most recent BMat models [18, 2].

While the model was fit to all data provided by the country, it placed higher weight on values with lower error variances, which were derived from calculating the random error in the data collection processes. As a result of incorporating error, the final BMat estimates had smaller uncertainty intervals for countries with higher-quality data [1]. By incorporating these uncertainty intervals, the BMat model could warn stakeholders about which statistics are more reliable, thus providing more nuanced information about maternal health trends.

#### 3.12 The Global Burden of Disease Estimates

Coordinated by the Institute of Health Metrics and Evaluation (IHME), the Global Burden of Disease Study (GBD) is an international scientific initiative that has benchmarked major diseases and risk factors since 1993 [5]. It encourages scientific debate by providing an alternate set of population health estimates from the UN [22]. Additionally, unlike the UN MMEIG estimates which are produced under consultation with Member States, GBD estimates are produced independently, removing potential bias [21]. Many GBD Studies have been published in high-impact, peer-reviewed journals like *Lancet* [5]. The GBD Study’s strong reputation is apparent in its use in national planning by a variety of governments, such as the United Kingdom, Norway, and China [5].

The 2021 GBD Study produced estimates of maternal mortality for 204 countries between 1990 and 2021 [6]. Similar to the UN MMEIG estimates, using a consistent methodological approach across countries enables GBD estimates to be used to monitor global trends in MMR and track progress toward the Sustainable Development Goals [gbd at 30 years]. Unlike the UN MMEIG estimates, the GBD estimates were informed by sub-national data as well as national data, allowing it to incorporate local realities into its predictions [6]. The data was then cleaned, standardised, and any deaths reported with an unclear or incorrect cause-of-death were probabilistically redistributed to a more likely cause of death [8]. This reduced the effect of misclassification errors on the model’s final outputs, improving the accuracy of its predictions.

GBD maternal mortality estimates were produced using cause of death ensemble modelling (CODEm) [4]. The ensemble consisted of linear mixed effects regression (LMER) and spatiotemporal Gaussian process regression (ST-GPR) models [4]. Like the Bayesian hierarchical models discussed above, LMER models use random effects to quantify super-region, region, country, and age-level trends [7]. The ST-GPR model first uses a LMER model without country-specific effects to give a basic maternal mortality prediction [7]. Then, it performs additional regression on the initial prediction errors using smoothing, where each datapoint’s individual error was replaced by the weighted average of errors in its neighbourhood [7]. The neighbourhood was defined over time, geographic space, and age [7]. The smoothing operation reduces noise, with the extent of smoothing increasing for countries with less data [7].

This technique is a strength of the GBD Study, as it enables CODEm to estimate the MMR of countries with sparse data using information from similar regions. Additionally, by adding the computed residuals to the model’s initial predictions, the model can capture trends in the data not represented by the covariates [7].

The relationship between all chosen covariates and maternal mortality was statistically significant [4]. Additionally, each covariate had a causal link with maternal mortality, as established by existing scientific literature and expert analysis [4]. The GBD 2021 Study used more covariates than the MMEIG, with the former using 19 covariates while the latter only used three [4, 1]. More specifically, the GBD study estimated maternal mortality using covariates including, but not limited to, age-specific fertility, maternal education, neonatal mortality ratio, skilled birth attendance, age-specific HIV mortality in females 10 to 54 years old, and age-standardised wasting [4].

The final maternal mortality prediction was the mean of 1000 CODEm ensemble predictions, with each prediction generated by one individual, component model. The likelihood of each model being chosen was determined by its weight, which was calculated from the base model’s out-of-sample predictive performance [4]. The 1000 draws allowed the construction of a 95% uncertainty interval [4]. This approach was another strength of the CODEm model, as it leverages patterns learned by a variety of models and increases generalisation, as discussed in the Background Information [cite].

#### 3.13 Limitations of the UN MMEIG and GBD Study’s MMR Estimates

While the UN MMEIG’s BMat model and the GBD Study’s CODEm model are widely respected, and their estimates are used to inform policy, they have some limitations [un mmeig report and GBD at 30 years].

The MMR estimates produced by the BMat and CODEm models were based on sparse, low-quality data [2,3]. Inaccurate data can cause the models to produce misleading MMR estimates, reducing their ability to inform policy. Unfortunately, limited data about the extent of underreporting and misclassification errors in different countries and systems make the errors difficult to correct, reducing the efficacy of the UN’s BMis model and the GBD Study’s data cleaning mechanisms [3]. This was especially true in countries without CRVS systems, which have lower data quality and quantity [13]. As a result, MMR estimates for countries with less developed data collection systems tend to be predicted with wide uncertainty bounds, reducing their informativeness [1,3,13].

