March 2020

BAYESIAN METHODS FOR THE ASSESSMENT OF REPORTING ERRORS FOR DATA-SPARSE POPULATION-PERIODS WITH APPLICATIONS TO ESTIMATING MORTALITY

Emily Peterson
University of Massachusetts Amherst

BAYESIAN METHODS FOR THE ASSESSMENT OF REPORTING ERRORS FOR DATA-SPARSE POPULATION-PERIODS WITH APPLICATIONS TO ESTIMATING MORTALITY

A Dissertation Presented

by

EMILY N PETERSON

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

February 2020

Biostatistics and Epidemiology

BAYESIAN METHODS FOR THE ASSESSMENT OF REPORTING ERRORS FOR DATA-SPARSE POPULATION-PERIODS WITH APPLICATIONS TO ESTIMATING MORTALITY

A Dissertation Presented

by

EMILY N PETERSON

Approved as to style and content by:

Leontine Alkema, Chair

Ken Kleinman, Member

Laura Balzer, Member

Michael Lavine, Member

Lisa Chasan-Taber, Department Head
Biostatistics and Epidemiology

DEDICATION

To my mother and father who have been my champions, and supported me at every step of my journey. Thank you for your unlimited love and support.

ACKNOWLEDGMENTS

Thanks to my advisor, Leontine Alkema, for her incredible support and mentorship during these past five years. You have set an example of excellence as a researcher and as a role model. I have felt so blessed to have been able to learn from you, and also to have had your support and care during some of the more challenging times. I will forever cherish our friendship and my experience with you.

Thanks to my committee members, Laura, Ken and Michael, who have given thoughtful and constructive feedback on this disseration, but who have also dedicated their own time to help guide me as a student.

Thanks to the collaborators from the United Nations Maternal Mortality Estimation Group and John's Hopkins University for their collaboration, and sponsorship of these projects. In particular, thanks to Doris Chou, Lale Sale, and Stephane Helleringer for giving me the opportunity to collaborate with them.

Thanks to my friends and family who have inspired, motivated, and supported me throughout my academic career. Thank you Tarzan, Olivia, and Roary, the best pets in the world.

ABSTRACT

BAYESIAN METHODS FOR THE ASSESSMENT OF REPORTING ERRORS FOR DATA-SPARSE POPULATION-PERIODS WITH APPLICATIONS TO ESTIMATING MORTALITY

FEBRUARY 2020

EMILY N PETERSON

B.S., DAVIDSON COLLEGE

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Leontine Alkema

Population level mortality data is often subject to substantial reporting errors due to misclassification of cause of death, misclassification of death status, or age reporting errors. Accuracy of error-prone data sources can be assessed by comparing such data to gold standard data for the same population-period. We present Bayesian methods for assessing the extent of reporting errors across different population-periods and generalizing those to settings where gold-standard data are lacking. Firstly, we investigate misclassification errors of maternal cause of death reporting in civil registration vital statistics data. We use a Bayesian hierarchical bivariate random-walk model to estimate country-year specific sensitivity and specificity in countries with at least one period where vital registration data overlaps with gold standard data. For countries without gold standard data, we developed a sequential approach, in which fixed global estimates of sensitivity and specificity are used. Additionally, we propose a new approach to incorporate temporal structure of misclassification parameters. Secondly, we investigate misreporting of adult mortality in sibling

survival history data. Sibling survival histories data suffers from reporting errors due to respondent misreporting of birth year and age at death of their maternal siblings. We perform an exploratory analysis of data collected in Malawi and propose a candidate parametrization for reporting errors in cohort survival probabilities by 5-year age groups. We introduce parameters to capture age-group specific age-at-death errors and birth year reporting errors and define the data generating processes that relate sibling survival data to true survival probabilities while accounting for reporting errors. This framework allows for the estimation of age-group specific survival probabilities in settings where only error-prone sibling survival history data is available.

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Population level mortality data is often subject to substantial reporting errors. To address this concern, we developed a Bayesian misreporting model framework for the assessment of the extent of reporting errors across different population-periods, using gold standard data to inform estimates of misreporting. However, in the case of national and global mortality estimation, gold standard data is limited and sparse, therefore, the challenge is to estimate the true outcome of interest given the observed error-prone data available. In the absence of better quality data, and limited gold-standard data, we developed a new approach that allows us to extrapolate misreporting estimates to population-periods lacking gold-standard data. This approach is a 2-step process in which we aimed to estimate global levels of misreporting parameters using all country-periods with available information on misreporting. Subsequently, we used a sequential modeling approach, to extrapolate global estimates of reporting errors for country-periods without gold standard data into a larger mortality estimation model. This improves upon a more common and simpler approach in which, in the absence of a modeling different reporting error processes, ratios of mortality rates are used to assess differences between the true and biased rates. We applied our proposed Bayesian mispreporting framework in the context of maternal mortality estimation and adult cohort specific survival probability estimation.

In the estimation of national trends of maternal mortality, estimates are constructed using a Bayesian hierarchical time series regression model, referred to as BMat (UN MMEIG 2019), which uses civil registration vital statistics (CRVS) data to inform model based estimates. However, CRVS data is prone to substantial reporting errors. Specifically, reporting errors are introduced in the misclassification of cause of death. A main concern

is accounting for bias introduced in the misclassification of maternal deaths. Based on a systematic review of studies, which report information on levels of misclassification within countries, we developed a Bayesian bivarirate random walk model to assess misclassification errors across different population-periods. We model misclassification of maternal deaths using country-year estimates of sensitivity and specificity. Our aim is to generalize these results to population-periods without available data on misclassification. To do so, we use a sequential approach in which we apply global estimates of sensitivity and specificity for countries without information given by misclassification studies. Therefore, within the larger BMat model, we estimate the true proportion of maternal deaths given error-prone CRVS maternal mortality data, and fixed global estimates of misclassification.

Sibling survival history surveys (SSH) is an indirect method used to estimate age-group cohort specific adult mortality rates for countries with limited CRVS data. SSH data consists of respondent reported information on the vital status, current age, and age at death of all their maternal siblings. However, reporting errors occur when a respondent misreports the vital status and/or ages of their siblings, which are broken down into different reporting error processes. Reporting errors due to age at death misreporting results in vital status errors (misreporting of death status) or siblings that are omitted/added to the population. Age misreporting, referred to as birth year misreporting, results in siblings being classified into incorrect age cohorts, which will inflate the false age-cohort mortality rate and conversely deflate the true age-cohort mortality rate. We extend upon the approach used for maternal mortality estimation to incorporate misreporting parameters related to both age at death misclassification, and birth year misreporting. We propose parametrization to assess the extent of both age at death and birth year reporting errors, based on limited preliminary data.

The Bayesian misreporting framework we propose is applicable in multiple settings, which assess mortality trends. Namely, the Institute of Health Metrics publishes the Global Bur-

den of Disease Study (IHME 2014), which describes mortality from major diseases at global and national levels. In this application, assessment of global trends in cause-specific mortality rates may be improved by a Bayesian misclassification model that incorporates global and country level estimates of misclassification as with maternal mortality estimation. Additionally, Masquelier et al. (2018), assessed mortality of children under 5 years of age for years 1990-2016. This estimation was based on surveys, vital statistics, and census data, which suffer from the same reporting errors as both CRVS data and sibling survival histories. Therefore, this framework provides a new approach to account for reporting errors in applications that assess cause or age-group specific mortality rates in many applications.

The paper is organized as follows: 1) In Chapter I, we first introduce the Bayesian misclassification model used to estimate global and country levels of misreporting in maternal mortality, 2) In Chapter II, we describe the sequential approach taken to incorporate results from Chapter I into a larger estimation model. Lastly, in Chapter III, we propose a preliminary parametrization to assess misreporting of age-group cohort specific survival probabilities, which is an extension of both Chapters I and II.

CHAPTER 1

A BAYESIAN HIERARCHICAL BIVARIATE RANDOM WALK MODEL TO ESTIMATE SENSITIVITY AND SPECIFICITY OF REPORTING OF MATERNAL CAUSE OF DEATH IN NATIONAL CIVIL REGISTRATION VITAL STATISTICS SYSTEMS

1.1 Introduction

A maternal death is "the death of a woman whilst pregnant or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes" as defined in International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) (World Health Organization, 2010). National civil vital registration (CRVS) systems record the number of deaths to women of reproductive ages, as well as the cause associated with each death using ICD coding. Based on the number of all-cause and maternal deaths, the proportion of deaths that are of a maternal cause, referred to as the proportion maternal (PM), can be constructed.

Under ideal circumstances, when all deaths are captured and all causes are accurately classified, CRVS systems provide perfect information on the number of maternal deaths within the country. However, even if routine registration of deaths is in place, maternal deaths may be reported incorrectly if deaths are unregistered or misclassified, where misclassification of deaths refers to incorrect coding in vital registration systems, due either to error in the medical certification of cause of death or error in applying the correct ICD code. The accuracy of CRVS systems can be assessed by comparing CRVS-based observed PMs to those obtained from specialized studies, which are rigorous assessments of maternal mortality for a given country-period. Prior work comparing the ratio of study-based

PMs to CRVS-based PMs, referred to as CRVS adjustment factors, found that these ratios are around 150%, thus suggesting that PMs obtained from CRVS do not adequately capture all maternal deaths (Wilmoth et al., 2012).

The United Nations Maternal Mortality Estimation Inter-agency group (UN MMEIG) is responsible for publishing internationally comparable estimates of maternal mortality for UN reporting. Since 2015, UN MMEIG estimates have been produced using a Bayesian hierarchical time series regression model, referred to as BMat (UN MMEIG 2015, Alkema et al 2017, UN MMEIG 2019). BMat uses an input database which is based upon nationally representative data available from Civil Registration Vital Statistics (CRVS), population-based surveys such as DHS and MICS, censuses, and specialized surveillance. A more general explanation of these data sources and their limitations is included in the UN MMEIG 2019 report (UN MMEIG 2019). In BMat, estimates of the PM are produced based on the available input data for the respective country-period, taking account of data quality issues in reporting. Based on the Wilmoth et al. analyses, the UN MMEIG has applied adjustments to CRVS data, to reduce bias in CRVS-based derived data in settings where CRVS systems are subject to error (UN MMEIG 2015, Alkema et al. 2017). The approach was subject to limitations (Alkema et al. 2017).

In this paper, we develop a new approach to estimate reporting errors associated with misclassification in maternal death reporting in CRVS data that improves upon limitations of the UN MMEIG 2015 approach. The next section introduces terminology and the framework used to describe errors in reporting of maternal mortality in CRVS systems. Section 1.3 provides information on the data available to inform estimation of the extent of incorrect reporting. Section 3.4 introduces a Bayesian model to estimate the extend of misclassification in the reporting of maternal deaths in CRVS systems. The estimation is based on summarizing misclassification in terms of sensitivity and specificity, and modeling these two indicators for all country-years with CRVS data using a bivariate hierarchical random walk model. Finally, we present findings and the results of validation exercises.

This work provides a new approach to modeling misclassification errors that quantifies dependence between sensitivity and specificity of reporting and allows for extrapolation to country-periods without validation data. Previous work on Bayesian estimation models of the extent of cause of death misclassification include work by Paulino et al. (2004) and Stamey et al. (2008). Paulino et al. (2004) used a Bayesian approach to account for misclassification of binomial data in a logistic regression model. In their approach, independent Beta priors were assigned to sensitivity and specificity. Similarly, Stamey et al. (2008) proposed a Bayesian approach to adjust for misclassification in death count data using Poisson regression model. They used informative Beta priors to account for lack of observed information on sensitivity and specificity. These existing approaches do not allow for extrapolation to country-periods without validation data, taking account of correlation between sensitivity and specificity.

1.2 Reporting errors in CRVS systems

The diagram in Figure 1.1 illustrates the breakdown of total deaths to women of reproductive age by CRVS-reporting status (columns) and true maternal cause (rows). In a complete-CRVS setting, meaning that all deaths are registered, the number of missed deaths (3rd column) is equal to zero, such that reporting errors are solely due to misclassification of deaths. Inaccurate attribution of cause of death is either due to error in the medical certification of cause of death, and/or error in applying the correct code, which results in two misclassification biases regarding maternal deaths. Firstly, error occurs when a maternal death is misclassified as non-maternal, referred to as a false negative (F^-) maternal death. Secondly, if a non-maternal death is misclassified as maternal, the death is labeled as a false positive maternal death (F^+). Correctly classified maternal and non-maternal deaths are indicated by true positive (T^+) and true negative (T^-) maternal deaths, respectively. From the individual categories in Figure 1.1, cumulative totals are calculated summing

across rows and columns, i.e. CRVS reported maternal deaths is the sum of $T^{(+)}$ and $F^{(+)}$, whereas, the true number of maternal deaths within the CRVS is the sum of $T^{(+)}$ and $F^{(-)}$. In incomplete CRVS systems, missed deaths include unregistered maternal deaths, referred to as $U^{(+)}$ deaths, and unregistered non-maternal deaths $U^{(-)}$.

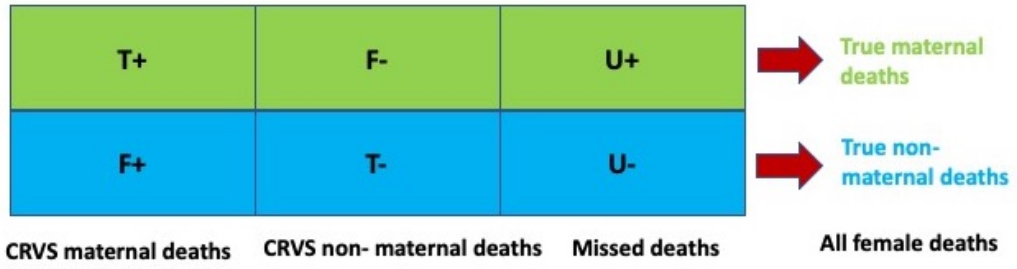

Figure 1.1: Diagram of breakdown of total deaths to women of reproductive age for a country-year, by CRVS-reporting status (columns) and true maternal cause (rows). $T^{(+)}$ and $F^{(-)}$ deaths refer to maternal deaths that are correctly registered as maternal deaths, and incorrectly registered as non-maternal deaths, respectively. Similarly, $F^{(+)}$ and $T^{(-)}$ maternal deaths refers to non-maternal deaths that are incorrectly registered as maternal deaths, and correctly registered as non-maternal deaths, respectively. $U^{(+)}$ refers to unregistered maternal deaths, and $U^{(-)}$ refers to unregistered non-maternal deaths.

1.3 Data

Information on CRVS misclassification errors and unregistered deaths was obtained from comparing information from specialized studies to CRVS reported deaths. This section discusses both types of data.

1.3.1 CRVS data and completeness assessment

The WHO Mortality Database maintains data from CRVS systems. Using this database, we obtained information on the number of maternal deaths reported in the CRVS and the number of deaths to women aged 15-49 reported in the CRVS (CRVS envelope).

Completeness of the reporting of deaths into the CRVS system was assessed by comparing CRVS reported deaths to WHO estimates of deaths to women of reproductive age, obtained from life tables for WHO Member States. We first calculated the annual ratio of female deaths reported in the CRVS over deaths estimated by the WHO for all years with CRVS

data, based on a moving window of 5-year periods (five-year periods were used to obtain less variable ratios for countries with smaller populations). If the ratios, more specifically, the upper bounds of 95% confidence intervals when accounting for stochastic uncertainty in the ratio, are greater than 0.95 for all years with CRVS data, we assumed that the CRVS was complete in the country during the entire period. Otherwise, CRVS completeness was given by the ratio for each individual year (UN MMEIG 2019).

1.3.2 Specialized studies

A specialized study is defined as the assessment of maternal mortality for a country-period, either independent of CRVS reported data or based on the checking of CRVS reported deaths. These studies provided counts of the number of true maternal deaths (first row in Figure 1.1) or possibly individual categories, i.e. the number of false negative maternal deaths. We assumed that the study envelope was equal to the envelope reported by the CRVS system, unless specified otherwise in the study. Specialized studies were obtained through (1) a literature review, (2) the UN MMEIG 2015 maternal mortality data base (UN MMEIG 2015), and (3) information provided by countries based on a follow-up survey, sent to countries in response to discussions with the Pan-American Health Organization (PAHO), and during country consultation. Detailed information on the compilation of specialized studies data is given in Appendix Section A.0.2.

1.3.3 Data availability

A total of 50 study documents contributed data to inform the CRVS adjustment model, referring to 33 unique countries and 221 unique country-periods (observations). The majority of included study documents were obtained through the systematic search ($n = 22$). In addition, 18 study documents were obtained from the UN MMEIG 2015 database (UN MMEIG 2015). Additional information from follow-up surveys and communication with countries during country consultation yielded 10 additional study documents (Appendix

Section A.0.2).

Reported information varied greatly across observations. While some studies reported a detailed breakdown of false positive and/or false negative maternal deaths, the majority of studies reported only the confirmed total number of maternal deaths for a given country-period, see Table 1.1. The majority of studies reported on the true number of maternal deaths within the CRVS (184 observations, 30 countries). Information on both false negative and false positive breakdowns was available for 18 observations (4 countries). Most studies with breakdown information solely reported on false negative breakdowns, 38 observations from 4 countries. Data regarding the relative difference between the proportion of maternal deaths among CRVS-reported deaths and the proportion of maternal deaths among unregistered deaths was very limited: only 13 observations reported information that included U^+.

Reported counts	# of observations	# of countries
True maternal in CRVS only	162	27
True maternal in CRVS and U+	2	1
F- and F+ and U+	10	2
F- and F+ only	8	2
F- and U+ only	1	1
F- only	38	4
Total	221	33

Table 1.1: Overview of data available from specialized studies.

1.4 Methods

1.4.1 Summary of modeling approach

Based on the 6-box model, refer to Figure 1.1, for each country c and year t, we assumed a multinomial data generating distribution as follows:

$$\boldsymbol{y}_{c,t}|y_{c,t}^{(\text{tot})}, \boldsymbol{\rho}_{c,t} \sim Multinom\left(y_{c,t}^{(\text{tot})}, \boldsymbol{\rho}_{c,t}\right), \tag{1.1}$$

where (leaving out subscripts (c,t) for improve readability):

$$\boldsymbol{y} = \left(y^{(\text{U-})}, y^{(U+)}, y^{(\text{T-})}, y^{(\text{T+})}, y^{(\text{F-})}, y^{(\text{F+})}\right),$$
$$\boldsymbol{\rho} = \left(\rho^{(\text{U-})}, \rho^{(U+)}, \rho^{(\text{T-})}, \rho^{(\text{T+})}, \rho^{(\text{F-})}, \rho^{(\text{F+})}\right),$$

with $y^{(\text{b})}$ the number of deaths reported for category b in $B = \{T^+, T^-, F^+, F^-, U^+, U^-\}$ and $y^{(\text{tot})} = \sum_{b \in B} y^{(\text{b})}$. Similarly, $\rho^{(\text{b})}$ denotes the probability of a death in category b and $\sum_{b \in B} \rho^{(\text{b})} = 1$. Lastly, observed proportions are denoted with $p^{(\text{b})} = y^{(\text{b})}/y^{(\text{tot})}$. Focusing on deaths captured in the CRVS data only, hence categories $B^{(\text{CRVS})} = \{T^+, T^-, F^+, F^-\}$, we define the total number of deaths in the CRVS as $y^{(\text{CRVS})} = \sum_{b \in B^{(\text{CRVS})}} y^{(\text{b})}$, CRVS-based probabilities $\gamma^{(\text{b})} = \gamma^{(\text{b})} / \sum_{b \in B^{(\text{CRVS})}} \rho^{(\text{b})}$, and CRVS-based proportions $q^{(\text{b})} = p^{(\text{b})} / \sum_{b \in B^{(CRVS)}} p^{(\text{b})}$. The proportion of CRVS-based deaths that is reported as being maternal (the CRVS-based observed PM) is given by $q^{(\text{matCRVS})} = \left(y^{(\text{T+})} + y^{(\text{F+})}\right) / y^{(\text{CRVS})}$.

The question of interest is how to estimate the true probability of a maternal death, $\rho^{(\text{truemat})} = \rho^{(\text{T+})} + \rho^{(\text{F-})} + \rho^{(U+)}$, based on the CRVS-reported maternal deaths $y^{(\text{matCRVS})}$, total CRVS-reported deaths $y^{(\text{CRVS})}$, and total deaths $y^{(\text{tot})}$. Based on Eq. 1.1 we find

$$\begin{aligned} y^{(\text{matCRVS})} | y^{(\text{CRVS})}, \gamma^{(\text{matCRVS})} &\sim Bin\left(y^{(\text{CRVS})}, \gamma^{(\text{matCRVS})}\right), \\ \gamma^{(\text{matCRVS})} &= \left(\rho^{(\text{T+})} + \rho^{(\text{F+})}\right) / \rho^{(\text{CRVS})}, \end{aligned} \tag{1.2}$$

where $\gamma^{(\text{matCRVS})}$ refers to the probability of reporting a death as being maternal in CRVS. For country-years with complete CRVS, this probability $\gamma^{(\text{matCRVS})}$ can be expressed as a function of the true probability $\rho^{(\text{truemat})}$, and misclassification parameters sensitivity $\lambda^{(+)}$ and specificity $\lambda^{(-)}$ as follows:

$$\gamma^{(\text{matCRVS})} = \lambda^{(+)} \rho^{(\text{truemat})} + \left(1 - \lambda^{(-)}\right)\left(1 - \rho^{(\text{truemat})}\right) \tag{1.3}$$

with sensitivity $\lambda^{(+)} = \frac{\gamma^{(T+)}}{\gamma^{(T+)}+\gamma^{(F-)}}$, the probability of correctly identifying a maternal death reported in the CRVS as such, and specificity $\lambda^{(-)} = \frac{\gamma^{(T-)}}{\gamma^{(T-)}+\gamma^{(F+)}}$, the probability of correctly identifying a non-maternal death reported in the CRVS as such.

For countries with incomplete CRVS systems, we define $\omega^{(\text{truematUNREG})}$ to be the probability of a maternal death among unregistered deaths, and $\gamma^{(\text{truematCRVS})} = \gamma^{(T+)} + \gamma^{(F-)}$ to be the probability of a maternal death among CRVS-registered deaths. For these countries, Eq. 1.2 still holds true but the relation between $\gamma^{(\text{matCRVS})}$ and $\rho^{(\text{truemat})}$ changes if $\omega^{(\text{truematUNREG})}$ differs from $\gamma^{(\text{truematCRVS})}$. In such settings, the relation between $\gamma^{(\text{matCRVS})}$ and $\rho^{(\text{truemat})}$ can be written as follows:

$$\gamma^{(\text{matCRVS})} = \frac{\lambda^{(+)} \cdot \rho^{(\text{truemat})}}{\rho^{(\text{CRVS})} + (1-\rho^{(\text{CRVS})}) \cdot \kappa} + (1-\lambda^{(-)}) \cdot \left(1 - \frac{\rho^{(\text{truemat})}}{\rho^{(\text{CRVS})} + (1-\rho^{(\text{CRVS})}) \cdot \kappa}\right), \tag{1.4}$$

where κ refers to the ratio of probabilities of a maternal death outside versus inside the CRVS:

$$\kappa = \frac{\omega^{(\text{truematUNREG})}}{\gamma^{(\text{truematCRVS})}}.$$

We aimed to estimate sensitivity, specificity, and κ (or a related parameter to summarize the relative difference between the probability of a maternal death outside versus inside the CRVS) for all country-years with CRVS data, such that CRVS data can be used to inform the estimation of maternal mortality among all deaths while accounting for CRVS misclassification errors and underregistration. However, given that data on the relative difference in maternal risk among CRVS-registered and unregistered deaths was so limited (see Table 1), we were unable to estimate this relative difference. Instead, we focused on the estimation of sensitivity and specificity using CRVS-based data only (221 observations, see Table 1). We developed a bivariate hierarchical random walk model for estimating sensitivity and specificity for all country-years, as explained in Section 1.4.2. We used all available CRVS-based data for model fitting, including data on the total number of maternal deaths only, as explained in Section 1.4.3.

1.4.2 Bivariate hierarchical random walk model for sensitivity and specificity

We developed a bivariate hierarchical random walk model to estimate sensitivity $\lambda_{c,t}^{(+)}$ and specificity $\lambda_{c,t}^{(-)}$ for all countries c with CRVS data for some year(s) t. We constrained sensitivity (se) to be within 0.1 and 1, and specificity (sp) to be within 0.95 and 1 using transformations:

$$\eta_{c,t}^{(+)} = log\left(\frac{\lambda_{c,t}^{(+)} - 0.1}{1 - \lambda_{c,t}^{(+)}}\right),$$
$$\eta_{c,t}^{(-)} = log\left(\frac{\lambda_{c,t}^{(-)} - 0.95}{1 - \lambda_{c,t}^{(-)}}\right).$$

Sensitivity and specificity (after transformation) were modeled using bivariate distributions to account for possible correlation between the two misclassification parameters. Accounting for this correlation is important for estimating misclassification parameters, i.e. see Chu et al. 2006. The model set-up used is a hierarchical random walk process. In reference year t_c, here chosen as the midyear of the country-specific observation period, we assume a hierarchical distribution for transformed sensitivity and specificity:

$$\begin{pmatrix} \eta_{c,t_c}^{(+)} \\ \eta_{c,t_c}^{(-)} \end{pmatrix} \sim N_2\left(\begin{bmatrix} \eta_{global}^{(+)} \\ \eta_{global}^{(-)} \end{bmatrix}, \begin{bmatrix} \sigma^{(+)^2} & \phi \cdot \sigma^{(+)} \cdot \sigma^{(-)} \\ \phi \cdot \sigma^{(+)} \cdot \sigma^{(-)} & \sigma^{(-)^2} \end{bmatrix}\right) \tag{1.5}$$

For years prior to the country-specific reference year, i.e. $t < t_c$:

$$\begin{pmatrix} \eta_{c,t}^{(+)} \\ \eta_{c,t}^{(-)} \end{pmatrix} \sim N_2\left(\begin{bmatrix} \eta_{c,t+1}^{(+)} \\ \eta_{c,t+1}^{(-)} \end{bmatrix}, \begin{bmatrix} \delta^{(+)^2} & \phi \cdot \delta^{(+)} \cdot \delta^{(-)} \\ \phi \cdot \delta^{(+)} \cdot \delta^{(-)} & \delta^{(-)^2} \end{bmatrix}\right). \tag{1.6}$$

For years after the country-specific reference year, i.e. $t > t_c$:

$$\begin{pmatrix} \eta_{c,t}^{(+)} \\ \eta_{c,t}^{(-)} \end{pmatrix} \sim N_2 \left(\begin{bmatrix} \eta_{c,t-1}^{(+)} \\ \eta_{c,t-1}^{(-)} \end{bmatrix}, \begin{bmatrix} \delta^{(+)^2} & \phi \cdot \delta^{(+)} \cdot \delta^{(-)} \\ \phi \cdot \delta^{(+)} \cdot \delta^{(-)} & \delta^{(-)^2} \end{bmatrix} \right).$$

The following prior distributions were assigned to the global mean parameters:

$$\begin{aligned} \lambda_{global}^{(+)} &\sim Unif(0.1, 1), \\ \lambda_{global}^{(-)} &\sim Unif(0.995, 1), \\ \eta_{global}^{(+)} &= log\left(\left(\lambda_{global}^{(+)} - 0.1\right)/(1 - \lambda_{global}^{(+)})\right), \\ \eta_{global}^{(-)} &= log\left(\left(\lambda_{global}^{(-)} - 0.1\right)/(1 - \lambda_{global}^{(-)})\right), \end{aligned}$$

Prior distributions for the correlation and standard deviations of the random walk were as follows:

$$\phi \sim Unif(-0.95, 0.95), \tag{1.7}$$

$$\sigma^{()} \sim N_{T(0,\infty)}(0, 1), \tag{1.8}$$

$$\delta^{()} \sim N_{T(0,\infty)}(0, 1), \tag{1.9}$$

where $N_{T(0,\infty)}(0,1)$ denotes a half-normal distribution (a truncated normal distribution with lower bound at 0).

We explored the use of indicators gross domestic product (GDP), the general fertility rate (GFR), the proportion of ill-defined causes, CRVS completeness, and ICD coding (ICD10 or earlier) as possible covariates to inform estimates of sensitivity and specificity. However, exploratory analyses suggested no substantially meaningful relations and were excluded from the final model. Illustrative plots are included in Appendix Section A.0.3.

1.4.3 Model fitting

Our goal is to estimate sensitivity and specificity using data from all country-years with CRVS-based specialized study data. Based on the assumption of a multinomial data generating process from Eq.1.1, we assumed the following data generating process for study counts from the ith study in country $c[i]$ in reference year $t[i]$:

$$\boldsymbol{z}_i | z_i^{(\text{CRVS})}, \boldsymbol{\gamma}_{c[i],t[i]} \sim \textit{Multinom}\left(z_i^{(\text{CRVS})}, \boldsymbol{\gamma}_{c[i],t[i]}\right), \tag{1.10}$$

with study counts $\boldsymbol{z}_i = \left(z_i^{(\text{T-})}, z_i^{(\text{T+})}, z_i^{(\text{F-})}, z_i^{(\text{F+})}\right)$, $z_i^{(\text{CRVS})} = \sum_{b \in B^{(\text{CRVS})}} z_i^{(\text{b})}$, and unknown probability vector $\boldsymbol{\gamma}_{c,t} = \left(\gamma_{c,t}^{(\text{T-})}, \gamma_{c,t}^{(\text{T+})}, \gamma_{c,t}^{(\text{F-})}, \gamma_{c,t}^{(\text{F+})}\right)$. For studies that refer to one calendar year, the study counts corresponds to the counts for that specific year, $z_i^{(\text{b})} = y_{c[i],t[i]}^{(\text{b})}$, while for studies that refer to multiple years, study counts are aggregates over the observation period, i.e., $z_i^{(\text{b})} = \sum_{t=t1[i]}^{t2[i]} y_{c,t}^{(\text{b})}$ where $t1[i]$ and $t2[i]$ refer to the start and end years of the ith study, respectively. The 4 CRVS-based probabilities $\gamma_{c,t}^{(\text{b})}$ can be written in terms of the two misclassification parameters $\lambda_{c,t}^{(+)}$ and $\lambda_{c,t}^{(-)}$, and the true CRVS-based probability of a maternal death as follows:

$$\begin{aligned}
\gamma_{c,t}^{(\text{T+})} &= \lambda_{c,t}^{(+)} \cdot \gamma_{c,t}^{(\text{truematCRVS})}, \\
\gamma_{c,t}^{(\text{F-})} &= \gamma_{c,t}^{(\text{truematCRVS})} - \rho_{c,t}^{(\text{T+})}, \\
\gamma_{c,t}^{(\text{T-})} &= \lambda_{c,t}^{(-)} \cdot \left(1 - \gamma_{c,t}^{(\text{truematCRVS})}\right), \\
\gamma_{c,t}^{(\text{F+})} &= \left(1 - \gamma_{c,t}^{(\text{truematCRVS})}\right) - \rho_{c,t}^{(\text{T-})}.
\end{aligned}$$

Country-year model parameters are defined through the bivariate hierarchical random model on $\lambda_{c,t}^{(+)}$ and $\lambda_{c,t}^{(-)}$, and vague independent priors on $\gamma_{c,t}^{(\text{truematCRVS})}$:

$$\gamma_{c,t}^{(\text{truematCRVS})} \sim U(0,1).$$

For studies that report on a specific set of non-overlapping categories, i.e. the number of false positive maternal deaths and/or the number of true positive maternal deaths, the corresponding likelihood function was obtained directly using the multinomial data generating process in Eq. 1.10.

However, the majority of studies only reported information on the number of true maternal deaths within the CRVS (see table 1.1). For each study that reported true maternal deaths within the CRVS, the study reported count of maternal deaths, $z_i^{(\text{truematCRVS})} = z_i^{(\text{T+})} + z_i^{(\text{F-})}$, overlaps with the CRVS-reported maternal deaths for the corresponding country-period, $z_i^{(\text{matCRVS})} = \sum_{t=t1[i]}^{t2[i]} y_{c[i],t[i]}^{(\text{T+})} + y_{c[i],t[i]}^{(\text{F+})}$. For each study period with information on overlapping categories, we obtained the exact likelihood function for the available death counts by summing over multinomial densities evaluated at each unique combination $\tilde{\boldsymbol{z}}_i = \left(\tilde{z}_i^{(\text{T-})}, \tilde{z}_i^{(\text{T+})}, \tilde{z}_i^{(\text{F-})}, \tilde{z}_i^{(\text{F+})}\right)$ that satisfied the observed set of counts. Specifically, for studies with information on the true number of maternal deaths $\left(z_i^{(\text{truematCRVS})}\right)$, and the number of maternal deaths observed in the CRVS $\left(z_i^{(\text{matCRVS})}\right)$, the likelihood function $f_i = f\left(z_i^{(\text{matCRVS})}, z_i^{(\text{truematCRVS})} | z_i^{(CRVS)}, \gamma_{c[i],t[i]}\right)$ is written as follows

$$f_i = \sum_{\tilde{z}_i^{(\text{T+})}=0}^{z_i^{(\text{matCRVS})}} p_z\left(\tilde{\boldsymbol{z}} | z_i^{(CRVS)}, \gamma_{c[i],t[i]}\right) \cdot 1\left(\tilde{z}_i^{(\text{T+})} + \tilde{z}_i^{(\text{F-})} = z_i^{(\text{truematCRVS})}\right) \cdot k_i$$

where $p_z\left(\tilde{\boldsymbol{z}} | z_i^{(\text{CRVS})}, \gamma_{c[i],t[i]}\right)$ refers to the multinomial density function for the 4 CRVS-based categories from Eq. 1.10. Additionally, to improve computational efficiency and remove combinations that result in values of sensitivity and specificity with negligible probabilities, we added additional constraints to possible combinations of $\tilde{\boldsymbol{z}}_i$, reflected in k_i with

$$k_i = 1\left(\tilde{z}_i^{(\text{T+})} \geq Bin_{2.5\%}\left(z_i^{(\text{truematCRVS})}, 0.1\right)\right) \cdot 1\left(\tilde{z}_i^{(\text{T-})} \geq Bin_{2.5\%}\left(z_i^{(\text{CRVS})} - z_i^{(\text{truematCRVS})}, 0.97\right)\right)$$

where $Bin_{2.5\%}(n,p)$ refers to the 2.5th percentiles of a Binomial distribution with sample size n and probability p, 0.1 is a lower bound for sensitivity, and 0.97 is a lower bound for specificity.

1.4.4 Computation

A Markov Chain Monte Carlo (MCMC) algorithm was employed to sample from the posterior distribution of the parameters with the use of the software *JAGS* (Plummer 2003). Ten parallel chains were run with a total of 40,000 iterations in each chain. Of these, the first of 10,000 iterations in each chain were discarded as burn-in and every 20th iteration after was retained. The resulting chains contained 1,500 samples each, with a total of 15,000 posterior samples. Standard diagnostic checks (using trace plots and Gelman and Rubin diagnostics (Gelman and Rubin 1992)) were used to check convergence.

1.4.5 CRVS adjustment factor

Based on estimates of sensitivity and specificity, for countries with complete CRVS systems, we defined the associated CRVS adjustment factor for country c in year t as follows:

$$\text{CRVSadj}_{c,t} = \frac{p_{c,t}^{(\text{truemat})}}{\lambda_{c,t}^{(+)} \cdot p_{c,t}^{(\text{truemat})} + \left(1-\lambda_{c,t}^{(-)}\right) \cdot \left(1-p_{c,t}^{(\text{truemat})}\right)}, \tag{1.11}$$

which varies with the true PM $p_{c,t}^{(\text{truemat})}$. For country-years without specialized studies, CRVS-adjustment factors follow from estimates of sensitivity and specificity, and the true PM.

1.4.6 Comparison to UN MMEIG 2015 approach

In the UN MMEIG 2015 approach, CRVS adjustment factors were obtained for all country-years with CRVS data and used directly in model fitting (Alkema et al. 2017). For countries with specialized studies, the CRVS adjustment in the UN MMEIG 2015 approach

was calculated for country-periods with studies by taking the ratio of the study-based observed proportion of maternal deaths to the observed CRVS-based proportion (Alkema et al. 2017). Linear interpolation was used to obtain adjustments in years in between observed adjustments. For forward extrapolation, the CRVS adjustment was kept constant at the level of the most recent observed CRVS adjustment. Backward extrapolations are explained below. The uncertainty of the adjustment was set equal to the variability associated with g, defined as follows:

$$\log(g)|G \sim N\left(\log(G), 0.25^2\right),$$

where G refers to the point estimate of the adjustment factor. For countries with CRVS data but no specialized studies, the UN MMEIG used a constant global adjustment factor of 1.5 for all country-years (Wilmoth et al. 2012, Alkema et al. 2017). For backward extrapolations in countries with studies, the CRVS adjustment was assumed to increase or decrease linearly to the same global adjustment factor of 1.5 in 5 years. The approach to obtaining CRVS adjustment with the CRVS-model differs from the UN MMEIG 2015 approach; the CRVS adjustment factor is obtained from estimates of sensitivity and specificity, and varies with the true PM, see Section 1.4.5.

1.4.7 Model validation

Model performance was assessed through two out-of-sample validation exercises. In the first exercise, 20% of the observations were left out at random to form a training data set. The process was repeated 20 times, i.e. 20 training sets were constructed with different samples left out in each set. In the second exercise, we left out the observation corresponding to the most recent study period in each country. In removing either 20% at random or the last observation, we assess how well the CRVS adjustment model performs in extrapolation of estimates within a given country. The CRVS adjustment model was fitted to each training set, and we obtained posterior samples for sensitivity and specificity in the

country-years with left-out specialized studies.

To validate model performance, we combined samples of sensitivity and specificity with information on study-based observed PMs to obtain samples of predicted CRVS-based PMs. We summarized the difference in terms of error, i.e., the difference between the observed CRVS-based PM and its point estimate, and coverage of 80% prediction intervals. We first summarize errors within countries, and the average across country specific measures to get overall predictive performance. The procedure is described in detail in Box 1.2.

Calculation of outcome measures in the validation exercise

1. Fit the CRVS adjustment model to the training data and obtain posterior samples $se_{c,t}^{(s)}$ and $sp_{c,t}^{(s)}$ for posterior samples $s = 1, 2, \ldots, S$ for country-years with left-out data in the test set.

2. Sample the CRVS-based reported number of maternal deaths using samples for sensitivity and specificity:

$$z_i^{(\text{matCRVS})^{(s)}} = z_i^{(\text{truematCRVS})} \cdot se_{c[i],t[i]}^{(s)} + \left(1 - z_i^{(\text{truematCRVS})}\right) \cdot \left(1 - sp_{c[i],t[i]}^{(s)}\right)$$

3. Calculate the difference between observed and estimated CRVS-based PM:

$$error_i^{matCRVS^{(s)}} = \left(z_i^{(\text{matCRVS})} - z_i^{(\text{matCRVS})^{(s)}}\right) / z_i^{(CRVS)}$$

   The median of the sampled errors is reported.

4. Calculate the proportion of CRVS-based PMs $z_i^{(\text{matCRVS})} / z_i^{(CRVS)}$ above and below their respective 80% prediction interval.

Figure 1.2: Overview of calculation of errors and coverage of prediction intervals in out-of-sample validation exercises.

1.5 Results

1.5.1 Validation Results

The CRVS adjustment model performs well in out-of-sample validation exercises, see Table 1.2. Median and relative errors are small in both exercises, and absolute errors are around 10% in predicting the CRVS-based PM. The model is well calibrated, the coverage

of the 80% prediction intervals is around 80%, with around 10% falling below (above) the lower (upper) bounds. Compared to the current UN-MMEIG approach, the CRVS adjustment model, median absolute error (MAE) showed improvement from 0.0009 to 0.0006, and 0.0010 to 0.0006, respectively.

Error in CRVS-PM								
Validation	**Model**	**# left-out obs**	**Median Errors**		**Relative Error** (%)		**outside** 80% **PI**	
			ME	MAE	MRE	MARE	% Below	% Above
Leave-out 20% at random	CRVSadj	43	0.00001	0.0006	0.5	9.9	0.11	0.11
	UN MMEIG 2015	43	-0.00010	0.0009	-1.8	15.9	0.08	0.05
Leave-out last observation	CRVSadj	20	0.0003	0.0006	2.0	10.8	0.10	0.10
	UN MMEIG 2015	20	-0.0003	0.0010	-4.0	14.4	0.10	0.10

Table 1.2: Validation results. The outcome measures are: median error (ME), median absolute error (MAE), relative error (MRE), absolute relative error (MARE), as well as the % of left-out observations below and above their respective 80% prediction intervals (PI) based on the training set.

1.5.2 Global findings

Table 1.3 lists the posterior estimates of the hyperparameters of the CRVS adjustment model. In the reference year, sensitivity is estimated at 0.586, 80% credible interval (CI) given by (0.511, 0.656), and specificity is 0.9993 (0.9990, 0.9996). The correlation between sensitivity and specificity was not significantly different from 0 (-0.095 $[-0.362, 0.183]$). There is substantial uncertainty associated with sensitivity and specificity in the reference year.

Figure 1.3 shows the relationship between true PM and the estimated CRVS adjustment factors, for specific values of specificity to illustrate their effect on the CRVS adjustment factor. When specificity equals one, the CRVS adjustment factor equals one over sensitivity, hence lower sensitivity results in a higher adjustment; conversely higher sensitivity results in a lower adjustment. When specificity is less than one, while keeping sensitivity fixed, the adjustment factor decreases with decreasing true PM. This effect is due to an in-

	10%	50%	90%
global sensitivity $\lambda^{(+)}_{global}$	0.511	0.586	0.656
global specificity $\lambda^{(-)}_{global}$	0.9990	0.9993	0.9996
correlation ϕ	-0.362	-0.095	0.183
sd sensitivity in t_c $\sigma^{(+)}_{tref}$	0.915	1.161	1.490
sd specificity in t_c $\sigma^{(-)}_{tref}$	0.871	1.293	1.842
sd sensitivity in RW $\delta^{(+)}$	0.161	0.201	0.255
sd specificity in RW $\delta^{(-)}$	0.508	0.673	0.857

Table 1.3: Posterior estimates of global parameters; median estimate (50%) and lower (10%) and upper (90%) bounds of 80% credible intervals.

creasing share of false positive maternal deaths among all deaths, and a decreasing share of false negative deaths, or, in other words, as the true PM decreases, the proportion of non-maternal deaths reported as maternal increases while the proportion of maternal deaths reported as non-maternal decreases. This relationship implies that keeping specificity and sensitivity constant in extrapolations will result in changing adjustment factors as the true PM changes. Specifically, the adjustment factor will decrease if the true PM decreases in forward projections. Similarly, when using a fixed value of sensitivity and specificity, the adjustment factor associated with these values will depend on the value of the true PM.

Moreover, small changes to values of specificity, with a given value of sensitivity, result in notable differences in CRVS adjustment factor. Shown in Figure 1.3, there are markable differences in resulting CRVS adjustment factor between specificity of 0.999 and 0.9999. This is due to small values of the proportion of maternal deaths.

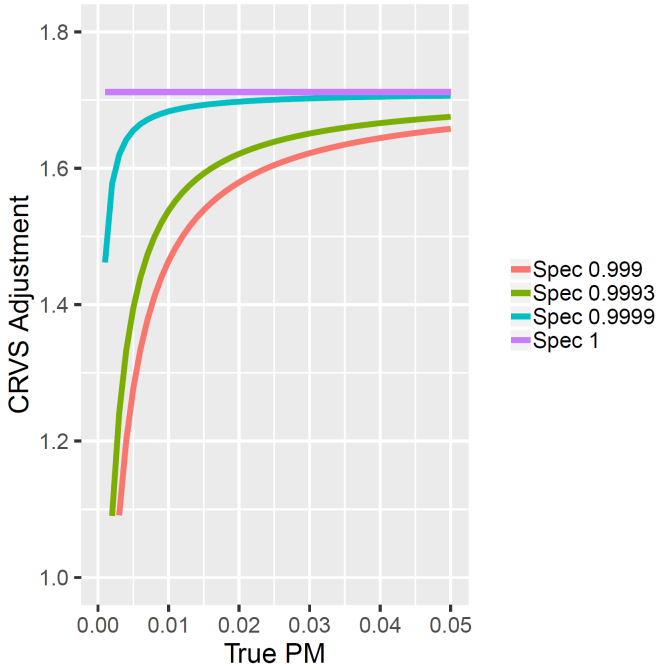

Figure 1.3: CRVS adjustment for different values of specificity, calculated at different levels of true PM when sensitivity is fixed at the global estimate of 0.586.

1.5.3 Country estimates

Sensitivity, specificity and CRVS adjustment estimates are shown for selected countries in Figure 1.4. Posterior estimates (blue) are shown with observed data (red) during the estimation period. Figure 1.4 illustrates how uncertainty in estimates of sensitivity and specificity depends on (i) what information is available, (ii) the number of deaths in the country, and (iii) the observation years. Most countries only have available data on true PM and CRVS-based PM across one or more periods. This is the case, for example, in Australia and the United Kingdom, in which we have observed true PM for multiple time periods. In these cases, sensitivity and specificity are unobserved, but are informed by observed data on true PM and CRVS-based PM. This results in larger uncertainty bounds for sensitivity and specificity estimates as compared to the same setting but with available information on breakdowns. An example country with breakdown information is Brazil, where sensitivity and specificity are recorded for recent years, and estimated with less uncertainty. In addition to availability of data, the number of deaths in the country also determines the

uncertainty in estimated sensitivity and specificity. For example, data in New Zealand is very uncertain due to the extremely small number of maternal deaths and total number of deaths to women of reproductive age. Uncertainty in sensitivity and specificity increases in years further away from years with data. This is illustrated in New Zealand, where data are available for recent years only; the uncertainty in sensitivity and specificity increases during periods without data.