Similarly, lack of data about the trends captured by the models’ parameters can force researchers to use wide prior distributions for the parameters [15]. When there is little empirical data available to update the parameters’ values, the parameters are largely informed by this wide prior distribution [15]. This can result in incorrect results with large uncertainty, again reducing the utility of the models’ outputs.

Additionally, when there is little to no data for a country, BMat would estimate non-HIV related MMR solely using regional effects and covariates [2]. Similarly, the CODEm LMER models would estimate MMR using the covariates, super-region and region effects [7]. CODEm ST-GPR models would increase smoothing in their residual regression analysis and use the region and super-region effects [7]. While this enables the models to estimate MMR for data-sparse countries, if there is regional heterogeneity, the country’s MMR estimates would be pulled towards an unrepresentative region-level estimate, resulting in inaccurate country-level predictions. Additionally, the lack of data may be due to an abrupt change or crisis, which smoothing may obscure, causing the model to lose important information about maternal mortality.

Prediction error may also be due to the assumption of a global relationship between the covariates and MMR. However, research has found that skilled birth attendance (SBA), a covariate used in both BMat and CODEm, only significantly reduces MMR when SBA coverage across the country is at least 40% [16]. Thus, countries with very low SBA coverage have a different relationship to MMR than countries with high coverage. If these countries are also missing data, the UN ARIMA model or GBD ST-GPR models could not adjust the covariate-driven estimates, contributing to inaccurate model predictions.

Another major limitation of BMat and CODEm is their consideration of only a small subset of relevant covariates [18]. There are a wide variety of factors that influence maternal health. For example, non-communicable disease (NCDs) are a leading cause of maternal mortality, with cardiovascular disease being one of the primary causes of maternal mortality between 2018 and 2020 in Australia [27, 28]. Other NCDs like diabetes, asthma and mental health conditions also commonly affect pregnant women, with anemia increasing probability of postpartum hemorrhage, the primary cause of global maternal mortality [27]. The literature also gives evidence for how excess maternal mortality is linked to quality of medical care, the incidence of infectious diseases like malaria, climate-related hazards, availability of contraception, financial constraints, violence in the woman’s region, the woman’s geographic remoteness and education, racism in the health system, and gender inequities that influence a woman’s ability to make decisions about childbearing and medical care [27, 28, 29, 30, 31].

Thus, BMat’s consideration of only 3 covariates limits its accuracy, as it did not account for other important socio-economic and health-related trends. While CODEm uses 19 covariates, only two of the covariates were unrelated to quality of care, fertility or mortality rates, limiting its consideration of socio-economic variables [4]. Consequently, BMat and CODEm did not model many of the factors that affect MMR, limiting their accuracy and reducing their ability to inform policymakers about which socio-economic factors should be targeted to reduce MMR.

##### 3.131 Specific Limitations of the UN MMEIG’s MMR Estimates

It may be challenging to use BMat to model different candidate policies. More specifically, it can be difficult to determine exactly how a policy would impact aggregate measures like GDP per capita and general fertility rate, which would be used to model the new MMR estimates produced as a result of the policy [9]. Additionally, while GDP per capita is a powerful predictor of MMR, the mechanism through which it impacts maternal mortality involves a variety of other factors. Therefore, using BMat to determine that increasing GDP per capita would reduce MMR would not provide politicians with enough information to craft policies that can effectively reduce MMR.

Unfortunately, researchers have observed that, when tested on out-of-sample data from more recent years, BMat can overestimate decreases in maternal mortality in low-income countries [2]. This overestimation may be due to the limitations in the modelling process discussed above. Consequently, BMat’s authors indicated the need for further exploration of possible modelling techniques [2].

##### 3.132 Specific Limitations of the GBD Study’s MMR Estimates

CODEm’s use of a combination of complex models could make it difficult to interpret the underlying associations between covariates and maternal mortality. This could hinder the estimates’ ability to effectively inform health policy.

Researchers have also noted that the GBD Study’s authors do not have access to all available national and sub-national data due to data privacy restrictions [20]. This has produced discrepancies between GBD Study and government estimates [20]. For example, researchers found that the estimates of the number of diabetes-related deaths from the German federal health reporting system were outside the uncertainty intervals of the GBD’s estimates [20]. This restricted access to data affects the accuracy of GBD estimates, and thus their ability to inform national health policy [20].

#### 3.14 The Global Maternal Health Microsimulation Model (GMatH) Estimates

The GMatH model was first proposed in 2023 [9]. The authors of the GMatH model motivated their approach by describing how the models produced by the UN MMEIG and GBD Study may inadequately describe intra-country trends [9]. More specifically, they described how the MMEIG and GBD estimates were based on statistical relationships between aggregate country-level factors and MMR, preventing them from modelling variation within a specific country [9]. In contrast, the GMatH model simulates individual women’s reproductive lifecycles to determine estimates of maternal mortality, with differences in how those lifecycles are simulated used to reflect country-level heterogeneity [9]. These estimates were produced for 200 countries and territories for every year since 1990. MMR estimates were also forecasted to 2050. Additionally, the calibrated model was used to make projections for each year up to 2050 [9].

The GMatH model used monthly cycles to simulate each stage of pregnancy and childbirth [9]. At each stage, the model estimates the probabilities of pregnancy, termination, and complications as a result of individual-level, social, and institutional risk factors [9]. Parameters governed the transition probabilities to different states within the model [9]. These parameters were estimated from probability distributions based on empirical data where possible, and on expert opinion when data was unavailable [9]. Relationships between parameters were similarly derived through a mixture of empirical data and expert opinion [9]. Parameters’ prior probability distributions were based on a hierarchical model with up to five levels (global, country income group, continent, region, and country). This allowed the model to determine parameter values for regions with sparse data, which was a strength of the GMatH model. GMatH was then fit to empirical data [9].

GMatH used 5 sets of parameters, categorised into biological parameters, family planning parameters, health system parameters, obstetrical complications, and clinical interventions [9]. Examples of biological parameters include age-specific probability of pregnancy and anaemia status, while examples of family planning parameters include contraceptive preferences. Health system parameters include the type of care available at birth and underreporting of maternal deaths. Parameters representing obstetrical complications include the risk of postpartum haemorrhage and parameters representing clinical interventions include the use of elective interventions, such as caesareans [9]. The model’s use of a wide variety of parameters causally related to maternal mortality was a strength, enabling the model to produce robust MMR estimates [9]. Additionally, by GMatH used parameters specific to demographic groups, allowing model estimates to more clearly represent local conditions [9].

To test the model’s predictive accuracy, the authors calibrated the model’s maternal death estimates using CRVS data collected between 1990 and 2015, then compared the model’s estimates for 2016 to 2020 to the CRVS estimates for the same time period. The mean absolute error for the total number of maternal deaths in test set was 47.5.

GMatH’s authors argue that, by simulating causal relationships between risks and the stage of a woman’s reproductive lifecycle, their model can use causal-inference to predict maternal outcomes more robustly than the MMEIG and GBD correlation-based approaches [9]. Additionally, and in contrast to BMat and CODEm, GMatH’s breadth of parameters allows a wide variety of policies and health system barries to be modelled [9]. For instance, GMatH has been used to investigate differences in maternal mortality between women in rural versus urban areas, as well as for women with different education levels [10]. This analysis showed the importance of addressing women’s education as an avenue for reducing maternal mortality [10]. In contrast, it is difficult to produce an effective policy to reduce MMR from observing that BMat’s MMR estimates are primarily predicted by GDP, which is a difficult outcome for politicians to change [9].

##### 3.141 Limitations of the Global Maternal Health Microsimulation Model Estimates

Similar to the limitations of BMat and CODEm discussed above, a primary limitation of the GmatH model was its use of sparse and low-quality data, again reducing its prediction accuracy [9]. For example, there were multiple instances of parameters for high-income countries using the prior distribution calculated for upper-middle income countries [23]. These parameters were generally informed by Demographic and Health Surveys, which collected data solely from lower-income countries, preventing informative priors from being generated for high-income countries [23]. Other parameters that lacked supporting empirical evidence were instead informed by expert opinion, which may not reflect the local reality [23]. Additionally, these parameters were often estimated using hierarchical models, with the associated limitations discussed in Section 3.13 [23]. These uninformative or unrepresentative priors could reduce the model’s accuracy, thus decreasing its ability to inform policy.