Figure 1.5 shows 2017 estimates of sensitivity and specificity and associated uncertainty for all countries. In countries such as Austria, Denmark, New Zealand and Sweden, there is large uncertainty in sensitivity due to a very small number of maternal deaths. In contrast, in countries such as Brazil and Chile, there is an abundance of information on true maternal deaths and the breakdown of true maternal deaths, and therefore, uncertainty surrounding their estimates is reduced. Similarly, in the 2017 estimates of specificity, countries with information on the breakdown of maternal deaths show reduced levels of uncertainty surrounding their estimates. The United States and Thailand show much lower estimates of specificity when compared to other countries.

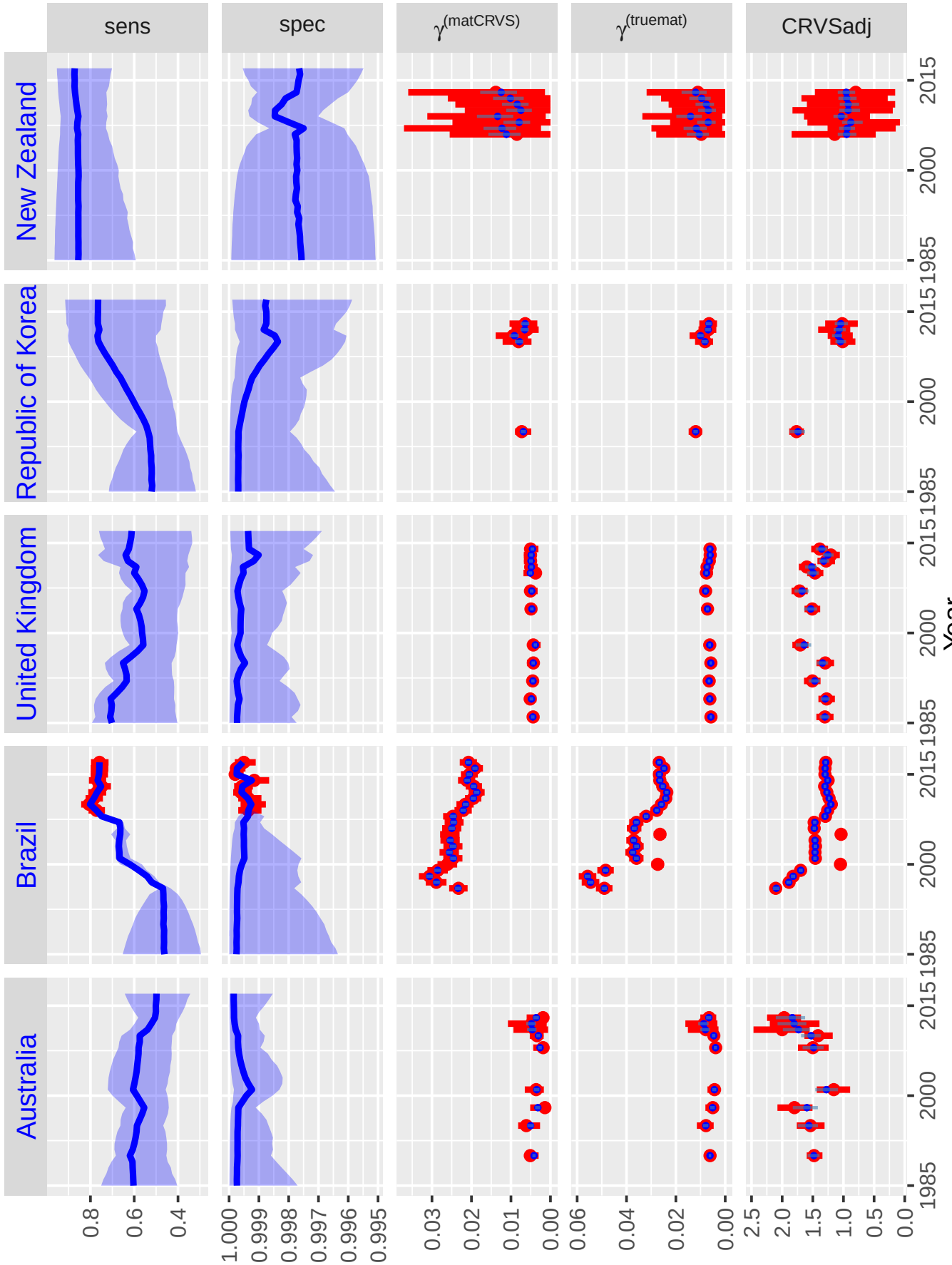

Figure 1.4: Illustration of CRVS adjustment model data and estimates for Australia, Brazil, the United Kingdom, Republic of Korea, and New Zealand. Parameters plotted consist of CRVS-based PM, true PM, sensitivity, specificity, and CRVS adjustment factors. The plots include: 1. observed data with associated observation-based 80% confidence intervals (red), 2. posterior estimates with 80% credible intervals (blue).

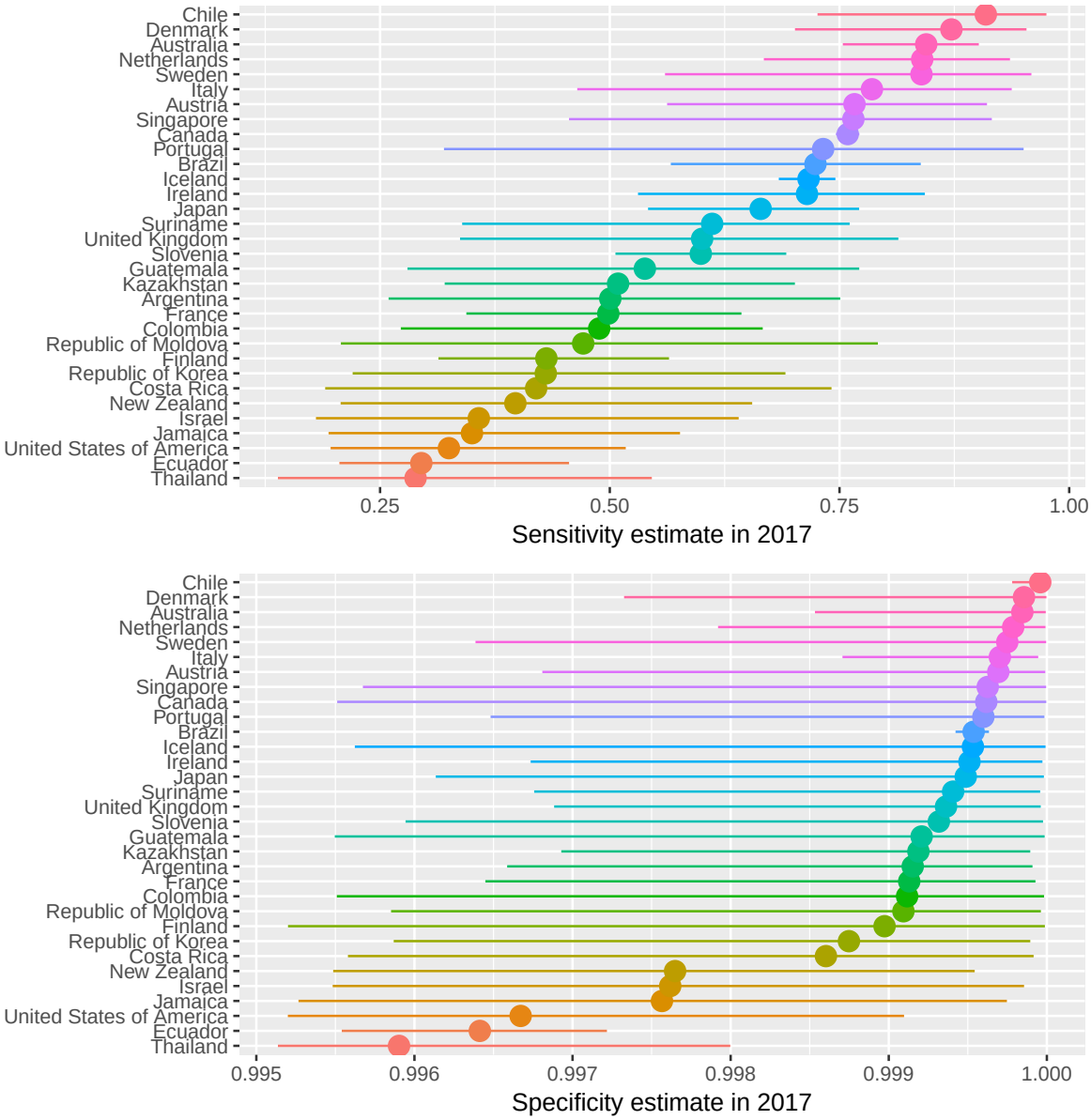

Figure 1.5: Ranked 2017 estimates of sensitivity and specificity across all countries with at least one specialized study.

1.6 Summary

In this paper, we presented a Bayesian hierarchical random walk model to assess maternal mortality misclassification errors in the CRVS with uncertainty. The model is based on the assessment of sensitivity and specificity of maternal mortality reporting, and captures differences therein between countries and within countries over time. Validation exercises

suggest that the model performs well in terms of predicting CRVS-based PM for country-periods without specialized studies.

The new model improves upon limitations of the 2015 UN MMEIG approach. In the UN MMEIG 2015 round of estimation, for countries with specialized studies that overlapped with CRVS data, adjustments were calculated directly from available data (i.e. the study's reported PM to CRVS-based PM) and kept constant in extrapolations. The rationale for keeping adjustments constant in the 2015 approach for countries with studies was to implement "no change in quality of reporting". However, when measuring quality of reporting in terms of sensitivity and specificity, the adjustment is not constant but varies with the true PM when keeping quality metrics constant, as illustrated in Figure 1.3. The CRVS model-based approach to obtaining adjustment factors improves upon this limitation of the UN MMEIG 2015 approach because its projections, which are based on constant sensitivity and specificity, are aligned with the assumption of constant quality of reporting. Finally, uncertainty assessments differ between the old and new approach. In the old approach, uncertainty in adjustments was assumed to be around 50% for all country-periods. In the new approach, uncertainty in the adjustment factor follows from the uncertainty in the estimates for sensitivity and specificity and resulting adjustments are more certain in settings with recent information about quality of reporting.

CHAPTER 2

ESTIMATING MATERNAL MORTALITY USING DATA FROM NATIONAL CIVIL REGISTRATION VITAL STATISTICS SYSTEMS: PRODUCING ESTIMATES OF SENSITIVITY AND SPECIFICITY FOR COUNTRY-YEARS WITHOUT VALIDATION DATA

2.1 Introduction

In Chapter I, we proposed a new modeling approach to estimate the extent of misclassification in the reporting of maternal mortality in civil registration vital statistics systems (CRVS). The indicator of interest is the proportion of all deaths to women of reproductive age that is maternal (PM). The CRVS misclassification model is a bivariate random walk process for sensitivity and specificity in reporting.

The CRVS model and its findings formed the direct basis for estimating the true proportion maternal deaths (true PM) from CRVS data in UN MMEIG estimates of maternal mortality (Peterson et al. 2019, UN MMEIG 2019). This chapter describes the additional steps needed to incorporate CRVS-model-based output into the UN MMEIG estimation approach. After review of the CRVS model in Section 2.2, we discuss how we incorporate CRVS model output in UN MMEIG estimation approach. Specifically, we discuss how to obtain point estimates and associated (co)-variances of sensitivity and specificity for countries with at least one country-period in which CRVS and validation data are available, and for countries with CRVS data and no validation studies. The main contribution of this chapter is the proposal of a new approach to obtain point estimates and associated (co)-variances of sensitivity and specificity for countries without validation studies. We introduce the approach in Section 2.3.

2.2 CRVS misclassification error model

We developed a bivariate hierarchical random walk model for estimating sensitivity and specificity for all country-years, described in detail in Peterson et al. (2019), see Chapter I. In summary, the random walk model is a non-stationary distribution that results in constant point estimates in backward and forward extrapolation.

Sensitivity and specificity (after transformation) were modeled using bivariate distributions to account for possible correlation between the two misclassification parameters.

The model set-up used is a hierarchical random walk process. In a reference year t_c, we assume a hierarchical distribution for transformed sensitivity and specificity:

$$\begin{pmatrix} \eta_{c,t_c}^{(+)} \\ \eta_{c,t_c}^{(-)} \end{pmatrix} \sim N_2 \left(\begin{bmatrix} \eta_{global}^{(+)} \\ \eta_{global}^{(-)} \end{bmatrix}, \begin{bmatrix} \sigma^{(+)^2} & \phi \cdot \sigma^{(+)} \cdot \sigma^{(-)} \\ \phi \cdot \sigma^{(+)} \cdot \sigma^{(-)} & \sigma^{(-)^2} \end{bmatrix} \right) \tag{2.1}$$

For years prior to the country-specific reference year , ie $t < t_c$:

$$\begin{pmatrix} \eta_{c,t}^{(+)} \\ \eta_{c,t}^{(-)} \end{pmatrix} \sim N_2 \left(\begin{bmatrix} \eta_{c,t+1}^{(+)} \\ \eta_{c,t+1}^{(-)} \end{bmatrix}, \begin{bmatrix} \delta^{(+)^2} & \phi \cdot \delta^{(+)} \cdot \delta^{(-)} \\ \phi \cdot \delta^{(+)} \cdot \delta^{(-)} & \delta^{(-)^2} \end{bmatrix} \right) \tag{2.2}$$

For years after the country-specific reference year , ie $t > t_c$:

$$\begin{pmatrix} \eta_{c,t}^{(+)} \\ \eta_{c,t}^{(-)} \end{pmatrix} \sim N_2 \left(\begin{bmatrix} \eta_{c,t-1}^{(+)} \\ \eta_{c,t-1}^{(-)} \end{bmatrix}, \begin{bmatrix} \delta^{(+)^2} & \phi \cdot \delta^{(+)} \cdot \delta^{(-)} \\ \phi \cdot \delta^{(+)} \cdot \delta^{(-)} & \delta^{(-)^2} \end{bmatrix} \right) \tag{2.3}$$

For forward extrapolation of estimates, after the country-specific reference period $i.e, t > t_c$, the random walk is a bivariate normal centered around the estimates of the preceding year.

Similarly, for backward extrapolation of estimates, before the country-specific reference period $t < t_c$, the estimates are modeled bivariate normal centered around the estimates of the proceeding year.

2.3 UN MMEIG 2019 approach to estimating maternal mortality from CRVS data

UN MMEIG uses a Bayesian model, referred to as BMat, to estimate the proportion of maternal deaths for all countries (Alkema et al. 2017). BMat combines a process model for the risk of a maternal death with data models that account for bias and uncertainty associated with available data. In summary (ignoring model specifications for handling HIV/AIDS maternal deaths), the process model for the log-transformed PM, $log(\gamma_{c,t}^{(\text{truemat})})$, combines a country specific intercept, a function of covariates, and an ARIMA(1,1,1) process. The BMat 2019 data model for CRVS data is given in Appendix A.0.4.1. In summary, we assumed that for country c and year t, the expected value of number of CRVS reported maternal deaths, $E_{c,t} = E(y_{c,t}^{(\text{matCRVS})} | \gamma_{c,t}^{(\text{truematCRVS})}, y_{c,t}^{(\text{CRVS})})$, is given by

$$E_{c,t} = y_{c,t}^{(\text{CRVS})} \cdot \left(\lambda_{c,t}^{(+)} \gamma_{c,t}^{(\text{truematCRVS})} + \left(1 - \lambda_{c,t}^{(-)}\right) \left(1 - \gamma_{c,t}^{(\text{truematCRVS})}\right) \right), \quad (2.4)$$

where, following notation from Chapter I, $y_{c,t}^{(\text{matCRVS})}$ refers to the number of maternal death as observed in CRVS in country c in year t, $\gamma_{c,t}^{(\text{truematCRVS})}$ is the true probability of a maternal death among all registered deaths, and $y_{c,t}^{(\text{CRVS})}$ is the total number of deaths registered in CRVS.

To incorporate CRVS misclassification parameters $\lambda_{c,t}^{(+)}$ and $\lambda_{c,t}^{(-)}$ into the estimation of maternal mortality in BMat, two approaches can be used. The first approach is jointly modeling misclassification parameters and parameters for the outcome of interest, i.e. incorporating the CRVS model into BMat to jointly estimate the true PM as well as sensitivity and

specificity of reporting using all available data. The second approach is to implement a sequential set-up, in which there is a stand alone model for misclassification parameters $\lambda_{c,t}^{(+)}$ and $\lambda_{c,t}^{(-)}$, which yields parameter estimates $\hat{\lambda}_{c,t}^{(+)}$ and $\hat{\lambda}_{c,t}^{(-)}$ that are used as fixed inputs into the larger model (Plummer, 2014). The Bayesian joint modeling approach correctly accounts for uncertainties in the data and allows simultaneous estimation of misclassification and outcome of interest. However, a disadvantage of this approach is that in settings with sparse data and potential model miss-specification, both misclassification and outcome of interest parameters may be poorly estimated (Bennett and Wakefield, 2001). In BMat 2019, the choice was made to use the sequential approach to avoid the updating of misclassification information in countries without information on data quality directly. This decision is outside the scope of this paper.

We implemented a sequential approach to estimating maternal mortality from CRVS data as follows: we first obtained point estimates as well as associated uncertainty of country-specific sensitivity and specificity using the CRVS model, as explained in the next section. Model fitting in BMat used this information in the form of point estimates $\hat{\lambda}_{c,t}^{(+)}$ and $\hat{\lambda}_{c,t}^{(-)}$ for country-year sensitivity and specificity, estimated variances $\hat{v}_{c,t}^{(+)}$ and $\hat{v}_{c,t}^{(-)}$, and $\hat{u}_{c,t}$, the estimated covariance between sensitivity and specificity in the same country-year (see Appendix A.0.4.1). Covariances across time in sensitivity and specificity were not incorporated, which is a model limitation.

2.3.1 Construction of estimates of misclassification parameters for countries with at least one specialized study

The estimates for sensitivity and specificity and associated outcomes need to be informed by all information available regarding misclassification in a country. For the (global) CRVS model as discussed in Chapter I, studies were used only if they provided exact information on death counts among deaths that were registered in the CRVS. Studies that reported only on the total number true maternal deaths in country-periods with incomplete CRVS sys-

tems, inclusive of missed maternal (U+) deaths, were excluded in the global assessment of misclassification because of lack of information on the relative difference between the true probability of a maternal death among registered versus unregistered deaths. In addition, studies that reported on partial calendar years were excluded. The exclusion decisions were made for the global model to avoid having to make additional assumptions that may affect the global estimates of misclassification. However, for constructing country-specific estimates, we aimed to include all available information, including data points that were excluded from the global model, if inclusion was possible based on reasonable assumptions.

To produce country-specific estimates, using all available data, we obtained country-specific fits of the CRVS model while keeping global parameters fixed at the estimates from the global CRVS model, referred to as a one-country model fit. For each country, all available studies were used, including studies that only provide information that includes missed maternal deaths (explained in Section 2.3.1.1), as well as studies that have partial overlap only with CRVS data. In the one-country model, all model parameters that do not vary across countries or by time are fixed at point estimates from the global CRVS model fit. The process model used for sensitivity and specificity equals the global process model otherwise.

In summary, estimates of misclassification parameters for countries with at least one specialized study are obtained as follows:

1. Fit CRVS model to global data base to obtain estimates of hyperparameters based on Eq. 2.1.

2. Fit CRVS model, described in Eqs. 2.1-2.3, to all data from country only, using estimates of hyperparameters, $\hat{\eta}^{(+)}_{global}, \hat{\eta}^{(-)}_{global}, \sigma^{(+)}, \sigma^{(-)}$ and ϕ from the global model fit in step 1.

2.3.1.1 Likelihood function for studies counting all maternal deaths in country-periods with incomplete CRVS

For a study that reported the total number of true maternal deaths, i.e. those within CRVS plus unregistered maternal deaths, the study-reported count of maternal deaths $z_i^{(\text{truemat})} = z_i^{(\text{T+})} + z_i^{(\text{F-})} + z_i^{(U+)}$ overlaps with CRVS-reported maternal deaths for the corresponding period. Similarly to studies that reported true maternal deaths inside CRVS, we obtain the exact likelihood function for available death counts by summing over multinomial densities evaluated at each combination $\tilde{\boldsymbol{z}}_i = \left(\tilde{z}_i^{(\text{T+})}, \tilde{z}_i^{(\text{T-})}, \tilde{z}_i^{(\text{F+})}, \tilde{z}_i^{(\text{F-})}, \tilde{z}_i^{(U+)}, \tilde{z}_i^{(\text{U-})}\right)$ that satisfied the observed counts. The likelihood function for i^{th} study (f_i) is written as follows:

$$f_i = \sum_{\tilde{z}_i^{(U+)}=0}^{z_i^{(UNREG)}} \sum_{\tilde{z}_i^{(\text{T+})}=0}^{z_i^{(\text{matCRVS})}} p_z\left(\tilde{\boldsymbol{z}}_{[1:4]} | z_i^{(CRVS)}, \boldsymbol{\gamma}_{c[i],t[i]}\right) \cdot 1\left(\tilde{z}_i^{(U+)} + \tilde{z}_i^{(\text{T+})} + \tilde{z}_i^{(\text{F-})} = z_i^{(\text{truemat})}\right) \cdot k_i \cdot h_i, \tag{2.5}$$

where $z_i^{(\text{UNREG})}$ refers to the number of unregistered deaths and $p_z\left(\tilde{\boldsymbol{z}}_{[1:4]} | z_i^{(\text{CRVS})}, \boldsymbol{\gamma}_{c[i],t[i]}\right)$ refers to the multinomial density function for the 4 CRVS-based categories from Eq. 1.10. To improve computational efficiency and remove combinations that result in values of sensitivity and specificity with negligible probabilities, we added constraints to possible combinations of $\tilde{\boldsymbol{z}}_i$, reflected in k_i with

$$k_i = 1\left(\tilde{z}_i^{(\text{T+})} \geq Bin_{2.5\%}(\tilde{z}_i^{(\text{truematCRVS})}, 0.1)\right) \cdot 1\left(\tilde{z}_i^{(\text{T-})} \geq Bin_{2.5\%}(z_i^{(CRVS)} - \tilde{z}_i^{(\text{truematCRVS})}, 0.97\right)$$

where $Bin_{2.5\%}(n, p)$ refers to the 2.5th percentile of a Binomial distribution with sample size n and probability p, 0.1 is a lower bound for sensitivity, and 0.97 is a lower bound for specificity. Lastly, we included combinations with expected ratios of the proportion maternal inside and outside the CRVS that vary between 0.5 and 2, reflected in h_i with

$$
\begin{aligned}
h_i &= 1\left(\tilde{z}_i^{(U+)} \geq Bin_{2.5\%}\left(z_i^{(\text{UNREG})}, p = 0.5 \cdot \frac{\tilde{z}_i^{(\text{truematCRVS})}}{z_i^{(\text{CRVS})}}\right)\right) \\
&\quad \cdot 1\left(\tilde{z}_i^{(U+)} \leq Bin_{97.5\%}\left(z_i^{(\text{UNREG})}, p = 2 \cdot \frac{\tilde{z}_i^{(\text{truematCRVS})}}{z_i^{(\text{CRVS})}}\right)\right).
\end{aligned}
$$

2.3.2 Construction of estimates of misclassification parameters for countries without specialized studies

We used global estimates of sensitivity, specificity and associated outcomes for all countries without specialized studies, obtained directly from fit of the CRVS model to the global data base, in the BMat data model in Eq A.1. Given the hierarchical set-up of the CRVS model (Eq. 2.1), the model can be used directly to produce a predictive distribution of sensitivity and specificity for countries without specialized study data in a reference year. The random walk model (Eqs. 2.2) is used for forward and backward extrapolations, and results in constant point estimates of sensitivity and specificity. Specifically, for a country c^* without specialized studies, we set point estimates for sensitivity and specificity equal to their respective global estimates from the global CRVS model fit, $\hat{\lambda}^{()}_{c^*,t} = \hat{\lambda}^{()}_{global}$.
However, in the bivariate random walk set-up, uncertainty in sensitivity and specificity is increasing as the time lag between the year of interest and the reference year increases, i.e. $Var(\lambda_{c^*,t_{ref}+l}) > Var(\lambda_{c^*,t_{ref}})$ for reference year t_{ref} and time lag $l > 0$. Lacking a natural choice of a reference year for countries without studies, we used constant estimates for the variance, and covariance terms, i.e. we set $\hat{v}^{()}_{c^*,t} = \hat{v}^{()}_{c^*,t_{ref}+l}$, $\hat{u}_{c^*,t} = \hat{u}_{c^*,t_{ref}+l}$ for all years t, fixed lag l and t_{ref} referring to the year where the hierarchical distribution of Eq. 2.1 applies. We used a validation exercise to determine the optimal value of time lag l, which resulted in the choice to use the uncertainty associated with the distribution of sensitivity and specificity in the reference year (Eq. 2.1).