There were also a variety of limitations unique to the GMatH model. For example, any misspecification of the causal relationships between maternal mortality and feature variables would reduce accuracy [24]. Additionally, while the model can consider a wide variety of explanatory variables, each variable is associated with uncertainty, especially in the case of countries with little empirical data for parameter-tuning [24]. By progressively adding variables, the model may become overly influenced by uncertainty, with its estimates becoming dominated by the parameters’ uncertainty [24]. As a result, the variation in the estimates may increase beyond the point at which the estimates themselves are informative, as they cover too wide a range of outcomes [24]. Small inaccuracies in each of the parameters’ values may also accumulate, further decreasing accuracy [26]. Moreover, the more parameters included in the model, the greater its complexity, and thus the greater the chance of overfitting, reducing the model’s ability to generalise [25]. This is particularly relevant for low-income countries with little data.

GmatH’s accuracy may also be affected by how it orders simulated events. According to the original paper, the model “progresses in monthly intervals”, indicating that the simulated women’s states are updated at discrete timesteps [9]. Many of the parameters in the model are inter-related, and as a result the order they are updated can affect the model’s final estimates [24]. For example, if a woman experiences a severe complication, her chance of mortality would chance substantially depending on whether she was treated before or after occurrence of a secondary infection [25].

The validity of GmatH’s estimates is also affected by the starting state of its simulated population, which is a sampled population and thus could result in the model’s final estimates being unrepresentative [9, 24].

#### 3.15 Comparison of MMR Estimates from the Bayesian Maternal Mortality Model, Cause of Death Ensemble Model, and Global Maternal Health Microsimulation Model

On average, GMatH’s country-level MMR estimates were 22% higher than CODEm’s estimates and 19% than BMat’s estimates [49]. 85.8% of CODEm’s MMR estimates were contained within GMatH’s 95% confidence intervals compared to 88.1% of estimates from BMat [49]. The correlation between GMatH and CODEm’s MMR estimates was 0.828 versus 0.879 between GMatH and BMat [49].

Despite these similarities, Ward et al. (2025) found large variation across the models’ estimates for certain countries, such as Nigeria and Afghanistan [49]. The inter-model variation is likely due to their different methodologies and input datasets. For example, Ward et al. (2025) noticed that inter-model variation was often greatest when the only data available about a country was survey-based data about pregnancy-related mortality [49]. In contrast to maternal mortality, the cause of pregnancy-related death does not need to be related to pregnancy, childbirth, or termination [49]. To be used in the BMat and CODEm models, pregnancy-related mortality must be converted into a model-recognisable metric using a series of calculations and assumptions [49]. In contrast, the pregnancy-related mortality data can be inputted directly into GMatH, preventing the model from needing to make potentially invalid assumptions that bias the model’s estimates [49]. This difference in pre-processing may explain the high inter-model variation for these cases [49].

The variation in MMR estimates can produce confusion and uncertainty about the type of policy that should be implemented [49]. Consequently, the authors of the GMatH model describe their hope that their intrinsically different modelling approach could provide further insight into the reason why the models’ estimates have diverged [9].

### 3.2 Bayesian Maternal Mortality Model, Cause of Death Ensemble Model, & Global Maternal Health Microsimulation versus Decision-Tree Based Machine Learning

The BMis/BMat, CODEm, and GMatH models are part of the ‘data modelling culture’, as they estimate MMR by modelling the processes that generate the input data [32]. To do so, they make assumptions about the data’s structure and the relationships that exist within the data, which inform their choice of priors and covariates. In contrast, decision-tree (DT) based machine learning techniques are part of the ‘algorithmic modelling culture’, where the model focuses on predicting the outcome of interest instead of trying to learn how the data is generated, [32]. Using an algorithmic modelling approach may avoid error from uninformative priors, regional heterogeneity, and misspecification of causal relationships between MMR and feature variables.

Unlike DT models, BMat and CODEm must consider the effect of multicollinearity when selecting their feature variables. This process is described explicitly in CODEm’s documentation [32]. Multicollinearity occurs when features are linearly dependent, which makes it difficult to attribute change in MMR to a specific feature [35]. Linearly dependent variables contain similar information, making the model more likely to learn noise in the data, overfit and generalise less easily [35]. However, use of a small subset of features can cause the models to ignore valuable information about how other health-related and socio-economic variables affect MMR, reducing their predictive accuracy.