2.4 Improving the estimation of (co-)variance terms for sensitivity and specificity for countries without validation studies: approximating the bivariate random walk with a vector autoregressive process

There are a number of limitations associated with the BMat 2019 approach in producing estimates of (co-)variance terms for sensitivity and specificity for countries without any validation data. Firstly, options considered were constrained to be based on the distribution associated with sensitivity and specificity in a year that is lag l away from the reference year. Secondly, temporal correlation is not assessed.

To overcome limitations of the BMat 2019 approach, we developed a new approach to incorporate results from the CRVS misclassification model into BMat for countries without misclassification data. In summary, we approximated the non-stationary random walk model, using fixed global mean estimates of sensitivity and specificity, with a stationary vector autoregressive process of order 1. The approximation approach for is summarized in Figure 2.1.

Sequential process to obtain global parameters related temporal correlation

1. Fit the global CRVS adjustment model to the global dataset of misclassification data and obtain posterior global estimates of transformed sensitivity and specificity $\hat{\eta}_{global} = (\hat{\eta}^{(+)}_{global}, \hat{\eta}^{(-)}_{global})$.
2. Use global estimates from step 1 as fixed inputs in a bivariate vector autoregressive VAR(1) set-up for log-transformed sensitivity and specificity. We approximate the non-stationary process in Step (1) with a stationary VAR(1) process.
3. Use global estimates of sensitivity, specificity and VAR(1) associated variance, covariance, and correlations in the larger BMat model.

Figure 2.1: Overview of sequential methods to obtain global estimates of sensitivity, specificity and related temporal variance and covariance parameters.

2.4.0.1 A vector autoregressive model for bivariate misclassification time series parameters

The vector autoregression (VAR) model is a flexible and natural extension of the univariate autoregressive model, which captures linear interdependencies among multiple time series (Lütkepohl, 2005). Each variable is written as a function based on its own lagged values, and lagged values of other model variables plus an error term.

The following section describes the analysis of the nonstationary time series of transformed sensitivity and specificity using a stationary VAR(1) framework that incorporates co-integration of relationships.

Following notation from Section 1.4.2, $\boldsymbol{\eta}_{c,t}$ denotes the (2×1) vector of time series transformed sensitivity and specificity. Let $\xi_{c,t}$ denote a zero-mean bivariate vector autoregressive (VAR) process of lag 1. This is an autoregressive structure in which an estimate at time t is solely dependent on the previous values at time $t-1$. The transformed sensitivity and specificity at time t is deterministically related to the sum of the fixed estimate of the global mean and a zero-mean VAR(1) process.

$$\eta_{c,t}^{(+)} = \hat{\eta}_{global}^{(+)} + \xi_{c,t}^{(+)}$$
$$\eta_{c,t}^{(-)} = \hat{\eta}_{global}^{(-)} + \xi_{c,t}^{(-)}$$

The zero-mean bivariate VAR(1) stochastic process is given by

$$\boldsymbol{\xi}_t = A\boldsymbol{\xi}_{t-1} + \boldsymbol{u}_t \tag{2.6}$$

in which $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is a 2×2 coefficient matrix representing the correlation between $\boldsymbol{\xi}_t$ and $\boldsymbol{\xi}_{t-1}$, and $\boldsymbol{u}_t$ is a 2×1 unobservable zero-mean innovation process (serially uncorrelated, ie. with time invariant covariance matrix Σ_u, and $E(u_t u_s^{'}) = 0$ for $s \neq t$.

Written more simply, we have

$$\xi_t^{(+)} = a_{11}\xi_{t-1}^{(+)} + a_{12}\xi_{t-1}^{(-)} + u_t^{(+)}$$
$$\xi_t^{(-)} = a_{21}\xi_{t-1}^{(+)} + a_{22}\xi_{t-1}^{(-)} + u_t^{(-)}$$
$$\Sigma_u = \begin{bmatrix} \delta^{(+)^2} & \psi \cdot \delta^{(+)} \cdot \delta^{(-)} \\ \psi \cdot \delta^{(+)} \cdot \delta^{(-)} & \delta^{(-)^2} \end{bmatrix}$$

in which $\xi_t^{(+)}$ and $\xi_t^{(-)}$ denote the zero-mean stochastic process for transformed sensitivity and specificity at time t, respectively. Based on derivations shown in Appendix B, vectors $\boldsymbol{\xi}_1, ..., \boldsymbol{\xi}_t$ are uniquely determined by $\boldsymbol{\xi}_0, \boldsymbol{u}_1, ..., \boldsymbol{u}_t$.

Vague prior distributions were assigned to the variance-covariance and correlation parameters for Σ_u. Autoregressive parameters a_{11} and a_{22} were given uniform(0,1) priors to enforce positive correlation across time specific estimates. The cross-correlation parameters a_{12} and a_{21} were set to zero such that we assume only the errors are correlated.

$$\delta^{(+)} \sim N_{T[0,]}(0,1)$$
$$\delta^{(-)} \sim N_{T[0,]}(0,1)$$
$$\psi \sim Unif(-1,1)$$
$$a_{11} \sim Unif(0,1)$$
$$a_{22} \sim Unif(0,1)$$
$$a_{11} = 0$$
$$a_{22} = 0$$

Appendix B shows the calculation of the unconditional expectation and variance. As such, the complete distribution of $\boldsymbol{\xi}_1, ..., \boldsymbol{\xi}_t$ can be written as the conditional distribution applied to the stationary time series for $t = 2, ..., T$, and the unconditional expectation and covariance at time $t = 1$.

$$\boldsymbol{\xi}_t | \boldsymbol{\xi}_{t-1} \sim N_2(A\boldsymbol{\xi}_{t-1}, \Sigma_u), \text{ for } t = 2, ..., T$$
$$\boldsymbol{\xi}_1 \sim N_2(\boldsymbol{\mu}, \Gamma_\xi(0))$$

In the zero-mean VAR(1) process, the unconditional expectation $\boldsymbol{\mu} = \mathbf{0}$, and the unconditional variance-covariance $\Gamma_\xi(0) = A\Gamma_\xi(0)A' + \Sigma_u$. Note that the unconditional variance $\Gamma_\xi(0)$ is written as a function of the coefficient matrix A and the global variance-covariance parameters in Σ_u. Refer to Appendix B for derivation of the unconditional mean and variance-covariance structure in detail.

Based on the bivariate VAR(1) framework, we obtain global variance-covariance estimates that incorporate temporal structure. Specifically, we obtain global estimates for; (1) autoregressive parameters for transformed sensitivity and specificity (a_{11}, a_{22}), (2) variance parameters $\delta^{(+)}, \delta^{(-)}$, and (3) global covariance-correlation between sensitivity and specificity $\delta^{(+-)} = \psi \cdot \delta^{(+)} \cdot \delta^{(-)}, \psi$.

2.5 Model validation

Model predictive performance is compared between the current approach of incorporating variance-covariance estimates from the random-walk reference year distribution versus the proposed VAR(1) approach, which accounts for temporal correlation. For each country in the global data set, we predicted country-year specific CRVS-based PMs using the random-walk reference year distribution, and using a VAR(1) process. We first summarize errors within countries to calculate country-specific mean errors, and then avaerage errors across

countries to report summary measures. As such countries are treated as independent units. The validation process is described in Box 2.2.

Calculation of outcome measures in the validation exercise
For each observation in the global data set:

1. Sample CRVS-based PMs using samples of sensitivity and specificity and observed true PM. Let i denote country-year with observed true PM for country c.

   - Transformed $se^{(s)}_{c[i],t[i]}$ and $sp^{(s)}_{c[i],t[i]}$ are drawn from unconditional stationary distributions from the reference year of the random-walk model in the first setting.

   $$\boldsymbol{\eta}^{(s)}_{c[i],t[i]} \sim N_2(\hat{\boldsymbol{\eta}}_{global}, \Sigma_{global})$$

   - Transform $se^{(s)}_{c[i],t[i]}$ and $sp^{(s)}_{c[i],t[i]}$ are drawn from stationary VAR(1) distribution in VAR(1) model in the second setting.

   $$\boldsymbol{\xi}^{(s)}_{c[i],t[i]} \sim N_2(\boldsymbol{A}\hat{\boldsymbol{\xi}}_{t[i]-1}, \Sigma_u)$$

   - Calculate samples of CRVS-based PM using estimates of sensitivity and specificity.

   $$z_i^{(\text{matCRVS})^{(s)}} = p_i^{(\text{truematCRVS})} \cdot se^{(s)}_{c[i],t[i]} + (1 - sp^{(s)}_{c[i],t[i]}) \cdot (1 - p_i^{(\text{truematCRVS})})$$

2. Calculate the mean error of CRVS-based PM for observation i, $Err_i = \sum_1^S p_i^{(\text{matCRVS})} - z_i^{(\text{matCRVS})^{(s)}}/S$.

3. Calculate the proportion of CRVS-based PMs above and below their respective 80% prediction interval.

4. Calculate the mean error of CRVS-based PM rate of change for observations $i-1$ to i.

$$\begin{aligned} Err_i &= \sum_1^S r_i^{(\text{truematCRVS})} - q_i^{(\text{truematCRVS})^s}/S \\ r_i^{(\text{matCRVS})} &= p_i^{(\text{matCRVS})} - p_{i-1}^{(\text{matCRVS})} \\ q_i^{(\text{matCRVS})^s} &= z_i^{(\text{matCRVS})^s} - z_{i-1}^{(\text{matCRVS})^s} \end{aligned}$$

5. Calculate the proportion of true PM rate of change above and below their respective 80% prediction interval.

6. Calculate the mean error and coverage across countries.

Figure 2.2: Overview of calculation of errors and coverage of prediction intervals in out-of-sample validation exercises.

2.5.1 Computation

A Markov Chain Monte Carlo (MCMC) algorithm was employed to sample from the posterior distribution of the parameters with the use of the software *JAGS* (Plummer 2003). Ten parallel chains were run with a total of 40,000 iterations in each chain. Of these, the first of 10,000 iterations in each chain were discarded as burn-in and every 20th iteration after was retained. The resulting chains contained 1,500 samples each, with a total of 15,000 posterior samples. Standard diagnostic checks (using trace plots and Gelman and Rubin diagnostics (Gelman and Rubin 1992)) were used to check convergence.

2.6 Results

2.6.1 Validation Results

Table 2.1 shows summary measure of bias across all countries. We compare the predictive performance of CRVS-based PM and the CRVS-based PM rate of change for both methods. In the case of predictive estimates of CRVS-based PM, the mean bias is comparable between the random-walk reference year and the VAR(1).

When we predict the difference in CRVS-based PM from observation $i-1$ to i, we account for temporal correlation by assessing the rate of change across observations. In this validation exercise, mean bias was slightly lower for the VAR(1) model, but differences were extremely small between methods. Coverage of 80% PIs was lower than the nominal 80% for the VAR(1) approach and higher for the RW reference year set-up.

Model	N	ME	SD Error	MAE	SD Absolute Error	% Below 80% PI	% Above 80% PI
	CRVS-based PM						
RW in reference year	37	-0.0001	0.0077	0.0014	0.0065	0.0601	0.1616
VAR(1)	37	-0.0004	0.0077	0.0013	0.0066	0.0096	0.1144
	CRVS-based PM Rate of Change						
RW in reference year	28	2.73×10^{-5}	0.0082	0.0004	0.0080	0.0273	0.0395
VAR(1)	28	1.56×10^{-5}	0.0083	0.0004	0.0081	0.1340	0.1842

Table 2.1: Summary measures of Mean Error, SD Error, Proportion below 80% PI, Proportion above 80% PI across countries. Measured for both true PM and the true PM rate of change between observations, summarized by the RW reference year and VAR(1) model set-ups.

2.6.2 Summary of global parameters

Table 2.2 lists the posterior estimates of the variance-covariance parameters for both RW and VAR(1) approaches. The correlation between sensitivity and specificity greatly differs between the two approaches, -0.15 and -0.09, respectively. There is substantial autoregression within sensitivity and specificity, $a_{11} = 0.985$ and $a_{22} = 0.911$, respectively. Lastly, uncertainty associated with sensitivity and specificity is larger in the VAR(1) framework.

	VAR(1)			RW reference year		
Parameter	10%	50%	90%	10%	50%	90%
correlation	-0.437	-0.15	0.13	-0.36	-0.09	0.18
autoregressive correlation (se) a_{11}	0.977	0.985	0.989			
autoregressive correlation (sp) a_{22}	0.85	0.911	0.950			
stationary sd sensitivity	0.211	0.258	0.310	0.16	0.20	0.25
stationary sd specificity	0.688	0.857	1.06	0.51	0.67	0.86

Table 2.2: Posterior estimates of global variance-covariance parameters; median estimate (50%) and lower (10%) and upper (90%) bounds of 80% credible intervals.

2.6.3 Assessment of bivariate distributional properties

Figure 2.3 shows the bivariate density distributions for transformed specificity against transformed specificity. At top the figure shows the bivariate normal distribution of transformed sensitivity and specificity, based on variance-covariance estimates of the reference year. At bottom the figure shows the bivariate normal distribution based on variance-covariance estimates obtained from the VAR(1) method. We see that the estimated correlation between sensitivity and specificity is more negative in the VAR(1) model fit. Additionally, Figure 2.4 illustrates bivariate normal distribution plots of sensitivity at time $t+1$ against sensitivity at time t (Left), and similarly specificity at time $t+1$ against specificity at time t (Right). Based on the VAR(1) model fits, we see there is high estimated autocorrelation within sensitivity and specificity.

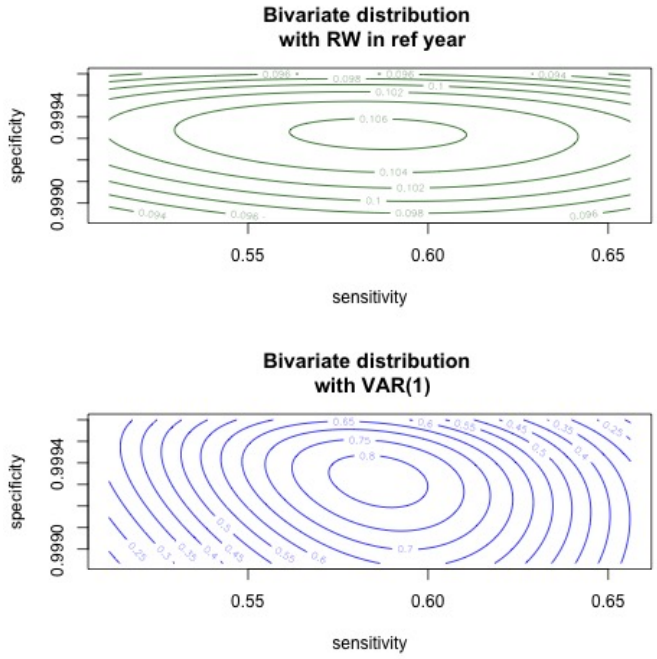

Figure 2.3: Bivariate density distribution plots. (Left) Bivariate normal distribution of transformed specificity and sensitivity, using uncertainties in the reference year, based on the RW method. (Right) Bivariate normal distribution of transformed specificity and sensitivity based on uncertainty estimates from the VAR(1) method.

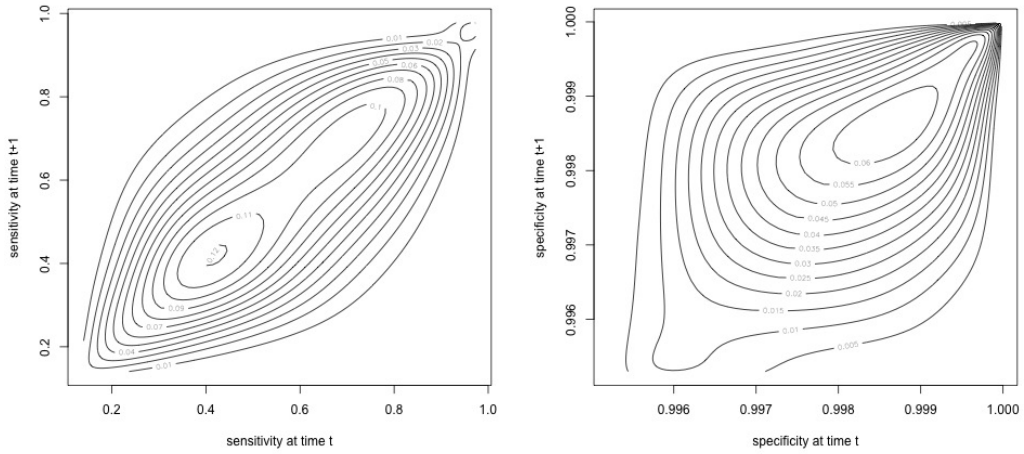

Figure 2.4: (A) Bivariate density distribution plot of $se_{t+1} \sim se_t$, (B) Bivariate density distribution plot of $sp_{t+1} \sim sp_t$.

2.7 Summary

In this chapter, we discussed how to use the CRVS adjustment model set-up and findings from chapter 1 to estimate maternal mortality in population-periods without validation studies in BMat, the model used for estimating maternal mortality by the UN MMEIG. We used a sequential set-up, in which point estimates and associated (co)variances of sensitivity and specificity of maternal mortality reporting are obtained first from the CRVS model, and then used as fixed inputs in BMat. Firstly, we developed an approach to obtain these estimates for countries with at least one specialized study, that may include maternal deaths not captured in the CRVS. Secondly, we developed two approaches to obtaining estimates of sensitivity and specificity for countries without validation data. The first approach is based directly on the random walk model. In the second approach, we approximated the bivariate random walk model with a vector autoregressive process. We summarized model performance using validation results of predicted CRVS-based PM and CRVS-based PM rate of change.

CHAPTER 3

A NEW PARAMETRIZATION OF REPORTING ERRORS IN SIBLING'S SURVIVAL HISTORIES FOR THE ESTIMATION OF AGE-GROUP SPECIFIC SURVIVAL PROBABILITIES

3.1 Introduction

Sibling survival history (SSH) data is a commonly used indirect method to obtain demographic information in countries with limited civil registration vital statistics (CRVS) systems (Graham et al. 1989). In countries with limited civil registration vital statistics (CRVS) systems, national estimates of adult mortality, for age groups 15-49 by sex, are derived from information obtained by respondents on the vital status, current age or age at death of their maternal siblings. Age-cohort specific adult mortality and survival rates are estimated directly using information on date of birth and date of death as reported by the respondents (Helleringer et al. 2014). However, due to respondent reporting errors, adult deaths may be incorrectly reported , i.e. a respondent may report the incorrect age of their sibling or incorrect age at death.

Reporting errors have potentially complex effects on SSH estimates of survival probabilities. Helleringer et al. (2014) classify reporting errors into four distinct types; 1) *List errors:* Respondents do not report a sibling, or respondents include non-maternal siblings, 2) *Vital status errors:* A respondent reports that one of her live siblings is dead or that a deceased sibling is alive, 3) *Age errors:* Inaccurate reporting of the current age or age at death, and 4) *Date errors:* Inaccurate reporting of year in which sibling(s) died. Additionally, in their assessment of the extent of reporting errors, they found there was substantial reporting errors in regards to misreporting of age and date errors, but less so in regards to

misreporting of vital status.

Health and demographic surveillance (HDSS) systems can be used to assess accuracy of SSH data by comparing reported age-cohort specific survival probabilities to those obtained from HDSS data. HDSS systems consist of monitoring over time an entire population located in a small geographic area (Pison 2005). By comparing SSH data to HDSS data, we identify reporting errors on an individual level to obtain information on siblings that have been misreported.

In this paper, we extend on the work presented by Helleringer et al. (2104), we present a new parametrization approach, which should be taken as an exploratory analysis in the absence of sufficient data to implement and validate. Vital status errors, as described by Helleringer, are reparametrized into misclassification metrics of sensitivity (true positive rate) and specificity (true negative rate) based on the parametrization used by Peterson et al. 2019, which accounts for misclassification of maternal deaths across country-periods without information on misclassification directly. Age at death errors that occur before the start of the reference period are parametrized into the probability of an omitted, and conversely, an added sibling, which is discussed further in Section 3.3. Birth year errors are parametrized into transition probabilities, which refer to the probability of a sibling being reported in age group j, given they are in age group i (Asmussen, 2003, Caswell, 2001, Leslie, 1945). Our final objective, in estimation of age at death and birth year reporting errors, is to propose a data generating mechanism to relate the true probability of survival related to female sibling survival history data, while accounting for misreporting bias. The objective in defining a data generating mechanism is to relate SSH error-prone data to the true probabilities given we have fixed estimates of misreporting parameters.

Section 3.2 briefly summarizes the preliminary data used in our exploratory analysis of reporting errors. Section 3.3 describes the different reporting error processes. Section 3.4

outlines our proposed parametrization to model misreporting errors by age at death and birth year, which is followed by Section 3.5 where we show preliminary graphical analysis based on the current limited data.

3.2 Data

Information on SSH misreporting errors was obtained from comparing information from health demographic surveillance systems to sibling survival history data. This section discusses both types of data.