In contrast, DT models are particularly suited to working with high-dimensional data, as splits in the individual DTs are determined by the feature partition that best reduces error [36, 37]. Thus, if three feature variables are highly correlated and one of the features is already used in a split, the others are less likely to be chosen for future splits because they would not add additional information. Therefore, DT methods can include all features that could influence MMR with a lower risk of overfitting, unlike the previously described approaches. Additionally, if a variable does not help reduce predictive error because it is not correlated with MMR, it will not be used in any splits and thus will be ignored [36, 37]. Therefore, error in redundant and uninformative variables will contribute less to uncertainty in the final predictions from DT models than GMatH, which relies on all variables. Consequently, DT models can work well with high-dimensional datasets.

Another strength of DT models is their treatment of missing values. They use surrogate splits or default directions to handle sparse data, as described in the background. This prevents them from needing to use imputation methods that may introduce error or Bayesian hierarchical models that may overly smooth regional heterogeneity. As the number of dimensions increase, the likelihood that a specific section of the input space contains data decreases, thus increasing sparsity [6]. As a result, DT models’ ability to work with sparsity may help them avoid further limitations of high-dimensional data [50].

As a consequence of these benefits, researchers have stated the importance of exploring how algorithmic modelling approaches can be used to improve prediction of maternal health outcomes [34]. Thus, I propose the use of a decision-tree based machine learning model to estimate MMR. This model will provide an alternate set of estimates to help form a consensus out of the three existing estimates presented in the literature.

### 3.3 Algorithmic Machine Learning Techniques in Public Health

To further motivate my use of algorithmic machine learning (ML) models, I present examples of how these approaches have been used in public health research. In this section, I refer to all algorithmic models as ML models to use their common name despite the data modelling approaches discussed above technically also being ML models.

#### 3.31 General Overview

Machine learning models are being applied to a wide variety of public health research problems to take advantage of the large quantities of health data being generated by wearable devices, clinical records, and social media [40]. For example, they can be used for image-based medical diagnostic tasks, improving operational efficiency, predicting patient-specific risks, and drug discovery [40]. ML models can be categorised as ‘white box’, ‘grey box’, and ‘black box’ depending on their level of interpretability, where ‘white box’ models are the most interpretable and ‘black box’ are the least [39]. ‘Black box’ models include deep learning and neural networks and are frequently used in image-based diagnosis [39]. Unfortunately, their complexity and lack of interpretability make it difficult to identify the feature variables with the highest predictive power, reducing their ability to inform policy targets [39]. In contrast, ‘white-box’ models like decision tree-based algorithms can be used to predict medical risk factors and complications in an interpretable manner, allowing the model user to understand the factors used by the model to produce its results [39].

Mahajan et al.’s (2023) review discussed the use of ensemble-based models in public health. They found that bagging and boosting algorithms were the most popular in the surveyed literature, as they were used in 41 and 37 of 45 studies, respectively [48]. However, they were only the most accurate algorithms evaluated in the study in 26.8 and 40.5% of instances [48]. In contrast, stacking and voting were less frequently used (23 and 7 out of 45 studies), but they had the highest accuracies 82.6% and 71.4% of the time, respectively [48]. Stacking models’ high performance was attributed to their ability to learn the best base estimators [48]. As a result, stacking models have been used to predict incidence of diabetes, heart disease, liver disease, and skin cancer [48].

#### 3.32 Estimations of Cause Specific Maternal Mortality and Risk of Mortality

Many studies that predict maternal health outcomes using DT based methods focus on estimating patient risk and cause-specific maternal mortality. As a result, much of the ML research in this domain uses classification models, which can categorise a woman’s mortality risk as ‘high’, ‘medium’, or ‘low’. While I use a regression model in my thesis, I include examples of classification models to contextualise how DT-based ML is being used in maternal health research. As described below, DT and boosting models generally had the highest, or among the highest, performance, motivating their use in my thesis.

Akazawa et al. (2021) used ML models to classify a woman’s risk of postpartum haemorrhage, a leading cause of maternal mortality, to inform treatment [41]. They compared the performance of logistic regression, DT, random forest (RF), boosted tree, and deep learning models trained on 11 clinical variables [41]. The boosted tree model had the highest accuracy [41]. However, the model was trained and tested on data from the same institution, potentially reducing its generalisability [41].