3.2.1 Sibling survival histories

In countries with limited civil vital registration data, national estimates of adult mortality use information on a respondent's close relatives collected during census or surveys, referred to as the sibling survivorship method or sibling survival histories (SSH), and is considered an indirect estimation method for non-cause-specific mortality (Graham et al. 1989). SSH data include questions on the survival of a respondent's maternal siblings, i.e. siblings born to their biological mother, which include questions on each sibling's sex, survival status, and current age or age at death, as well as time elapsed since death if sibling is deceased (Helleringer et al., 2014). The current SSH survey sample consists of information reported on 1016 unique siblings via 410 respondents for one population in Malawi, extracted in 2018. These are a subset of siblings that have been linked to HDSS data using unique sibling identifiers.

3.2.2 Health and demographic surveillance data

Health and demographic surveillance systems (HDSS) consist of monitoring over time an entire population located in a small geographic area (Pison 2005). They include a baseline census, followed by continuous registration of demographic events (i.e. births, deaths,

marriages, migrations) affecting this population. Event registration happens yearly or more frequently, in which interviewers visit every household and ask for information on recent demographic events among household members (Helleringer et al, 2014). To link SSH reported siblings to HDSS records, HDSS information including names, sex, and resident of each sibling were used. We use HDSS data as gold standard data to assess the extent of misreporting of vital status errors and birth year errors among SSH data.

Table 1 gives information on the number of living siblings in the current 2018 Malawi SSH survey, broken down by those siblings alive at start of reference period (2013), and at time of survey (2018). Sibling information is given by data source, ie. HDSS and SSH reported information. Additionally, our outcome of interest is survival probabilities by age-cohort across different populations. Figure 3.1, illustrates the observed survival probabilities based on data from Malawi obtained in a SSH survey 2018 (red) versus the true survival probabilities (blue).

	SSH	HDSS
Number alive at start of survey period	$(No.)$	$(No.)$
15-19	66	65
20-24	84	85
25-29	95	93
30-34	119	111
35-39	132	122
40-44	133	124
45-49	76	69
Number alive at end of survey period	$(No.)$	$(No.)$
15-19	46	47
20-24	75	77
25-29	80	78
30-34	93	90
35-39	107	100
40-44	104	103
45-49	59	56
Number Deceased Siblings	208	205

Table 3.1: Characteristics of study respondents by HDSS and SSH reporting status. Number alive at start of survey period, is defined as total number of living siblings in 2013. Number alive at end of survey period, is defined as total number of living siblings at time of survey, 2018.

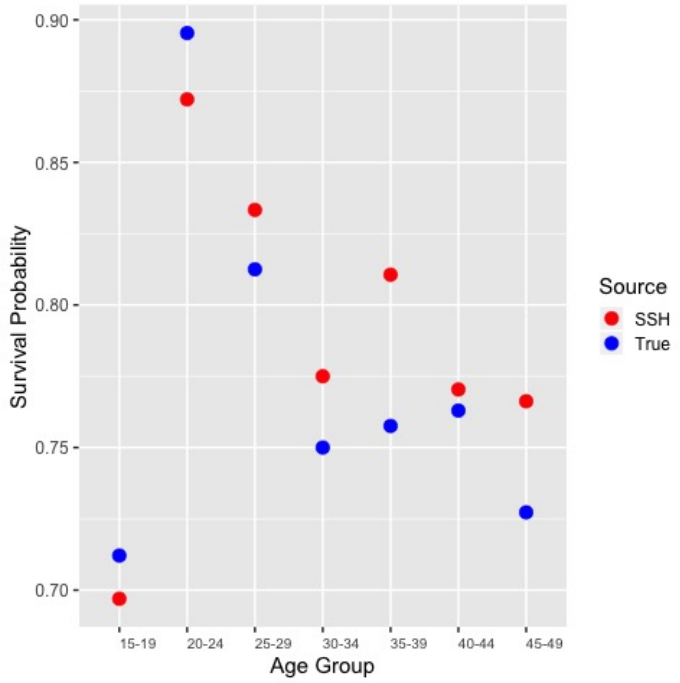

Figure 3.1: Observed survival probabilities by age cohort, comparing across data sources, for Malawi 2018 survey. (Red) indicates SSH reported survival probabilities, and (blue) indicates HDSS reported survival probabilities.

3.3 Reporting Errors

In Figure 3.1 we see a difference between SSH and HDSS age-cohort specific survival probabilities. The main objective is to parametrize reporting errors that result in these differences, such that we can relate SSH reported mortality data to true age-group specific survival probabilities. In doing so, we are able to learn information on the true survival probabilities given we only have error-prone SSH data across different population and age-groups.

Lexis diagrams (Figure 3.2) are commonly used devices to clarify relations between event/exposure segments for cohorts and event/exposure segments for periods. It is a two-dimensional figure in which age is one dimension and calendar year the other. (Preston 2001, Leslie 1945). We visualize cohort age-group specific mortality events using a Lexis diagram, in which diagonal lines display life lines of individual siblings. The reference

period is defined to be the respondent's interview date minus five years (t-5,t). Within this reference period, we calculate age group specific rates by 5-year age intervals from ages 15-49. Therefore, using the Lexis diagram, we can summarize the mortality experience of each female sibling, as well as the misreporting of age at death and birth year, within the given reference period, and by age-group. Figure 3.2, shown below, has illustrative examples of two siblings. Sibling (A) is in the 20-25 year age-group, and is a deceased sibling in the reference period. Sibling (B) is a sibling in the 25-30 year age-group, and is a living sibling at the end of the reference period.

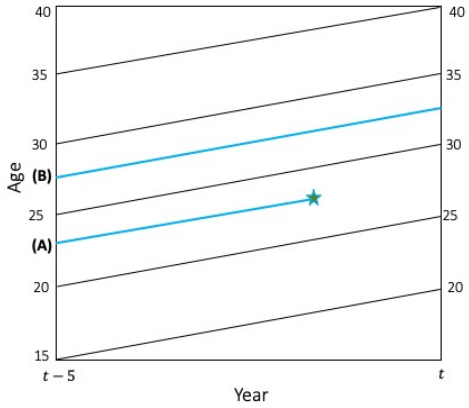

Figure 3.2: Illustrative Lexis diagram with reference period on the x-axis, and 5-year age groups on the y-axis. Sibling (A) is an example of deceased sibling in age group 20-25. Sibling (B) is an example of living sibling in age group 25-30.

3.3.1 Age at death errors

If the age at death is misreported, this changes the length of life lines, but does not influence the siblings reported age-group cohort. Therefore, the length of sibling life lines changes horizontally, but there is not adjustment vertically. Age at death reporting errors are broken down into two types of errors; (1) Age at death errors that occur before the start of the reference period, i.e. $<$ t-5, and (2) Age at death errors that occur during the reference period (t-5,t), which have been referred to as vital status errors previously by Helleringer et al. (2014).

(A) Errors after t-5 :

If a sibling is correctly identified as alive at the start of the reference period, (t-5), then age at death reporting errors occur during the reference period alone, ie (t-5,t). These errors can be broken down into false positive (F^+) and false negative (F^-) cases. In Figure 3.3 (A) we illustrate how false positive and false negative errors occur during the reference period. Sibling (A) was reported deceased within (t-5,t), but was classified alive at the end of the reference period t, which we label as a false positive, ie false death (F^+). Sibling (B) was reported alive at time t, but had died during the reference period (t-5,t), i.e. false living sibling (F^-). Lastly, sibling (C) died during the reference period, and was reported deceased, but age at death was misreported. However, if the sibling dies within the reference period, and is also reported deceased within the reference period, then no error occurs.

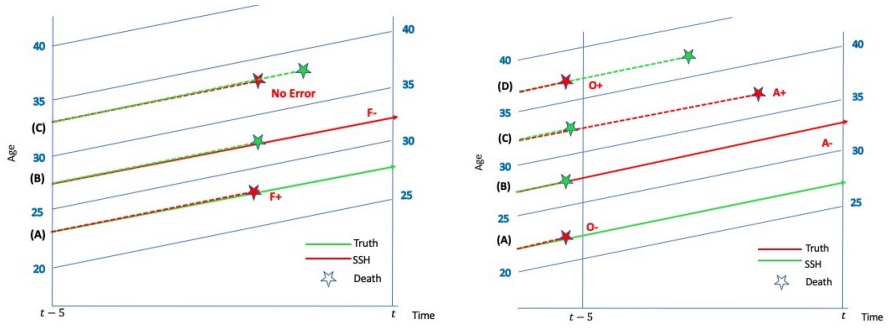

Figure 3.3: (A) Lexis diagram of vital status reporting error after t-5. (B) Lexis diagram of vital status reporting errors before t-5. Individual siblings are identified as (A),(B),(C),(D).

(B) Errors before t-5 :

At the start of the reference period, t-5, there are two misreporting errors that can occur; (1) A respondent incorrectly reports that a female sibling dies before the start of the reference period, which we refer to as an omitted sibling, and (2) A respondent incorrectly reports that a female sibling is alive at t-5 when the sibling is classified as deceased, which we refer to as an added sibling. In Figure 3.3 (B) we illustrate how the additions and omissions of siblings occur at t-5. Sibling (A) is an example of a sibling who was reported dead

before t-5, but was alive up to time t, labelled as omitted negative (O^-). Sibling (B) died before t-5, but was reported alive at time t, labelled as added negative (A^-). Sibling (C) was reported to have died past t-5, but had died before t-5, labelled as an added positive (A^+). Lastly, Sibling (D) was reported dead before t-5, but had died past t-5, labelled as omitted positive (O^+). To note, that the above definitions of omitted and added siblings are different from those used in previous work. Contrary to previous definitions used, we do not define omitted and added siblings to be those that are unmatched between HDSS and SSH data.

3.3.2 Birth Year

If the birth year is misreported, ie the sibling is moved to another age group, but vital status is correct, this changes the life line vertically, as shown in Figure 3.4 (A). For example, sibling (A) is a living sibling in the age group 20-25, but was reported to be in age group 25-30. Conversely, sibling (B) is a deceased sibling in age group 35-40, but was reported to be in age group 30-35.

To model birth year reporting errors, we account for the rate at which siblings have birth year misreported based on the degree of difference between SSH birth year and true birth year at time t-5, i.e. it is more common for a sibling to be reported in an age group directly above or below the true age compared to an age group with a large degree of difference. Figure 3.4 (B) shows a breakdown of siblings by true age group and SSH reported age group in a simplified example for 2 age groups $a = (1,2)$. There are siblings correctly classified in age group 1 (top left), and siblings in age group 1, incorrectly reported in age group 2 (top right). The sum of these boxes gives the total number of siblings that are truly in age group 1. Conversely, the total number of SSH reported siblings in age group 1 is obtained by summing those correctly classified (top left) and those in age-group 2, but reported in age group 1 (bottom left).

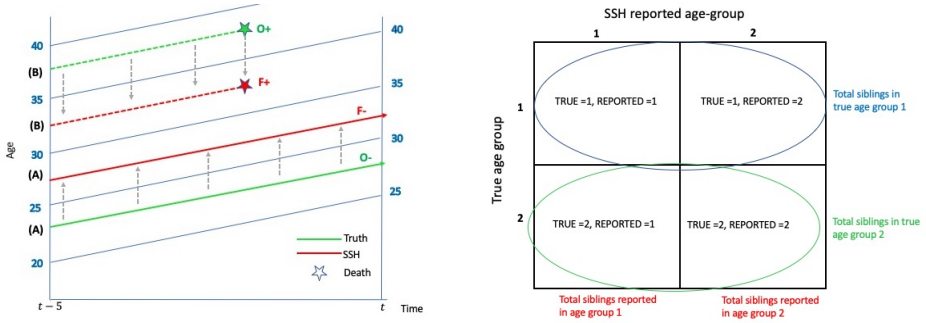

Figure 3.4: (A) Lexis diagram of birth year reporting errors. Siblings are moved vertically based on changes to their reported birth year, with vital status unchanged. (B) Illustrative example of breakdown of female siblings by true age group and SSH reported age group.

To capture the different rates at which siblings are transitioned, for each true age-group a, we estimate a vector of transition probabilities of SSH birth year reporting. The stochastic (transition) vector is defined by the probability of moving from position i to j, i.e $pr(j|i) = p_{i,j}$ (Asmussen, 2003). By definition, the sum of transition probabilities is equal to 1, $\sum_{j=1}^{J} p_{i,j} = 1$. Let $\delta_{a,\tilde{a}}$ denote the probability of a female sibling being reported in age group $\tilde{a}$, given the sibling is in true age-group a at time t-5. Given the simplified 2 age group example, let $y_{a,\tilde{a},t-5}$ refer to the total counts of living siblings corresponding to each cell, i.e. $y_{1,2,t-5}$ denotes the number of women reported in age group 2, in true age-group 1, at time t-5. The associated probability $\delta_{1,2}$ refers to the probability of being reported in age group 2, given the true age group is 1.

In summary, we parametrize birth year reporting errors using transition probabilities that relate the degree to which SSH reported age group is different from true age group at time t-5. We expect higher probabilities associated with lower degrees of difference between the true and SSH reported age-groups.

3.4 Methods

Our objective is to parametrize reporting errors related to SSH data and specify data generating mechanisms such that we relate the true cohort-specific probability of survival

for a given age-group a, denoted $\rho_a^{(\text{true})}$, to SSH reported total number of living siblings at times t-5 and t, denoted $S_{a,t-5}^{(\text{ssh})}, S_{a+1,t}^{(\text{ssh})}$, respectively. In the case where there is no reporting error, we can assume the data generating mechanism to be $S_{a+1,t}^{(\text{ssh})} | \rho_a^{(\text{true})}, S_{a,t-5}^{(\text{ssh})} \sim Binom(S_{a,t-5}^{(\text{ssh})}, \rho_a^{(\text{true})})$, in which the SSH reported living siblings accurately captures the true cohort survival probability $\rho_a^{(\text{true})}$. However, in the case where there is reporting errors, we parametrize the SSH associated survival probability $\pi_a^{(\text{ssh})}$ taking into account reporting errors related to age at death and birth year misreporting.

In Section 3.4.1 we first describe how we estimate age at death misreporting parameters (se_a, sp_a) absent of birth year reporting errors, and secondly, how we estimate birth year reporting errors, absent of age at death errors. Lastly, we account for the interaction between age at death and birth year reporting errors. We define the data generating process mechanism that accounts for these errors and the function in which we relate the SSH associated survival probability $\pi_a^{(\text{ssh})}$ to true survival probability $\rho_a^{(true)}$ and misreporting errors.

3.4.1 Breakdown of misreporting parameters for age at death

First, if we ignore birth year misreporting momentarily, we parametrize age at death errors. We first parametrize errors related to misclassification, sensitivity and specificity, among siblings reported by both HDSS and SSH data, i.e. Siblings that are correctly reported alive at the start of the reference period t-5. Subsequently, we extend to parametrize probabilities related to omitted siblings, and lastly to added siblings, i.e. Siblings that are misreported at time t-5. This is characterized in a 3 step process described below.

Step 1: We consider vital status reporting errors within women reported living at t-5 in both SSH and HDSS data. Specifically, we define vital status reporting errors across a 4-box model shown in Figure 3.5. Let $\tau_a^{(c)}$ denote the probability of a female sibling in category c given the sibling has been reported alive at t-5 by both sources, i.e. $\tau_a^{(\text{F+})}$ denotes the probability of a false positive death out of the 4 boxes. We assume a multinomial data

generating model for the 4-box model given the siblings are reported living at t-5 by both sources. As such $\tau_a^{(T^+)} + \tau_a^{(T^-)} + \tau_a^{(F^+)} + \tau_a^{(F^-)} = 1$. Misclassification metrics, sensitivity (se_a) and specificity (sp_a), for given age group a, are defined as:

$$se_a = \frac{\tau_a^{(T^+)}}{\tau_a^{(T^+)} + \tau_a^{(F^-)}} \qquad (3.1)$$
$$se_a = \frac{\tau_a^{(T^-)}}{\tau_a^{(T^-)} + \tau_a^{(F^+)}}$$

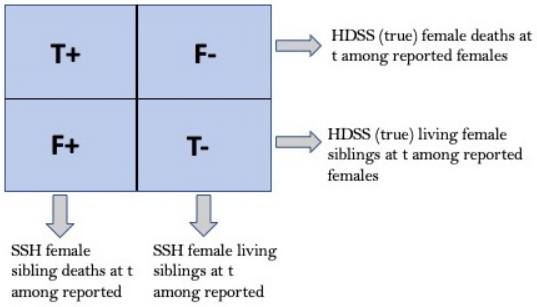

Figure 3.5: Diagram of age-group specific 4-box breakdown of female siblings reported alive at $t-5$ in both HDSS and SSH data.

Given definitions sensitivity and specificity in Eq. 3.1, we can derive the SSH-associated survival probability, for age-group a, among reported siblings, $\tau_a^{(\text{ssh})} = \tau_a^{(\text{T-})} + \tau_a^{(\text{F-})}$. We write the SSH-associated survival probability as a function of sensitivity, specificity, and the true survival probability for age group a among siblings living at t-5 in both sources, $\tau_a^{(\text{true})} = \tau_a^{(T^-)} + \tau_a^{(F^+)}$.

$$\tau_a^{(\text{ssh})} = sp_a \cdot \tau_a^{(\text{true})} + (1 - se_a) \cdot \left(1 - \tau_a^{(\text{true})}\right) \qquad (3.2)$$

Based on the 4 box model, we are able to estimate the true survival probability within women reported living at t-5 in both SSH and HDSS data, ie within the 4 boxes, using a binomial data generating assumption. Let $s*_{a,t-5}^{(\text{ssh})}$ refer to SSH reported total number of livings siblings among the 4 boxes alone, for age group a at time t-5.

$$s*_{a+1,t}^{(\text{ssh})} \,|\, \tau_a^{(\text{ssh})}, s*_{a,t-5}^{(\text{ssh})} \sim Binom(s*_{a,t-5}^{(\text{ssh})}, \tau_a^{(\text{true})}) \qquad (3.3)$$
$$s*_{a,t-5}^{(\text{ssh})} = y_{a,t-5}^{(\text{T-})} + y_{a,t-5}^{(\text{F-})} \qquad (3.4)$$

Step 2: In Eq. 3.2, we derive the SSH-associated survival probability, for age group a, as a function of the true survival probability, among those siblings reported in the 4 box figure.

However, to derive the true survival probability, for age group a, we extend the 4-box model to 6-boxes in Figure 3.6, which includes omitted female siblings, i.e. living siblings that have been reported to have died at time t-5. The total number of true living siblings at time t-5 is given by the sum of the 6 boxes. We assume a multinomial data generating process for the individual counts,

$$
\begin{aligned}
\boldsymbol{y}_{a,t-5} | \boldsymbol{\rho}_a, S^{(\text{true})}_{a,t-5} &\sim Multinom(S^{(\text{true})}_{a,t-5}, \boldsymbol{\rho}_a) \\
\boldsymbol{y}_{a,t-5} &= (y^{(\text{T+})}_{a,t-5}, y^{(\text{T-})}_{a,t-5}, y^{(\text{F+})}_{a,t-5}, y^{(\text{F-})}_{a,t-5}, y^{(O+)}_{a,t-5}, y^{(O-)}_{a,t-5}) \\
\boldsymbol{\rho}_a &= (\rho^{(\text{T+})}_a, \rho^{(\text{T-})}_a, \rho^{(\text{F+})}_a, \rho^{(\text{F-})}_a, \rho^{(O+)}_a, \rho^{(O-)}_a) \\
S^{(true)}_{a,t-5} &= y^{(\text{T+})}_{a,t-5} + y^{(\text{T-})}_{a,t-5} + y^{(\text{F+})}_{a,t-5} + y^{(\text{F-})}_{a,t-5} + y^{(O+)}_{a,t-5} + y^{(O-)}_{a,t-5} \\
\sum_{c \in C} \rho^{(c)}_a &= 1, \text{ for } C = \{T+, T-, F+, F-, O+, O-\}
\end{aligned}
\tag{3.5}
$$

The associated box probabilities $\rho^{(c)}_a$, for category c, refer to the probability of a sibling in category c out of $S^{(true)}_{a,t-5}$ (the sum of the 6 boxes). The true survival probability, for age group a, is given by Eq. 3.6.

$$
\rho^{(true)}_a = \rho^{(\text{T-})}_a + \rho^{(\text{F+})}_a + \rho^{(O-)}_a \tag{3.6}
$$

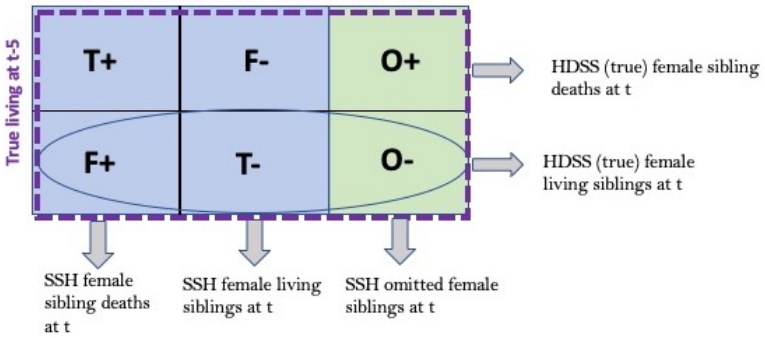

Figure 3.6: Diagram of 6-box multinomial breakdown of vital status error and omitted siblings for true living siblings at time $t-5$ in age-group a.

To relate the true probability of survival, for age group a, to Eq. 3.2, we introduce parameters related to the probability of omitted siblings. Specifically, let $\rho_a^{(\text{omit})} = \rho_a^{(O+)} + \rho^{(O-)}$ refer to the probability of an omitted sibling. Additionally, let $\kappa_a^{(O-)} = \frac{\rho_a^{(O-)}}{\rho_a^{(O-)}+\rho_a^{(O+)}}$ refer to the probability of being an omitted living sibling out of all omitted siblings.

$$\rho_a^{(\text{true})} = \tau_a^{(\text{true})} \cdot (1 - \rho_a^{(\text{omit})}) + \rho_a^{(\text{omit})} \cdot \kappa_a^{(O-)} \tag{3.7}$$

Conversely, using Eq.3.7, we write the true survival probability among siblings living at t-5 in both sources, $\tau_a^{(\text{true})}$, in terms of $\rho_a^{(\text{true})}$, and probabilities associated with omitted siblings, $\rho_a^{(\text{omit})}$, $\kappa_a^{(O-)}$, to be used in step 3.

$$\tau_a^{(\text{true})} = \frac{\rho_a^{(\text{true})} - \rho_a^{(\text{omit})} \kappa_a^{(O-)}}{1 - \rho_a^{(\text{omit})}} \tag{3.8}$$

Step 3: In Steps 1 and 2 we parametrized vital status reporting errors in terms of sensitivity and specificity, among siblings reported living at t-5 by both SSH and HDSS, and then extended the model to account for omitted living siblings in the true probability of survival. The last step is to relate the true probability of survival, for age group a, to the SSH-associated probability of survival by accounting for added living siblings in the SSH-reported survival probabilities.

Figure 3.7 shows the breakdown of living female siblings, at time t-5 based on SSH reported vital status (left) and true vital status (right). The SSH-reported total of living female siblings at t-5, $S_{a,t-5}^{(\text{ssh})}$, is the sum of the 2 columns shown on the left. The SSH reported survival probability is given by $\pi_a^{(\text{ssh})} = \pi_a^{(\text{T-})} + \pi_a^{(\text{F-})} + \pi_a^{(A-)}$, in which $\pi_a^{(c)}$ refers to the probability of a sibling being in category c out of $S_{a,t-5}^{(\text{ssh})}$ (the 6 boxes on left) . Conversely, the true total of living siblings at t-5, $S_{a,t-5}^{(true)}$, is the sum of the 2 rows shown on the right. The true survival probability is given by $\rho_a^{(true)} = \rho_a^{(\text{T-})} + \rho_a^{(\text{F+})} + \rho_a^{(O-)}$, in which $\rho_a^{(c)}$ refers to the probability of a sibling being in category c out of $S_{a,t-5}^{(true)}$.

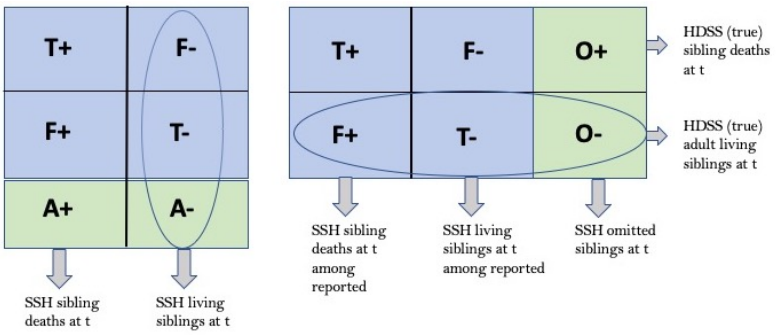

Figure 3.7: Diagram of true versus SSH multinomial breakdown of vital status error and omitted/added siblings.

To relate the true survival probability to the SSH reported survival probability, we incorporate parameters related to added siblings as follows; let $\pi^{(add)} = \pi_a^{(A+)} + \pi_a^{(A-)}$ denote the probability of being an added sibling among $S_{a,t-5}^{(\text{ssh})}$. Additionally, let $\kappa^{(A-)} = \frac{\pi_a^{(A-)}}{\pi_a^{(A+)} + \pi_a^{(A-)}}$ denote the probability of added living siblings among all added siblings. Lastly, $D_{a,t-5}^{(true)}$ denotes the true number of deceased siblings in age-group a at time t-5, and $\alpha_a^{(add)}$ refers to the probability of an added sibling out of $D_{a,t-5}^{(true)}$. $D_{a,t-5}^{(true)}$ is observed, given HDSS data is available, in which the true number of deceased siblings is the total number of siblings born in the original cohort minus those truly living at time t-5, $D_{a,t-5}^{(true)} = B_{a,t-5}^{(true)} - S_{a,t-5}^{(true)}$. We reparametrize the probability of an added sibling using previously defined parameters related to the omitted siblings $\rho_a^{(\text{omit})}$.

$$\pi_a^{(add)} = \frac{\alpha_a^{(add)} \cdot D_{a,t-5}^{(true)}}{(1-\rho_a^{(\text{omit})}) S_{a,t-5}^{(true)} + \alpha_a^{(add)} \cdot D_{a,t-5}^{(true)}} \tag{3.9}$$

Our goal is to estimate the true survival probability, $\rho_a^{(true)}$ based on $S_{a,t-5}^{(\text{ssh})}$ and $S_{a+1,t}^{(\text{ssh})}$, while accounting for reporting errors. Based on Eqs. 3.1 - 3.9, we relate the age-group specific true survival probability $\rho_a^{(true)}$ to the SSH associated survival probability $\pi_a^{(\text{ssh})}$ using misclassification parameters (se_a, sp_a), and parameters related to added and omitted deaths $\rho_a^{(\text{omit})}, \kappa_a^{(O-)}, \kappa_a^{(A-)}$, and $\alpha_a^{(add)}$, shown in Eq. 3.10

$$\pi_a^{(\text{ssh})} = \left(\frac{(1-\rho_a^{(omit)}) \cdot S_{a,t-5}^{(true)}}{(1-\rho_a^{(omit)}) \cdot S_{a,t-5}^{(true)} + \alpha_a^{(add)} \cdot D_{a,t-5}^{(true)}} \right) \cdot \tag{3.10}$$
$$\left(\frac{\rho_a^{(\text{true})} - \rho_a^{(\text{omit})} \kappa_a^{(O-)}}{(1-\rho_a^{(\text{omit})})} \cdot sp_a + \left(1 - \frac{\rho_a^{(\text{true})} - \rho_a^{(\text{omit})} \kappa_a^{(O-)}}{(1-\rho_a^{(\text{omit})})} \right) (1-se_a) \right) +$$
$$\frac{\alpha_a^{(add)} \cdot D_{a-1,t-5}^{(true)} \cdot \kappa_a^{(A-)}}{(1-\rho_a^{(\text{omit})}) S_{a,t-5}^{(true)} + \alpha_a^{(add)} D_{a,t-5}^{(true)}}$$

In the following section, we parametrize birth year reporting errors, in the absence of age at death errors, and subsequently propose a method to incorporate the combination of age at death errors and birth year reporting errors together.

3.4.2 Breakdown of birth year misreporting

In Section 3.4.1, we parametrized age at death errors ignoring birth year reporting errors. For the assessment of birth year reporting errors, we first parametrize these errors in the absence of age at death errors. Birth year reporting errors are parametrized into transition probabilities, defined as the probability of being reported in age-group $\tilde{a}$ given the sibling's true age group a. This parametrization captures the rate at which siblings are misreported from one age-group to another. A simplified example was shown in Figure 3.4 for 2 age-groups. We generalize to accommodate 5-year age groups between 15-49 years of age. Let $\boldsymbol{y}_{a,t-5} = (y_{a,\tilde{a}_{min},t-5}, ..., y_{a,\tilde{a}_{max},t-5})$ refer to the vector of counts of siblings that are reported in age groups $\tilde{a} = (1,2,...,7)$, given the true age group a, at time t-5. The corresponding vector of transition probabilities $\boldsymbol{\delta}_a = (\delta_{a,\tilde{a}_{min}}, ..., \delta_{a,\tilde{a}_{max}})$, refers to the probabilities of a sibling being reported in age group $\tilde{a}$ given the sibling is in true age group a at time t-5. Based on this parametrization, we assume a multinomial data generating process for each true age group, in which,

$$
\begin{aligned}
\boldsymbol{y}_{a,t-5}|\boldsymbol{\delta}_a, S^{(true)}_{a,t-5} &\sim Multinom(S^{(true)}_{a,t-5}, \boldsymbol{\delta}_a) \\
\boldsymbol{y}_{a,t-5} &= (y_{a,\tilde{a}_{min},t-5}, \ldots, y_{a,\tilde{a}_{max},t-5}) \\
\boldsymbol{\delta}_a &= (\delta_{a,\tilde{a}_{min}}, \ldots, \delta_{a,\tilde{a}_{max}}) \\
y_{a,\tilde{a},t-5}|\delta_{a,\tilde{a}}, S^{(true)}_{a,t-5} &\sim Binom(S^{(true)}_{a,t-5}, \delta_{a,\tilde{a}}) \\
S^{(true)}_{a,t-5} &= \sum_{\tilde{a}} y_{a,\tilde{a},t-5} \\
S^{(\text{ssh})}_{a,t-5} &= \sum_{a} y_{a,\tilde{a},t-5} \\
S^{(\text{ssh})}_{a+1,t} &= \sum_{a} y_{a+1,\tilde{a}+1,t} \\
\sum_{\tilde{a}} \delta_{a,\tilde{a}} &= 1
\end{aligned}
\tag{3.11}
$$

(3.12)

3.4.3 Combining birth year and age at death errors

In the previous sections, Section 3.4.1 and Section 3.4.2, we describe misreporting parameters based on age at death and birth year independently. However, we must extend the above parametrizations to account for the combination of birth year and age at death reporting errors. Figure 3.8 demonstrates the case in which a sibling can have a combination of birth year and age at death error. For example, a sibling in true age group $a = 1$, may be reported in age group $\tilde{a} = 2$. Additionally, the sibling may have their vital status reported as deceased, within the reference period, given they are classified as a living sibling at time t, i.e. they are a true living sibling at the end of the reference period. This results in being an omitted living sibling in age group a, and a

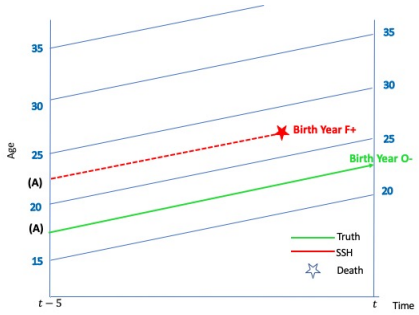

Figure 3.8: Illustrative example of the combination of age at death and birth year misreporting. Age at death errors given birth year errors are distinguished from previous age at death errors alone, using notation Birth year $F+$, Birth year $O-$.

false positive death in age group $\tilde{a}$. However, these omitted and false positive siblings have a different misreporting mechanism than those due to age at death errors alone. Therefore, within each combination $(a,\tilde{a})$, we have a multinomial breakdown of age at death reporting, based on the 6 box model in Eq. 3.5, given birth year reporting error. Using the simplistic example, if we look at true age-group $a = 1$, SSH age-group $\tilde{a} = 2$, we have multinomial counts

$$\boldsymbol{y}_{1,2,t-5} = (y_{1,2}^{(\text{T+})}, y_{1,2}^{(\text{T-})}, y_{1,2}^{(\text{F+})}, y_{1,2}^{(\text{F-})}, y_{1,2}^{(O+)}, y_{1,2}^{(O-)}, y_{1,2}^{(A+)}, y_{1,2}^{(A-)})$$

More generally, to account for the interaction between birth year and age at death misreporting, we assess the extent of age at death errors dependent on the size of birth year errors.

$$\begin{aligned}
\boldsymbol{y}_{a,\tilde{a},t-5} | \boldsymbol{\rho}_{a,\tilde{a}}, y_{a,\tilde{a},t-5} &\sim Multinom(y_{a,\tilde{a},t-5}, \boldsymbol{\rho}_{a,\tilde{a}}) \qquad (3.13)\\
y_{a,\tilde{a},t-5} &= \sum_{c \in C} y_{a,\tilde{a}}^{(c)}, \text{ for } C = (T+, T-, F+, F-, O+, O-)\\
\sum_{c \in C} \rho_{a,\tilde{a}}^{(c)} &= 1, \text{ for } C = (T+, T-, F+, F-, O+, O-, A-, A+)\\
S_{a,t-5}^{(true)} &= \sum_{\tilde{a}} \sum_{c \in C} y_{a,\tilde{a}}^{(c)}
\end{aligned}$$

To obtain the SSH-associated survival probability corresponding to true age-group a and SSH reported age-group $\tilde{a}$, $\pi_{a,\tilde{a}}^{(\text{ssh})}$, we extend the parametrization given in Eq. 3.10 to incorporate misclassification parameters, and probabilities associated with added and omitted siblings that are further broken down by true age group a and SSH reported age group $\tilde{a}$.

$$\pi_{a,\tilde{a}}^{(\text{ssh})} = \left(\frac{(1-\rho_{a,\tilde{a}}^{(omit)}) \cdot S_{a,t-5}^{(true)}}{(1-\rho_{a,\tilde{a}}^{(omit)}) \cdot S_{a,t-5}^{(true)} + \alpha_{a,\tilde{a}}^{(add)} \cdot D_{a,t-5}^{(true)}} \right) \cdot \tag{3.14}$$

$$\left(\frac{\tilde{\rho}_a^{(true)} - \rho_{a,\tilde{a}}^{(omit)} \kappa_{a,\tilde{a}}^{(O-)}}{(1-\rho_{a,\tilde{a}}^{(omit)})} \cdot sp_{a,\tilde{a}} + \left(1 - \frac{\tilde{\rho}_a^{(true)} - \rho_{a,\tilde{a}}^{(omit)} \kappa_{a,\tilde{a}}^{(O-)}}{(1-\rho_{a,\tilde{a}}^{(omit)})} \right) (1 - se_{a,\tilde{a}}) \right) +$$

$$\frac{\alpha_{a,\tilde{a}}^{(add)} \cdot D_{a,t-5}^{(true)} \cdot \kappa_{a,\tilde{a}}^{(A-)}}{(1-\rho_{a,\tilde{a}}^{(omit)}) S_{a,t-5}^{(true)} + \alpha_{a,\tilde{a}}^{(add)} D_{a,t-5}^{(true)}}$$

in which $\tilde{\rho}_a^{(true)}$ refers to the true survival probability of the true age group a, ie if there is no error in birth year misreporting $\tilde{\rho}_a^{(true)} = \rho_a^{(\text{true})}$. If $a = 1$ and $\tilde{a} = 2$, then $\tilde{\rho}_a^{(true)} = \rho_1^{(true)}$.

Based on Eqs. 3.13 and 3.14, we have expressions for the total number of SSH reported living siblings in true age group a, SSH reported age group $\tilde{a}$, at time t-5, $y_{a,\tilde{a},t-5}$, and the corresponding SSH associated survival probability $\pi_{a,\tilde{a}}^{(\text{ssh})}$. The SSH reported number of living siblings for age group $a+1$, reported age group $\tilde{a}+1$, at time t (the end of the reference period), is given by the data generating process below.

$$y_{a+1,\tilde{a}+1,t}^{(\text{ssh})} | y_{a,\tilde{a},t-5}^{(\text{ssh})}, \pi_{a,\tilde{a}}^{(\text{ssh})} \sim Binom(y_{a,\tilde{a},t-5}^{(\text{ssh})}, \pi_{a,\tilde{a}}^{(\text{ssh})}) \tag{3.15}$$

3.5 Exploratory Analysis

Due to limitations of the preliminary data, we use graphical exploratory analysis to assess age at death and birth year misreporting trends in the current data.

3.5.1 Age at death errors

The extent of vital status errors and added/omission errors are shown in Figure 3.9, in which proportions of added and omitted siblings, sensitivity and specificity are plotted with each age group. Due to limited data, estimates of sensitivity were equal to 1 for all age-groups, and specificity close to 1 for all age-groups. The proportion of added and omitted siblings

is notable, indicating that age at death errors at t-5 are more prevalent compared to vital status errors within the reference period, for this given population.

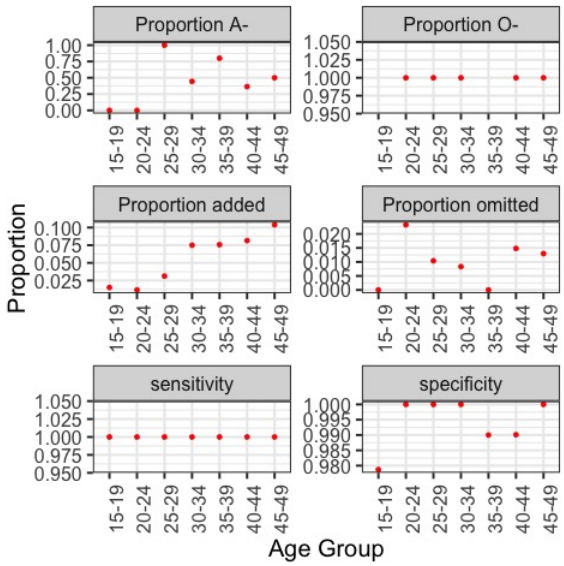

Figure 3.9: Probabilities of added and omitted siblings by age group. Proportion $A-$ refers to proportion of added negatives out of added siblings. Proportion $O-$ refers to proportion of omitted negatives out of omitted siblings. Proportion added refers to proportion added siblings out of SSH reported siblings at t-5. Lastly, proportion omitted refers to proportion omitted out of true living siblings at t-5.

3.5.2 Birth year errors

The extent of birth reporting errors, in absence of age at death errors, are shown in Figure 3.10. Based on the plot on the left, which shows SSH reported age against the true age-group, there is no systematic trend of under/over reporting of sibling age. The plot on the right shows for true age group a, the proportion of siblings reported in age group $\tilde{a}$, therefore, is a visualization of the observed $\delta_{a,\tilde{a}}$. The plot indicates that although a higher proportion of siblings have correct age reporting, the proportion of siblings with birth year reporting errors is substantial.

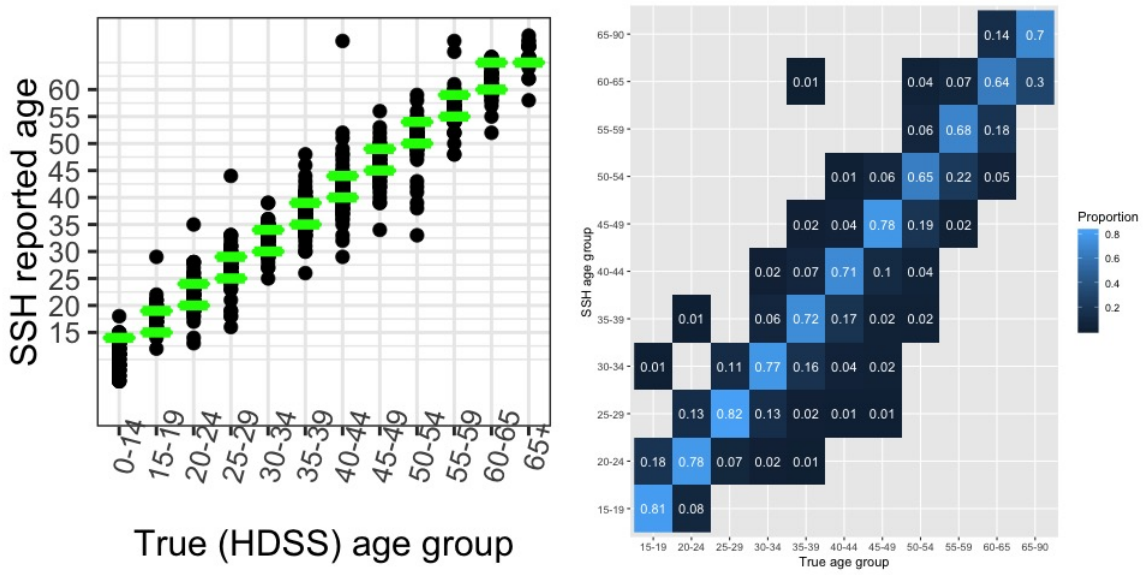

Figure 3.10: Diagram of birth year reporting errors by age group. Figure (A) shows siblings' SSH reported age group against the true age group, with green lines indicating the correct interval. Figure (B) shows a raster plot of the proportions of siblings by SSH reported age group against true age group, i.e. observed transition proportions.

3.5.3 Combination age at death and birth year errors

Lastly, in Figure 3.11 we illustrate misreporting probabilities by breakdown of age-group and a degree of difference between true age group a and reported age group $\tilde{a}$. Therefore, we visualize the combination of age at death and birth year reporting errors, for $|a-\tilde{a}|<4$, for ease of readability. Figure 3.11 suggests that due to limitations of the preliminary data, we cannot conclude the existence of an interaction between age at death and birth year reporting errors, and as such more data is needed to explore the possible relationship between the two reporting errors.

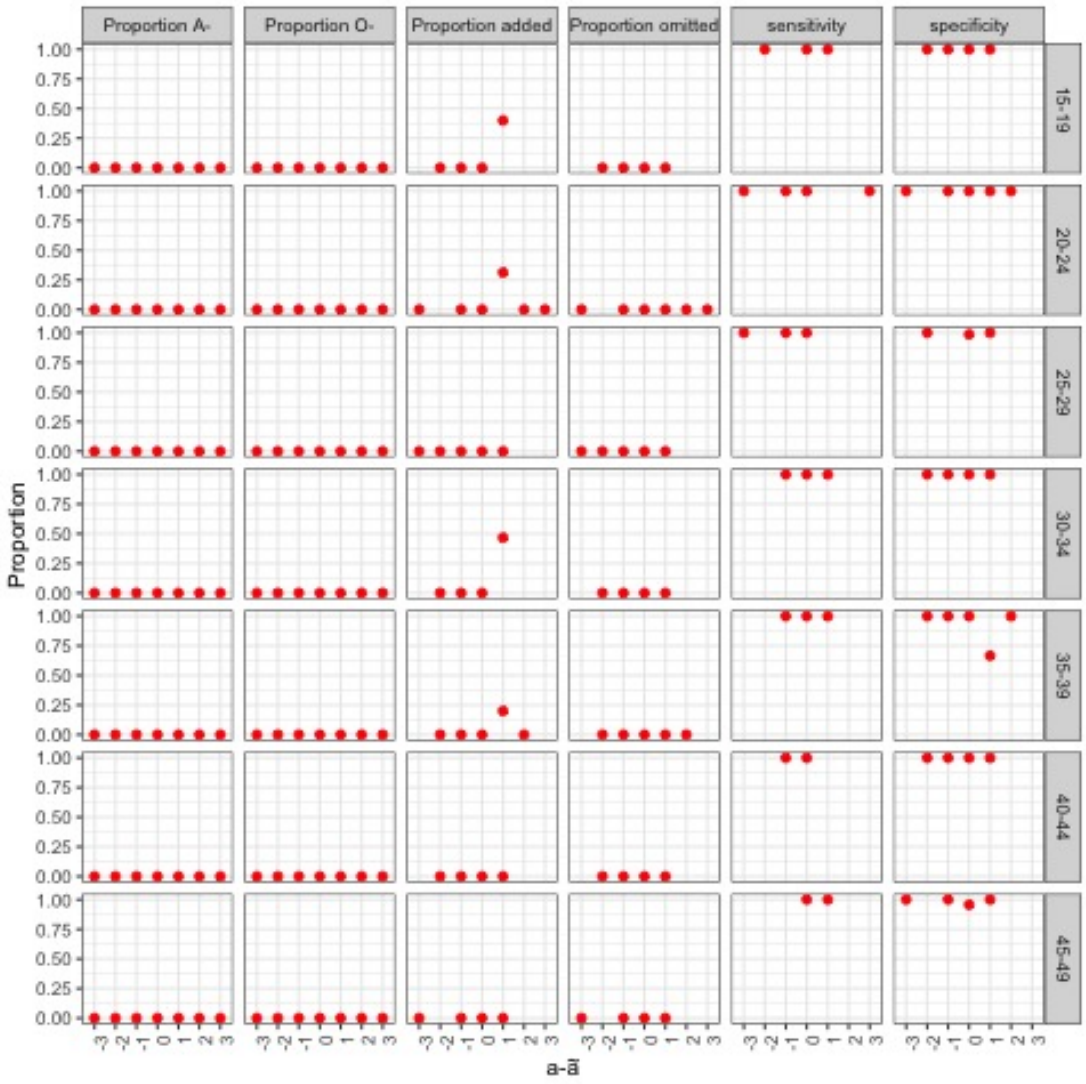

Figure 3.11: Probabilities of added and omitted siblings by true age group , and the difference between true age group a and SSH reported age-group $\tilde{a}$. Proportion $A-$ refers to proportion of added negatives out of added siblings. Proportion $O-$ refers to proportion of omitted negatives out of omitted siblings. Proportion added refers to proportion added siblings out of SSH reported siblings at t-5. Lastly, proportion omitted refers to proportion omitted out of true living siblings at t-5.