Similarly, researchers have used ML techniques to predict a woman’s risk of pre-eclampsia, another leading cause of maternal mortality, to improve identification and treatment of high-risk pregnancies [42]. The study trained Elastic Net and gradient boosting (GB) models on a range of medical and socio-demographic covariates, with similar performance between the two models [42]. However, the model was trained and tested on data from a single referral hospital, which had a higher proportion of high-risk patients, potentially reducing the model’s generalisability [42].

As a final example, Sylvain et al. (2025) predicted adverse pregnancy outcomes in Rwanda using logistic regression, DT, RF, GB models, support vector machines, and neural networks [46]. The RF and GB models had the highest accuracy (90.6% and 88.49%, respectively) [46]. The study also determined the most predictive variables [46]. However, the study predicted occurrence of adverse outcomes as a binary variable, with a negative outcome encompassing a wide variety of possible maternal and neonatal health events [46]. This could reduce nuance in the model’s predictions and thus its ability to inform treatment [46]. This model may also not generalise to rural regions, as it was only trained on data from district hospitals, which are only responsible for only roughly 35% of births in Rwanda [46].

Khadidos et al. (2024) used a stacking-ensemble model to classify maternal health risk in Bangladesh, where they trained GB, RF, DT, and k-Nearest Neighbours models as the base estimator, with each base estimator trialled as the meta-learner [47]. Using the GB model as the meta-learner had the highest precision (0.86), with all DT based stacking ensemble outperforming sole use of bagging or boosting [47].

Machine learning classifiers are also used to predict health system attributes. For instance, Taye et al. (2025) used a RF classifier to predict whether a birth was attended by a skilled birth attendant using a mixture of socio-economic and health system quality variables [44]. The RF model was also used to indicate the most predictive variables [44]. The model achieved 92% accuracy despite being trained and tested on survey data, which is known to be of lower quality [44]. This data was imputed, which may introduce bias into the model’s estimates [44]. However, this study shows how DT based ensembles can achieve high performance with low-quality data. Another example is Fredriksson et al.’s (2022) paper, which compared the performances of a RF model and artificial neural network to more classical statistical models when classifying the likelihood of a woman delivering her baby in a health facility [45]. The RF had the highest classification accuracy (74%), with the paper also reporting the most predictive variables [45]. The lower accuracy may be related to the authors’ use of imputation [45].

#### 3.33 Estimation of Maternal Mortality

There is a severe lack of studies that use algorithmic modelling techniques to estimate MMR and the number of maternal deaths. The only relevant study found in my literature review was published in 2025 and only estimated MMR for Bangladesh [38]. The authors compared the performance of a Bidirectional Recurrent Neural Network and an Elastic Neural Network to predict MMR, with the associated root mean square errors being 3.30 and 3.44 per 100,000 live births, respectively [38]. This study was severely weakened by its dataset size, as it reported having only 21 observations of MMR, which is insufficient to train a robust model [38]. No critique can be made of its feature variables, as these did not appear to be reported. Despite this study’s limitations, it serves as proof of concept for using algorithmic modelling techniques to estimate MMR. The lack of robust studies doing this type of analysis highlights a gap in the literature that my thesis aims to fill.

### 3.4 Conclusion

In conclusion, the UN MMEIG and Global Burden of Disease models are the primary modelling techniques for estimating global maternal mortality ratios. The GMatH is a newer method for MMR prediction. These three methods are limited by their low-quality, sparse input data. As a result, estimates for countries with less data may be estimated with parameters generated by uninformative priors. The parameter values and overall MMR estimates may also be oversmoothed, with the modelling techniques ignoring potential regional heterogeneity. Given that countries with less data tend to have higher MMR, this can reduce the estimates’ ability to inform national health policy. The models are also weakened by their consideration of only a small subset of the socio-economic and health-related variables that impact maternal mortality. Using DT based models would circumvent the need to model the data generating process, eliminating assumptions about data distribution and the need for priors. Additionally, DT models can handle a wide range of feature variables, allowing them to make more comprehensive estimates, with evaluation of feature importance covering a wider range of domains. In maternal health research, ML methods are generally used to classify a woman’s overall and cause-specific maternal mortality risk. I found only one algorithmic modelling approach to estimating MMR, with the study solely predicting MMR in Bangladesh. Thus, there is a gap in the literature about how a decision-tree based ML model can be used to estimate MMR at a global level.

As a final note, building a model with an entirely new methodology will produce another set of MMR estimates. My model’s estimates can be compared to the literature and contribute to resolving some of the lack of consensus around current MMR estimates.