3.6 Next Steps

In this paper, we proposed an extended parametrization of misreporting in SSH adult mortality data. The model accounts for misreporting in vital status using misclassification metrics of sensitivity and specificity, and age at death before t-5 using probabilities of

added and omitted siblings. Additionally, we account for misreporting of birth year using transition probabilities. The next steps in this process are to (1) Develop a model for estimating the reporting parameters for population-periods without validation data, and (2) Use information on reporting errors when estimating adult mortality when only SSH data is available. HDSS validation data sets, in a variety of settings, are being collected currently, and will be used to explore how errors vary across populations and age-groups. This analysis will inform candidate models for reporting error parameters.

3.6.1 Incorporation of reporting errors into a larger model for adult mortality

In Eq. 3.15 we describe the data generating process to model age at death and birth year misreporting parameters, given we have both HDSS and SSH data. However, in adult mortality estimation, our aim is to estimate true adult mortality rates for populations without information on age at death or birth year reporting errors. As such, we estimate $\rho_a^{(true)}$ given we only have SSH data available, and fixed misreporting parameter estimates. As stated previously in Eq.3.13, the assumed data generating process is a multinomial distribution for the breakdown of counts for each true age-group a, reported age-group $\tilde{a}$ combination.

$$\begin{aligned} \boldsymbol{y}_{a,\tilde{a},t-5} | \boldsymbol{\rho}_{a,\tilde{a}}, y_{a,\tilde{a},t-5} &\sim Multinom(y_{a,\tilde{a},t-5}, \boldsymbol{\rho}_{a,\tilde{a}}) \\ y_{a,\tilde{a},t-5} &= \sum_{c \in C} y_{a,\tilde{a}}^{(c)}, \text{ for } C = (T+, T-, F+, F-, O+, O-) \end{aligned}$$

What is observed in SSH data is the reported total number of living siblings at time t-5 for age group $\tilde{a}$, i.e. $S_{\tilde{a},t-5}^{(\text{ssh})} = \sum_a y_{a,\tilde{a},t-5}$. Therefore, $S_{\tilde{a},t-5}^{(\text{ssh})}$ is derived as the sum of non-identical binomial distributions, i.e. different populations and different probabilities, shown in Eq. 3.16, and as such does not have a closed form solution. The exact likelihood, given the sum of non-identical binomial distributions, is in Eq. 3.16.

$$P(S^{(\text{ssh})}_{a,t-5} = s) = P(\sum_{a} y_{a,\tilde{a},t-5} = s) \tag{3.16}$$
$$= \sum_{i=0}^{s} \sum_{j=0}^{s-i} \dots \sum_{k=0}^{s-i-j-\dots-k} P(y_{1,\tilde{a},t-5} = i) P(y_{2,\tilde{a},t-5} = j) \dots P(y_{7,\tilde{a},t-5} = s-i-j-\dots-k)$$

A saddlepoint approximation has been implemented previously in the case of sums of non-identical binomial distributions (Eisinga et al. 2013, Liu et al. 2017). Computation of the exact likelihood involves enumerating all possible combinations of each variable that sums to a given value, i.e. the reported number of living siblings at time t-5. This becomes computationally infeasible when the number of combinations becomes large. The saddle point approximation is an accurate way to model the exact likelihood function without computation inefficiency.

3.7 Summary

In this chapter, we proposed a new parametrization to capture reporting errors within SSH data related to age at death and birth year reporting errors. To capture age at death errors, we parametrized these errors into metrics of sensitivity, specificity, and probabilities related to added and omitted siblings. To capture birth year reporting errors, we parametrized these errors into transition probabilities. Lastly, we propose a data-generating mechanism to incorporate misreporting parameters into a larger model for adult mortality using the sum of multinomial distributions, which does not have a closed form solution. To account for misreporting in a larger adult mortality model, estimates of misreporting parameters will be used as fixed inputs into a model for adult age-group specific mortality estimation. Exploratory analysis of preliminary data suggest that age at death errors related to added/omitted siblings are sizeable where vital status errors were minimal to none based on this limited preliminary data. In addition, birth year misreporting is prevalent across all age-groups.

CONCLUSION

We presented a Bayesian misreporting model framework for the assessment of the extent of reporting errors across different population-periods, using gold standard data to inform estimates of misreporting. The approach taken is to estimate global levels of misreporting parameters using all country-periods with gold standard data available, and to subsequently extrapolate for all country-periods without gold standard data into a larger mortality estimation model. We applied our proposed framework in the context of maternal mortality, and presented a candidate parametrization for reporting errors of adult mortality within sibling survival history data.

In Chapter I, to assess the extent of cause of death errors in civil registration vital statistics (CRVS) systems, we compare CRVS-based observed proportion maternal (PM) to those obtained from specialized studies. We developed a new approach to parametrize reporting errors in terms of sensitivity and specificity, which are data quality parameters that are comparable across different population-periods. We modeled these indicators with a bivariate hierarchical random walk model to obtain global parameter estimates. Country-year specific CRVS adjustment factors were obtained using ratios of the CRVS-based PM to the true PM for all countries with at least one specialized study. Country results showed that for countries without breakdowns of false negative and false positive maternal deaths, sensitivity and specificity were estimated to be subject to substantial uncertainty. Assessment of the relationship between CRVS adjustment factor and true PM indicated that for a country with a low true PM value, the previous United Nations Maternal Mortality Interagency Group (UN MMEIG) approach would overestimate the respective CRVS adjustment factor. Subsequently, for a country with a higher value of true PM, the UN MMEIG 2015

approach would under estimate the respective CRVS adjustment factor. Lastly, validation results confirmed that the CRVS adjustment model showed improved predictive performance of CRVS-based PM when compared to the UN MMEIG 2015 approach.

In Chapter II, we developed a new approach to extrapolate estimates of sensitivity and specificity to countries without specialized studies, and to incorporate CRVS-model-based output into the larger Bayesian maternal mortality estimation model (BMat). We implemented a sequential approach in which we first obtained point estimates as well as associated uncertainty of country specific sensitivity and specificity using the CRVS model, and then constructed estimates of misclassification parameters, for countries without studies, using an approximation method. We compared a vector autoregressive approximation to the current 2019 approach. Validation results showed that improved coverage of predictive estimates of CRVS-based PM for VAR(1) approach compared to the random walk reference year.

In Chapter III, we proposed a new parametrization to capture reporting errors within sibling survival history data related to age at death and birth year reporting errors. To capture age at death errors, we parametrize these errors into sensitivity, specificity, and probabilities related to added and omitted siblings. To capture birth year reporting errors, we parametrize these errors into transition probabilities to capture the rate at which birth year is misreporting between age groups.

The question we address is how can we learn about true mortality when we only have error prone data available? In our applications, we assessed reporting errors in civil registration vital statistics maternal mortality data and sibling survival history mortality data. We proposed a framework in which we parametrized the breakdown of reporting errors, estimate reporting errors within populations with validation data, and then generalized to populations without validation data using a sequential approach. In the assessment of adult and child mortality, estimation of cause-specific mortality rates may be improved by apply-

ing this framework to account for the extent of misreporting across different population-periods. Additionally, this framework is applicable to assess the extent of reporting errors in multiple data settings.

APPENDIX A

CRVS MISCLASSIFICATION

A.0.1 Definitions

Term	Description
Maternal death	The death of a woman whilst pregnant or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes define with the International Statistical Classification of Diseases and Related Health Problems 10th revision (ICD-10)
CRVS	Civil registration vital statistics, national death registration statistics
Specialized Study	(1) A study conducted precisely for the purpose of assessing the extent of misclassification within the CRVS and/or the extent of "missingness" of maternal deaths, (2) A study conducted to independently assess cause of death classification among the true number of maternal deaths.
BMat	Bayesian maternal mortality estimation model, used by the UN MMEIG. BMat 2019 refers to the model used in the 2019 estimation round.
Sensitivity	(1) True positive rate, (2) Proportion of correctly classified maternal deaths to the true number of maternal deaths within CRVS systems.
Specificity	(1) True negative rate, (2) Proportion of correctly classified non-maternal deaths to the true number of non-maternal deaths within CRVS systems.
True positive maternal death	A maternal death correctly classified as maternal within CRVS.
True negative maternal death	A non-maternal death correctly classified as non-maternal within CRVS.
False positive maternal death	A non-maternal death misclassified as maternal within CRVS.
False negative maternal death	A maternal death misclassified as non-maternal within CRVS.
Missed/unregistered maternal death	A maternal death unregistered (missed) within CRVS, and therefore, unreported.
PM	The proportion of maternal deaths out of the total deaths to women of reproductive age (15-49).
CRVS-based PM	The proportion of CRVS reported maternal deaths out of the total deaths to women of reproductive age within CRVS.
CRVS adjustment	Relative adjustment needed to CRVS-based PM to obtain true PM.

A.0.2 Compilation of specialized studies data

A.0.2.1 Summary of systematic review process

The objective of the review was to assess the level of misclassification reported by national official agencies for all WHO Member States. In other words, what is the level of incorrect reporting of maternal deaths in national official CRVS reporting, e.g. what is the difference between official reported number of maternal deaths versus the number of maternal deaths identified through special maternal mortality studies, confidential enquiries and surveillance systems etc. And to what extent is the incorrect reporting of maternal death due to misclassification versus missed or unregistered maternal deaths?

This review identified studies that fulfilled inclusion criteria as follows:

	Inclusion Criteria
Population	Women of reproductive age (15-49 years) who died during pregnancy or up to one year after termination of pregnancy, irrespective of duration and the site of the pregnancy, from any cause.
Concept	Assessment of misclassification of maternal deaths by CRVS systems.
Study design	Cross-sectional study and retrospective cohort
Context	All WHO Member States reporting CRVS data

In addition, the following criteria has to be met for inclusion:

1. study is nationally representative;
2. mid-years of reported data are after 1990;
3. there is a matched comparison of CRVS data available in the study or in the WHO Mortality Database.

A.0.2.1.1 Search Strategy The search strategy was conducted for all relevant existing literature based on search terms relevant to the research questions restricted to the years 1990-2016, using the following online bibliographic databases: PubMed/MEDLINE, EMBASE, Global Index Medicus, EBSCO, Web of Science and Popline. The searches were

conducted without any language restrictions. Search terms are included in Box at the end of this document. A hand search was also conducted on all WHO Member States Ministries of Health (MoH) websites to identify pertinent MoH maternal mortality and confidential inquiries reports.

A.0.2.1.2 Data Extraction Data were extracted from full-text journal articles and reports which met the inclusion criteria. Data were extracted using a Microsoft Excel database. Information retrieved from the included studies included country, years assessed, study objectives, methodology /study design, number of maternal deaths, information on misclassification and incompleteness when available. Specifically, extraction focused on the assessment of the following:

1. The process by which the study retrieved and reviewed information on maternal deaths, including data source descriptions, definitions used by study, and whether the study reviewed all deaths to women of reproductive age or a description of the subset of deaths collected.

2. The number of maternal deaths, any information pertaining to misclassification of maternal cause of death by the CRVS system, any information regarding missed deaths by maternal cause.

3. Breakdown of maternal deaths by maternal cause of death was extracted if reported.

A.0.2.2 Compilation of data

The PRISMA diagram in Figure A.1 provides information on the number of study documents and associated study observations both identified and included by (1) systematic review, (2) WHO maternal mortality database, and (3) information obtained from follow-up surveys and country consultation. Lastly, it reports the number of studies excluded and

reason for exclusion at each stage of the screening process. Studies were excluded in 3 subsequent steps. Firstly, studies were excluded if they reported information that could not be used, i.e. if no information on maternal death counts in the CRVS or associated envelopes could be obtained (non-usable data). Secondly, a study was excluded if it was not nationally representative. Lastly, a study was excluded if an alternate study with more up-to-date or detailed information for the same country-period was available. The complete set of references of the included study documents is given in Box 1 at the end of this document.

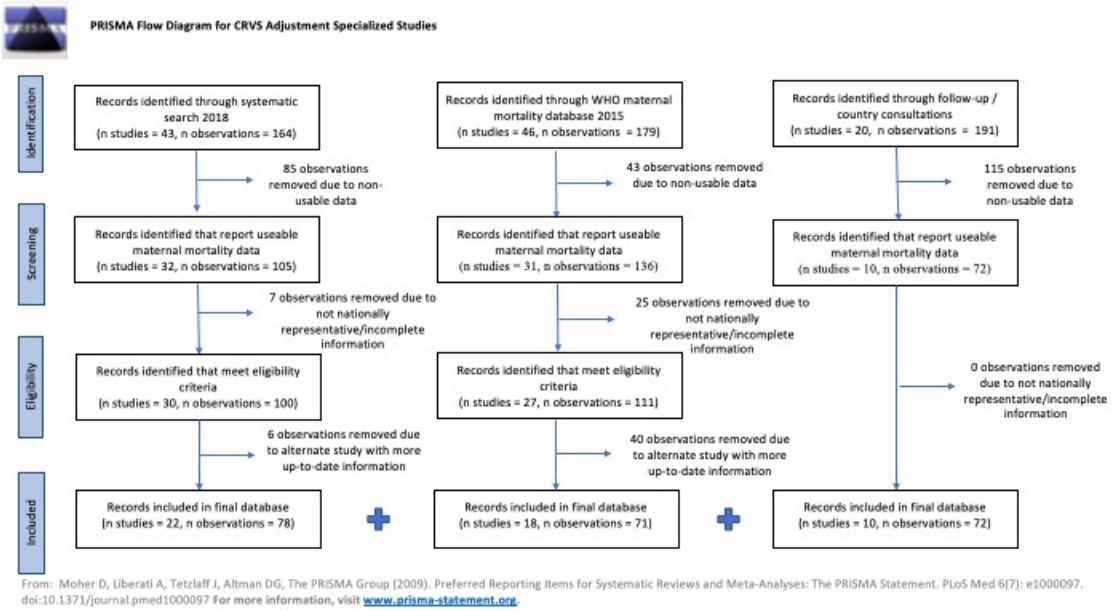

Figure A.1: PRISMA flow diagram of data compilation of specialized studies for inclusion in the CRVS adjustment model. The numbers of studies mentioned refer to study documents.

A.0.3 Covariate plots

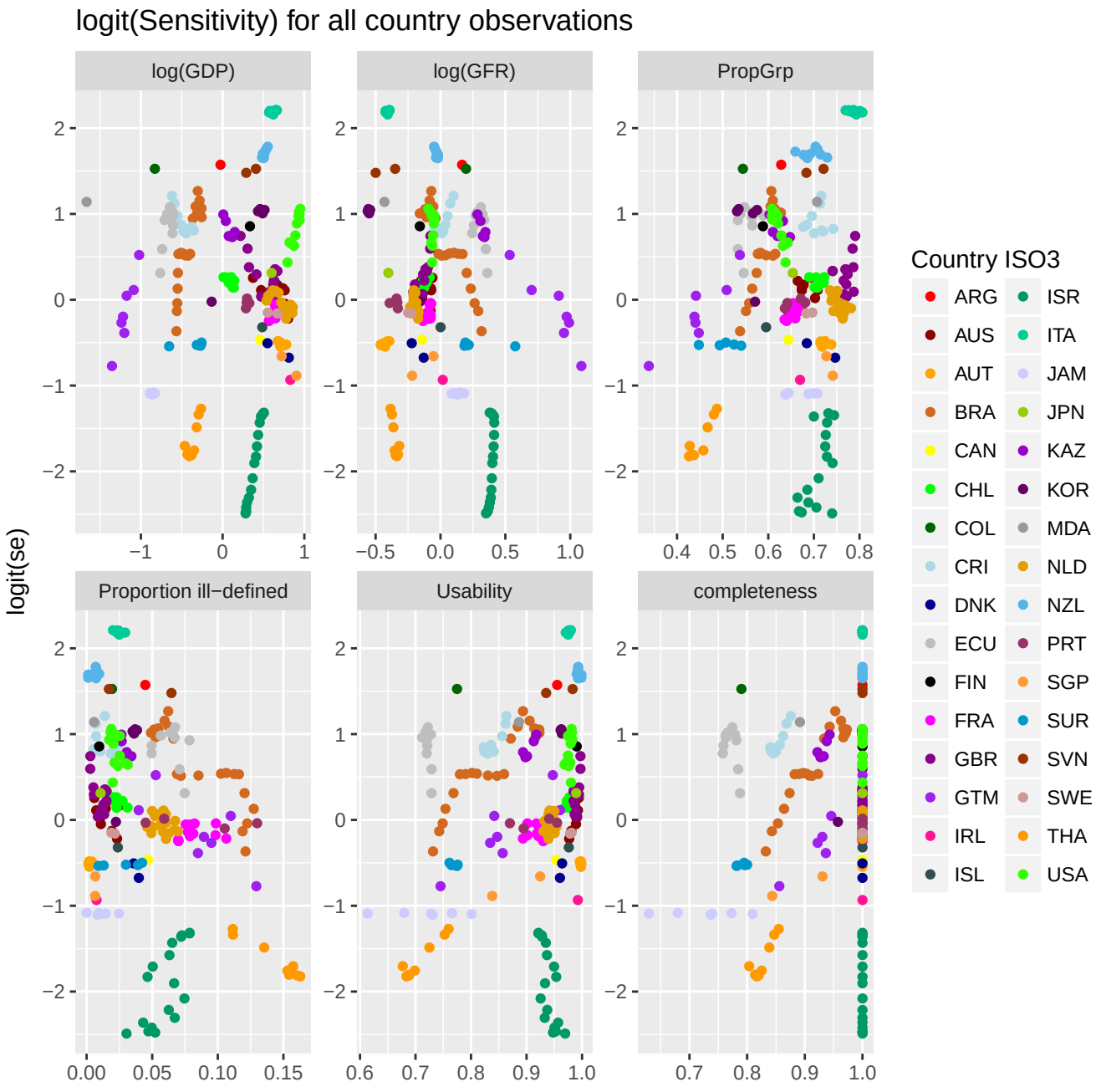

Figure A.2: Estimates of sensitivity (on logit-scale) plotted against covariates.

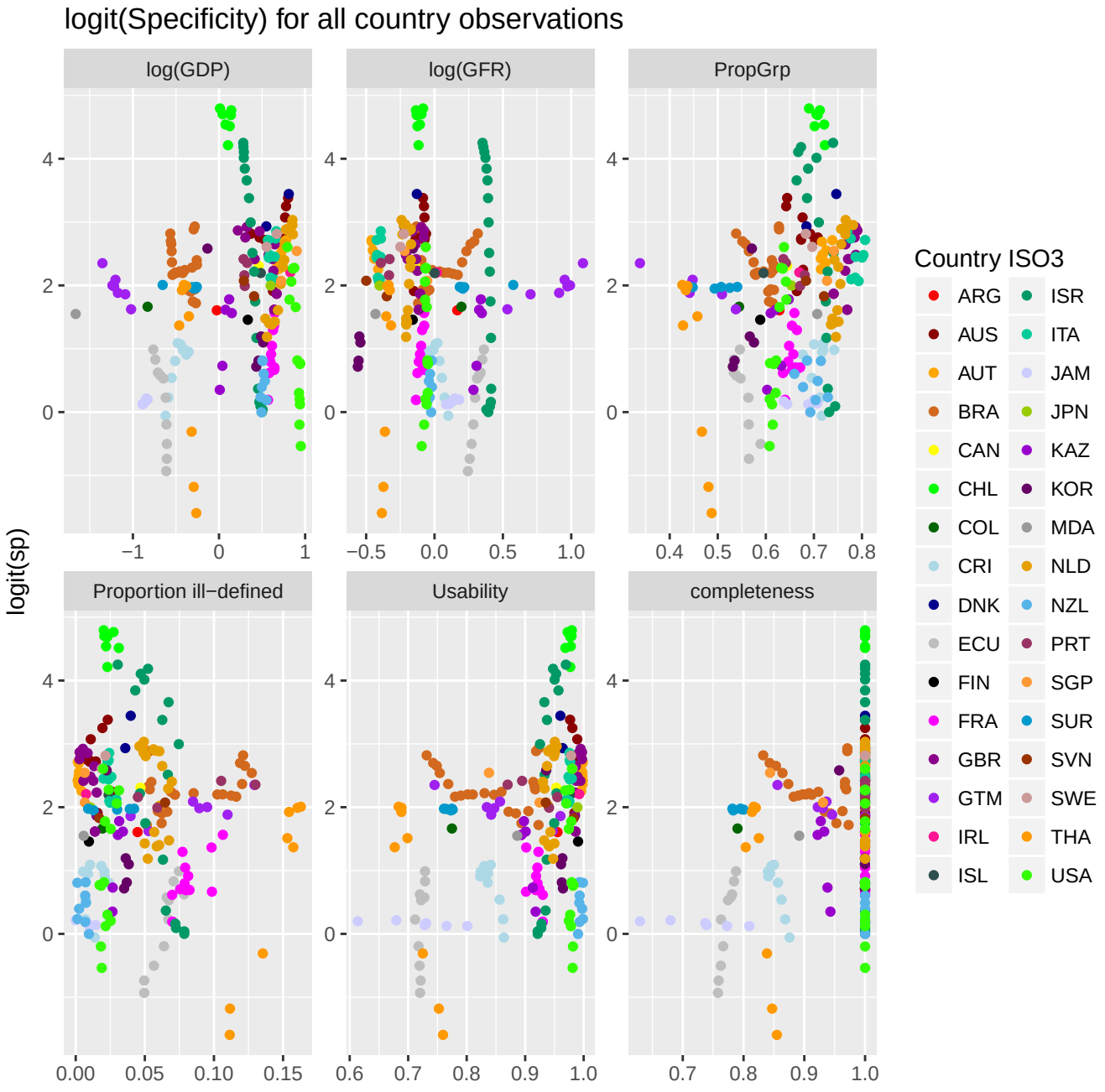

Figure A.3: Estimates of specificity (on logit-scale) plotted against covariates.

A.0.4 BMat 2019

The approach by which CRVS data are used to inform maternal mortality estimates in BMat 2019 builds upon the model for sensitivity and specificity in CRVS reporting and BMat 2015 assumptions. In summary, a two-step approach is taken:

1. We obtain point estimates of misclassification parameters from the CRVS model, as explained in Section 2.3.1.

2. The estimated misclassification parameters are used in BMat for country-years with CRVS data and without specialized studies, see Section A.0.4.1.

A.0.4.1 BMat 2019 data model for CRVS data

In BMat 2019, the data model for observed CRVS data is as follows:

$$y_{c,t}^{(\text{matCRVS})} | \rho_{c,t}^{(truemat)}, y_{c,t}^{(CRVS)} \quad \sim \quad NegBin\left(E_{c,t}, V_{c,t}\right), \tag{A.1}$$

where (following notation from the main paper), $y_{c,t}^{(\text{matCRVS})}$ refers to the number of maternal deaths as observed in the CRVS in country c in year t, $\rho_{c,t}^{(truemat)}$ is the true probability of a maternal death among all deaths, and $y_{c,t}^{(CRVS)}$ is the total number of deaths registered in the CRVS.