Citations:

1. Trends in maternal mortality estimates 2000 to 2023: estimates by WHO, UNICEF, UNFPA, World Bank Group and UNDESA/Population Division. Geneva: World Health Organization; 2025. Licence: CC BY-NC-SA 3.0 IGO
2. The Annals of Applied Statistics 2017, Vol. 11, No. 3, 1245–1274 DOI: 10.1214/16-AOAS1014 © Institute of Mathematical Statistics, 2017
   1. 2017 un mmeig estimates
3. <https://www.tandfonline.com/doi/full/10.1080/2330443X.2023.2286313#abstract>
   1. Bmis model information
4. <https://www.thelancet.com/cms/10.1016/S0140-6736(24)00367-2/attachment/b7b1d025-a598-408a-b62e-025ebc1f08d1/mmc1.pdf>
   1. Supplement to: GBD 2021 Causes of Death Collaborators. Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2024; published online April 3. https://doi.org/10.1016/ S0140-6736(24)00367-2.
5. <https://www.nature.com/articles/s41591-022-01990-1>
   1. Gbd at 30 years
6. <https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(24)00367-2/fulltext>
   1. Gbd 2021
7. <https://pophealthmetrics.biomedcentral.com/counter/pdf/10.1186/1478-7954-10-1.pdf>
   1. Foreman KJ, Lozano R, Lopez AD, Murray CJL. **Modeling causes of death: an integrated approach using CODEm**. Population Health Metrics. 2012; 10:1.
8. <https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-021-01501-1>
   1. Garbage codes in the gbd studies
9. <https://www.nature.com/articles/s41591-023-02310-x>
   1. Ward, Z.J., Atun, R., King, G. *et al.* Simulation-based estimates and projections of global, regional and country-level maternal mortality by cause, 1990–2050. *Nat Med* **29**, 1253–1261 (2023). <https://doi.org/10.1038/s41591-023-02310-x>
10. <https://www.sciencedirect.com/science/article/pii/S2589537024002323>
11. <https://www.sciencedirect.com/science/article/pii/S2666827025000714>
12. <https://academic.oup.com/bmb/article/121/1/121/2926155>
13. <https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(24)00560-6/fulltext>
    1. Lancet global determinants of MMR
14. <https://www.mdpi.com/1999-5903/15/8/255>
    1. Arima
15. <https://link.springer.com/article/10.3758/s13428-023-02204-3#Sec36>
    1. Bayesian hierarchical model tutorial
16. <https://pubmed.ncbi.nlm.nih.gov/17274999/>
    1. Maternal mortality, stillbirths, and measures of obstetric care
17. Onambele, L.; Guillen-Aguinaga, S.; Guillen-Aguinaga, L.; Ortega-Leon, W.; Montejo, R.; Alas-Brun, R.; Aguinaga-Ontoso, E.; Aguinaga-Ontoso, I.; Guillen-Grima, F. Trends, Projections, and Regional Disparities of Maternal Mortality in Africa (1990–2030): An ARIMA Forecasting Approach. Epidemiologia 2023, 4, 322–351. https://doi.org/ 10.3390/epidemiologia4030032
18. <https://pmc.ncbi.nlm.nih.gov/articles/PMC3886639/>
    1. 2014 un mmeig model
19. Zaba B, Calvert C, Marston M, Isingo R, Nakiyingi-Miiro J, Lutalo T et al. Effect of HIV infection on pregnancyrelated mortality in sub-Saharan Africa: secondary analyses of pooled community-based data from the network for Analysing Longitudinal Population-based HIV/AIDS data on Africa (ALPHA). Lancet.
    1. 2013;381(9879):1763-71 (<https://doi.org/10.1016/s0140-6736(13)60803-x>).
20. <https://link.springer.com/article/10.1007/s00103-018-2793-0>
    1. Germany gbd
21. <https://link.springer.com/article/10.1186/s13690-020-00458-3?fromPaywallRec=true>
    1. Who pressure
22. <https://www.tandfonline.com/doi/epdf/10.1016/S0968-8080%2811%2937550-7?needAccess=true>
    1. Ihme vs un
23. <https://static-content.springer.com/esm/art%3A10.1038%2Fs41591-023-02310-x/MediaObjects/41591_2023_2310_MOESM1_ESM.pdf>
    1. Gmath methods
24. <https://www.jstor.org/stable/pdf/2998681.pdf>
    1. Jstor microsimulation paper
25. <https://researchsystem.canberra.edu.au/ws/portalfiles/portal/8849112/rm2013000054_journal_article.pdf>
    1. UC microsimulation
26. <https://journals-sagepub-com.virtual.anu.edu.au/doi/epdf/10.1177/0894439310370085>
    1. More microsimulation
27. <https://onlinelibrary.wiley.com/doi/10.5694/mja2.52452>
    1. Non-communicable disease and pregnancy
    2. Ramson et al., 2024
28. Akselrod S, Banerjee A, Collins TE, Acharya S, Artykova N, Askew I, Berdzuli N, Diorditsa S, Eggers R, Farrington J, Jakab Z, Ferreira-Borges C, Mikkelsen B, Azzopardi-Muscat N, Olsavszky V, Park K, Sobel H, Tran H, Vujnovic M, Weber M, Were W, Yaqub N, Berlina D, Dunlop CL and Allen LN (2023) Integrating maternal, newborn, child health and non-communicable disease care in the sustainable development goal era. Front. Public Health 11:1183712. doi: 10.3389/fpubh.2023.1183712
29. [**https://pubmed.ncbi.nlm.nih.gov/27642018/**](https://pubmed.ncbi.nlm.nih.gov/27642018/)
30. [**https://pmc.ncbi.nlm.nih.gov/articles/PMC11117177/pdf/jogh-14-04128.pdf**](https://pmc.ncbi.nlm.nih.gov/articles/PMC11117177/pdf/jogh-14-04128.pdf)
31. [**https://www.sciencedirect.com/science/article/pii/S2589537024002323**](https://www.sciencedirect.com/science/article/pii/S2589537024002323)
32. [**https://projecteuclid.org/journals/statistical-science/volume-16/issue-3/Statistical-Modeling--The-Two-Cultures-with-comments-and-a/10.1214/ss/1009213726.full**](https://projecteuclid.org/journals/statistical-science/volume-16/issue-3/Statistical-Modeling--The-Two-Cultures-with-comments-and-a/10.1214/ss/1009213726.full)
    1. **Two cultures: statistics vs ML**
33. <https://www.thelancet.com/action/showPdf?pii=S2214-109X%2823%2900468-0>
    1. Global determinants of maternal health -> lancet
34. [**https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-025-23528-z**](https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-025-23528-z)
    1. Maternal hypertensive disorder projections
35. <https://www.mdpi.com/2227-7390/10/8/1283>
    1. Multicollinearity
36. <https://projecteuclid.org/journals/annals-of-statistics/volume-43/issue-4/Consistency-of-random-forests/10.1214/15-AOS1321.full>
    1. RF and high-dimensionality
37. <https://onlinelibrary.wiley.com/doi/epdf/10.1111/insr.12016>
    1. Decision tree info
38. Molla MM, Hossain MS, Ali MA, Islam MR, Sultana MP, Roy DC (2025) Exploring the achievements and forecasting of SDG 3 using machine learning algorithms: Bangladesh perspective. PLoS ONE 20(2): e0314466. <https://doi.org/10.1371/journal.pone.0314466>
39. Digital Object Identifier 10.1109/ACCESS.2024.3399827
    1. Box models
40. <https://eurjmedres.biomedcentral.com/articles/10.1186/s40001-025-02680-7>
    1. Brief ml pub health overview
41. <https://www.nature.com/articles/s41598-021-02198-y>
    1. Postpartum haemorrhage prediction
42. <https://www.sciencedirect.com/science/article/pii/S2589933320300306?via%3Dihub>
    1. Pre-eclampsia prediction
43. <https://pmc.ncbi.nlm.nih.gov/articles/PMC9379432/pdf/PAMJ-42-89.pdf>
    1. Medical supplies demand
44. <https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-025-22007-9#Sec2>
45. <https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2022.855236/full>
46. <https://link.springer.com/article/10.1186/s12911-025-02921-z?fromPaywallRec=false>
47. <https://www.nature.com/articles/s41598-024-71934-x#Abs1>
48. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10298658/>
    1. Ensemble review
49. <https://www.sciencedirect.com/science/article/pii/S2589537025004389>
    1. Gmat latest comparison (2025)
50. <https://www.nature.com/articles/s41592-018-0019-x>
    1. Curse of dimensionality