$E_{c,t}$ and $V_{c,t}$ are defined as follows:

$$\begin{aligned}
E_{c,t} &= y_{c,t}^{(CRVS)} \cdot \left(\hat{\lambda}_{c,t}^{(+)} \rho_{c,t}^{(truemat)} + \left(1 - \hat{\lambda}_{c,t}^{(-)}\right) \left(1 - \rho_{c,t}^{(truemat)}\right) \right), && \text{(A.2)} \\
V_{c,t} &= E_{c,t} + y_{c,t}^{2(\text{matCRVS})} \cdot \left(\tilde{V}_{1,c,t} + \tilde{V}_{2,c,t} \right), && \text{(A.3)} \\
\tilde{V}_{1,c,t} &= \hat{v}_{c,t}^{(+)} \cdot \rho_{c,t}^{2(truemat)} + \hat{v}_{c,t}^{(-)} \cdot \left(1 - \rho_{c,t}^{(truemat)}\right)^2 && \text{(A.4)} \\
&\quad -2 \cdot \rho_{c,t}^{(truemat)} \cdot \left(1 - \rho_{c,t}^{(truemat)}\right) \hat{u}_{c,t}, && \text{(A.5)} \\
\tilde{V}_{2,c,t} &= \hat{m}_{c,t} \cdot \rho_{c,t}^{2(truemat)} \cdot \left(\hat{e}_{c,t}^{(+)} + \hat{e}_{c,t}^{(-)} \right), && \text{(A.6)}
\end{aligned}$$

where $\hat{\lambda}_{c,t}^{(+)}$ and $\hat{\lambda}_{c,t}^{(-)}$ refer to point estimates for sensitivity and specificity, $\hat{v}_{c,t}^{(+)}$ and $\hat{v}_{c,t}^{(-)}$ to estimated variances for sensitivity and specificity, $\hat{u}_{c,t}$ to the estimated covariance between sensitivity and specificity, $\hat{e}_{c,t}^{(+)}$ to the estimated squared sensitivity and $\hat{e}_{c,t}^{(-)}$ to estimated squared (1- specificity). Finally, $\hat{m}_{c,t} = 0$ for country-years with complete CRVS. In country-years with incomplete CRVS, $\hat{m}_{c,t}$ is the estimated variance of $\theta_{c,t}$, with

$$\theta_{c,t} = 1/\left(\rho_{c,t}^{(\text{CRVS})} + \left(1-\rho_{c,t}^{(\text{CRVS})}\right)\kappa_{c,t}\right), \tag{A.7}$$

due to uncertainty in the ratio of probabilities of a maternal death among unregistered versus registered deaths $\kappa_{c,t}$ (see Section 1.4.1). $\hat{m}_{c,t}$ is approximated using a monte carlo approximation; we set $\hat{m}_{c,t} = Var(\theta_{c,t}^{(h)})$, where samples $\theta_{c,t}^{(h)}$ are constructed as follows:

$$\log\left(\kappa_{c,t}^{(h)}\right) \sim N(0,1), \tag{A.8}$$

$$\theta_{c,t}^{(h)} = 1/\left(\rho_{c,t}^{(\text{CRVS})} + \left(1-\rho_{c,t}^{(\text{CRVS})}\right)\kappa_{c,t}^{(h)}\right). \tag{A.9}$$

In summary, the variance in θ is determined by the variability in the ratio of probabilities κ. The lognormal distribution assigned to κ results in first and third quantiles of κ around 0.5 and 2, respectively, to reflect the uncertainty associated with this ratio.

The derivation of the data model for CRVS data in Eq. A.1 is based on the following assumptions:

$$y_{c,t}^{(\text{matCRVS})}|\gamma_{c,t}^{(\text{matCRVS})} \sim \textit{Poisson}\left(\gamma_{c,t}^{(\text{matCRVS})} \cdot y_{c,t}^{(CRVS)}\right), \tag{A.10}$$

$$\gamma_{c,t}^{(\text{matCRVS})}|\rho^{(\text{truemat})} \sim \textit{Gamma}(g_1, g_2), \tag{A.11}$$

with g_1 and g_2 such that $E\left(\gamma_{c,t}^{(\text{matCRVS})}|\rho^{(\text{truemat})}\right) = E_{c,t}/y_{c,t}^{(CRVS)}$ and $V\left(\gamma_{c,t}^{(\text{matCRVS})}|\rho^{(\text{truemat})}\right) = \tilde{V}_{1,c,t} + \tilde{V}_{2,c,t}$.

The data model in Eq. A.1 specifies which estimates of misclassification parameters are needed to include CRVS-based data into BMat: point estimates $\hat{\lambda}_{c,t}^{(+)}$, $\hat{\lambda}_{c,t}^{(-)}$, (co-)variance es-

timates $\hat{v}_{c,t}^{(+)}$, $\hat{v}_{c,t}^{(-)}$ and $\hat{u}_{c,t}$, and estimated squared sensitivity $\hat{e}_{c,t}^{(+)}$ and squared (1- specificity) $\hat{e}_{c,t}^{(-)}$.

A.0.4.2 Data model for specialized studies in BMat 2019

Let specialized studies be indexed by i, with the ith study referring to country $c[i]$, observation period $t1[i]$ to $t2[i]$ and midpoint $t[i]$. Let $\rho_{c,t1,t2}^{(truemat)}$ refer to the true probability of a maternal death in country c for the period from $t1$ to $t2$, obtained from the annual probabilities weighted by the total deaths in each year:

$$\rho_{c,t1,t2}^{(truemat)} = \frac{\sum_{t=t1}^{t2} \rho_{c,t}^{(truemat)} y_{c,t}^{(tot)}}{\sum_{t=t1}^{t2} y_{c,t}^{(tot)}}.$$

Data models are discussed separately for studies with complete envelopes $z_i^{(env)} = z_i^{(tot)}$, versus those with incomplete envelopes $z_i^{(env)} < z_i^{(tot)}$.

A.0.4.2.1 Studies with complete envelopes For specialized study i with envelope $z_i^{(env)} = z_i^{(tot)}$, we assumed

$$z_i^{(truemat)} | \rho_{c[i],t1[i],t2[i]}^{(truemat)} \sim Bin\left(z_i^{(tot)}, \rho_{c[i],t1[i],t2[i]}^{(truemat)}\right),$$

where as before, $z_i^{(truemat)}$ refers to the number of maternal deaths as observed in the specialized study, and $z_i^{(tot)}$ to its respective envelope of all-cause deaths.

A.0.4.2.2 Studies with incomplete envelope For specialized study i with incomplete envelope $z_i^{(env)} < z_i^{(tot)}$, we assumed (following assumptions and notation from the data model for CRVS data Eq.A.1):

$$z_i^{(truemat)} | \rho_{c[i],t1[i],t2[i]}^{(truemat)} \sim NegBin(E_i, V_i), \tag{A.12}$$

where, setting $\rho_i = \rho^{(truemat)}_{c[i],t1[i],t2[i]}$ to improve readability,

$$
\begin{aligned}
E_i &= z_i^{(env)} \cdot \left(\hat{\lambda}^{(+)}_{c[i],t[i]} \rho_i + (1 - \hat{\lambda}^{(-)}_{c,t})(1 - \rho_i) \right), && \text{(A.13)} \\
V_i &= E_{c,t} + z_i^{2(env)} \cdot (\tilde{V}_{1,i} + \tilde{V}_{2,i}), && \text{(A.14)} \\
\tilde{V}_{1,i} &= \hat{v}^{(+)}_{c[i],t[i]} \cdot \rho_i^2 + \hat{v}^{(-)}_{c[i],t[i]} \cdot (1 - \rho_i^2) && \text{(A.15)} \\
&\quad -2 \cdot \rho_i \cdot (1 - \rho_i) \hat{u}_{c[i],t[i]}, && \text{(A.16)} \\
\tilde{V}_{2,i} &= \hat{m}_i \cdot \rho_i^2 \cdot \left(\hat{e}^{(+)}_{c[i],t[i]} + \hat{e}^{(-)}_{c[i],t[i]} \right), && \text{(A.17)}
\end{aligned}
$$

where $\hat{m}_i$ is the estimated variance of θ_i, with

$$
\theta_i = \frac{1}{z_i^{(env)}/z_i^{(tot)} + \left(1 - z_i^{(env)}/z_i^{(tot)}\right) \kappa_i}, \qquad \text{(A.18)}
$$

due to uncertainty in the ratio of probabilities of a maternal death among uncaptured versus captured deaths κ_i. We set $\hat{m}_i = Var(\theta_i^{(h)})$, where samples $\theta_i^{(h)}$ are constructed as follows:

$$
\begin{aligned}
\log\left(\kappa_i^{(h)}\right) &\sim N(0,1), && \text{(A.19)} \\
\theta_i^{(h)} &= \frac{1}{z_i^{(env)}/z_i^{(tot)} + \left(1 - z_i^{(env)}/z_i^{(tot)}\right) \kappa_i^{(h)}}. && \text{(A.20)}
\end{aligned}
$$

Box 1: References for specialized studies included in the CRVS adjustment model.

Argentina
Abalos E, Duhau M, Escobar P, Fasola ML, Finkelstein JZ, Golubicki JL et al. Omisión de registros de causas maternas de muerte en Argentina: estudio observacional de alcance nacional. Rev Panam Salud Publica 43. 2019 : 1-10.

Australia
Sullivan EA, Ford JB, Chambers G, Slaytor EK. Maternal mortality in Australia 1973–1996. AIHW Cat. No. PER 24. Sydney: AIHW National Perinatal Statistics Unit. (Maternal Deaths Series No.1); 2004.

Australia
Report on Maternal Deaths in Australia, 1994–1996. Canberra: National Health and Medical Research Council (NHMRC), AIHW National Perinatal Statistics Unit; 2001.

Australia
Slaytor EK, Sullivan EA, King JF. Maternal deaths in Australia 1997–1999. AIHW Cat. No. PER 24. Sydney: AIHW National Perinatal Statistics Unit. (Maternal Deaths Series No.1); 2004.

Australia
Australian Institute of Health and Welfare (AIHW): Humphrey MD, Bonello MR, Chughtai A, Macaldowie A, Harris K, Chambers GM. Maternal deaths in Australia 2008–2012. Maternal Deaths Series no. 5. CAT. No. PER 70. Canberra: AIHW; 2015.

Australia
Maternal deaths in Australia 2012–2014. Cat. No. PER 92. Canberra: Australian Institute of Health and Welfare (AIHW); 2017.

Austria
Beck A, Vutuc C. Die Entwicklung der mütterlichen Mortalität in Österreich. Frauenarzt. 2008;49:21-6.

Brazil
Technical Note of the Ministry of Health of Brazil. Maternal mortality in Brazil: estimates versus reality. From country consultation in 2015. Ministry of Health, Secretariat of Health Surveillance [Brazil]; 2015.

Brazil
PAHO maternal mortality data set (reflecting communication on data requests collected via questionnaire). November–December 2018.

Brazil
Data received from country consultations. 2019

Canada
Turner LA, Cyr M, Kinch RAH, Liston R, Kramer MS, Fair M, et al. Under-reporting of maternal mortality in Canada: a question of definition. Chronic Dis Can. 2002; 23(1):22–30.

Chile
PAHO maternal mortality data set (reflecting communication on data requests collected via questionnaire). November–December 2018.

Colombia
PAHO maternal mortality data set (reflecting communication on data requests collected via questionnaire). November–December 2018.

Costa Rica
Proceso de Búsqueda Intencional y Reclasificación de Mortalidad Materna [Intentional Search Process and Reclassification of Maternal Mortality (ISPRMM)]. San José: Ministerio de Salud, Instituto Nacional de Estadística y Censos (INEC) [Costa Rica]; 2014 (in Spanish).

Denmark
Andersen BR, Westergaard HB, Bødker B, Weber T, Møller M, Sørensen JL. Maternal mortality in Denmark, 1985–1994. Eur J Obstet Gynecol Reprod Biol. 2009; 142:124–8.

Denmark
Bødker B, Hvidman L, Weber T, Møller M, Aarre A, Nielsen KM, et al. Maternal deaths in Denmark 2002–2006. Acta Obstet Gynecol Scand. 2009;88:556-62.

Ecuador
PAHO maternal mortality data set (reflecting communication on data requests collected via questionnaire). November–December 2018.

Finland
Gissler M et al. Pregnancy-associated deaths in Finland 1987-1994 – definition problems and benefits of record linkage. Acta Obstet Gynecol Scand. 1997;76(7):651-7.

France
Saucedo M, Deneux-Tharaux C, Bouvier-Colle MH; French National Experts Committee on Maternal Mortality. Ten years of confidential inquiries into maternal deaths in France, 1998-2007. Obstet Gynecol. 2013 Oct;122(4):752-60.

France
Rapport du Comité national d'experts sur la mortalité maternelle (CNEMM): 1999-2001

France

Les morts maternelles en France: mieux comprendre pour mieux prévenir. 5e rapport de l'Enquête Nationale Confidentielle sur les Morts Maternelles (ENCMM) 2010-2012. France: Institut National de la Santé et de la Recherche Médicale (INSERM), Sante Publique France; 2017 (in French). Additional information on numbers of maternal deaths was received during the country consultation 2019.

France
Data received during country consultation 2019.

Georgia
Serbanescu F, Tefft M, Shakhnazarova M, Williams D, Berdzuli N, Berg C. Reproductive Age Mortality Study, Georgia, 2008 – Part II: Maternal Mortality. Atlanta (GA): Georgian National Center for Disease Control, JSI Research & Training Institute, Inc (JSI) and CDC; 2009.

Georgia
Georgia Reproductive Age Mortality Study (RAMOS) 2014. Executive summary. Georgia; 2015.

Georgia
Data received for 2006, 2012, 2015, 2016 and 2017 during the country consultation 2019.

Guatemala
Schieber B, Stanton C. Estimates of Maternal Mortality in Guatemala 1996 – 1998. Guatemala; 2000.

Guatemala
Línea Basal de Mortalidad Materna para el Año 2000. Informe final. Ciudad de Guatemala: Ministerio de Salud Pública y Asistencia Social [Guatemala]; 2003 (in Spanish).

Guatemala
Situación de la Mortalidad Materna. Informe de País 2013. Guatemala: Ministerio de Salud Publica y Asistencia Social; 2015 (in Spanish).

Iceland
Birgisdottir H, Bjarnadottir RI, Kristjansdottir K, Geirsson RT. Maternal deaths in Iceland over 25 years. Acta Obstet Gynecol Scand. 2016 Jan;95(1):74-8. doi:10.1111/aogs.12797. Epub 2015 Nov 14.

Ireland
Confidential Maternal Death Enquiry in Ireland, Report for Triennium 2009–2011. Cork: Maternal Death Enquiry (MDE); 2012.

Israel
Data received for 2000-2016 during the country consultation 2019.

Italy
Data received for 2006-2014 during the country consultation 2019.

Jamaica
PAHO maternal mortality data set (reflecting communication on data requests collected via questionnaire). November-December, 2018.

Japan
Hidaka A, Fukuda H, Imoto H, Yamazaki T, Muranaka J. Nishimura J, et al. [Causes and ratio of maternal mortality, and its reliability]. Sanfujinka Chiryo [Obstetrical and gynaecological therapy]. 2009;99(1):85-95 (in Japanese).

Kazakhstan
Communication after mission 2014 with MoH (mission-related documentation) [e-mail].

Kazakhstan
Findings of a confidential audit of maternal mortality rates in the Republic of Kazakhstan in 2011–2013 [unofficial translation]. Astana City: The Central Commission on Confidential Audit (CCAC); 2014.

Moldovia
Hodorogea S, Friptu V. The Moldovan experience of maternal deaths reviews. BJOG. 2014;121(Suppl. 4):81-5. doi:10.1111/1471-0528.12945.

Netherlands
Schutte JM, Steegers EA, Schuitemaker NW, Santema JG, de Boer K, Pel M, Vermeulen G, Visser W, van Roosmalen J. Rise in maternal mortality in the Netherlands.; Netherlands Maternal Mortality Committee. BJOG. 2010 Mar;117(4):399-406. doi: 10.1111/j.1471-0528.2009.02382.x. Epub 2009 Nov 26.

Netherlands
Communication with the Dutch Maternal Mortality Committee (MMC) from the Netherlands Society of Obstetrics and Gynaecology [e-mail, 22 August 2013].

New Zealand
Perinatal and Maternal Mortality Review Committee (PMMRC). Eleventh Annual Report of the Perinatal and Maternal Mortality Review Committee. Reporting mortality and morbidity 2015. Wellington: Health Quality & Safety Commission [New Zealand]; 2017.

Portugal
Gomes, MC, Ventura MT, Nunes RS. How many maternal deaths are there in Portugal? J Matern Fetal Neonatal Med. 2012;25(10):1975-9. doi:10.3109/14767058.2012.668587.

Paraguay

Búsqueda Intencionada y Reclasificación de Muertes Maternas en Paraguay INFORME 2015. Ministerio de Salud Pública y Bienestar Social (MSPyBS) : 2016 (in Spanish).

Paraguay
Information retrieved during country consultation 2019.

Republic of Korea
Han Y, Doh S, Park J, Lee S. Maternal Mortality Ratio and Causes of Death in 1995-1996 in Korea. Sejong City: Korea Institute for Health and Social Affairs, Ministry of Health and Welfare [Republic of Korea]; 1997.

Republic of Korea
From country consultation in 2015. Daejeon: Vital Statistics Division of Statistics Korea (KOSTAT); 2015.

Singapore
Lau G. Are Maternal Deaths on the Ascent in Singapore? A review of Maternal Mortality as Reflected by Coronial Casework from 1990 to 1999. Ann Acad Med Singapore. 2002;31(3):261-75.

Slovenia
Kralj E, Mihevc-Ponikvar B, Premru-Sršen T, Balažic J. Maternal mortality in Slovenia: Case report and the method of identifying pregnancy-associated deaths. Forensic Science International Supplement Series. 2009;1:52-7. doi:10.1016/j.fsisup.2009.10.001.

Slovenia
Country communication with WHO post Kralj et l. 2009 study methodology

Surniname
Kodan LR, Verschueren KJC, van Roosmalen J, Kanhai HHH, Bloemenkamp KWM. Maternal mortality audit in Suriname between 2010 and 2014, a reproductive age mortality survey. BMC Pregnancy Childbirth. 2017;17:275. doi:10.1186/s12884-017-1466-6.

Sweden
Esscher A, Högberg U, Haglund B, Essén B. Maternal mortality in Sweden 1988–2007: more deaths than officially reported. Acta Obstet Gynecol Scand. 2013;92:40-6. doi:10.1111/aogs.12037.

Sweden
Grunewald C, Nilsson E, Cnattingius S, Westgren M, Stephanson O. Mödradödligheten underskattad i Sverige. Registerstudie av död i samband med graviditet, förlossning och post-partum. [Maternal mortality in Sweden underestimated. Registry study of death in connection with pregnancy, delivery and postpartum.] Läkartidningen (Sweden). 2008;105(34):2250-3 (in Swedish).

Thailand
Chandoevwit W, Phatchana P, Sirigomon K, Ieawsuwan K, Thungthong J, Ruangdej S. Improving the measurement of maternal mortality in Thailand using multiple data sources. Popul Health Metr. 2016;14:16. doi 10.1186/s12963-016-0087-z.

United Kingdom
Knight M, Kenyon S, Brocklehurst P, Neilson J, Shakespeare J, Kurinczuk JJ (editors) on behalf of MBRRACE-UK. Saving Lives, Improving Mothers' Care: Lessons learned to inform future maternity care from the UK and Ireland Confidential Enquiries into Maternal Deaths and Morbidity 2009–2012. Oxford: National Perinatal Epidemiology Unit, University of Oxford; 2014.

United Kingdom
Knight M, Nair M, Tuffnell D, Shakespeare J, Kenyon S, Kurinczuk JJ (editors) on behalf of MBRRACE-UK. Saving Lives, Improving Mothers' Care: Lessons learned to inform maternity care from the UK and Ireland Confidential Enquiries into Maternal Deaths and Morbidity 2013–2015. Oxford: National Perinatal Epidemiology Unit, University of Oxford; 2017.

United States of America
CDC's Pregnancy Mortality Surveillance System. From country consultation 2015.

United States of America
Creanga AA, Berg CJ, Syverson C, Seed K, Bruce FC, Callaghan WM. Pregnancy-related mortality in the United States, 2006-2010. Obstet Gynecol. 2015;125(1):5-12.

Uruguay
Data received for 2000-2017 during the country consultation 2019.

End of Box 1.

Box 2: SEARCH TERMS

EMBASEw
http://www.embase.com
No AGE, HUMAN
YEAR limits applied : [1990-2050]/py
Options Also search as free text was enabled.

#	Searches	Results
1	'maternal mortality'/exp OR 'maternal mortality' OR 'maternal mortalities'	22,873
2	'underreporting' OR 'under reporting' OR underreported OR 'under reported' OR 'data quality' OR 'official figures' OR 'record linkage' OR 'quality of information' OR 'officially reported' OR 'multiple sources' OR 'linkage' OR 'under registered' OR 'under registration' OR underregistered OR underregistration OR 'under registering' OR 'source of error' OR 'misclassification' OR 'misclassified' OR (errors AND ('registration'/exp OR registration)) OR 'late maternal mortality' OR 'confidential enquiries' OR 'confidential enquiry'	180888
3	'data collection method'/exp OR 'health survey'/exp AND (standard* OR method*)	558644
4	#1 AND (#2 OR #3)	1760
5	'pregnancy'/exp OR 'pregnancy complication'/exp OR 'pregnancy disorder'/exp OR 'abortion'/exp AND ('death'/exp OR deaths OR 'mortality'/exp OR fatal OR fatalities OR deceased)	105286
6	#2 AND #5	1143
7	#4 OR #6	2440
8	#4 OR #6 AND [1990-2016]/py	2335

PubMed
http://www.pubmed.gov
Filters: Publication date from 1990/01/01 to 2050/12/31

#	Searches	Results
1	"underreporting"[tiab] OR “under reporting “[tiab] OR underreported [tiab]OR ” under reported” [tiab]OR “data quality” [tiab] OR “official figures” [tiab] OR “record linkage” [tiab] OR “quality of information” [tiab] OR “officially reported “[tiab] OR “multiple sources” [tiab] OR” linkage” [tiab] OR “under registered” [tiab] OR” under registration” [tiab] OR “under registering” [tiab]	209385

	OR underregistered[tiab] OR underregistration[tiab] OR “under registering” [tiab] OR “source of error” [tiab] OR “misclassification” [tiab] OR “misclassified “[tiab] OR (errors[tiab] AND registration[tiab]) OR “late maternal mortality” [tiab] OR “confidential enquiries” [tiab] OR “confidential enquiry” [tiab] OR "Data Collection/methods"[Mesh] OR "Data Collection/standards"[Mesh] OR "Population Surveillance/methods"[Mesh] OR "Population Surveillance/standards"[Mesh]	
2	"Maternal Mortality"[Mesh] OR "maternal mortality" [Tw] OR "maternal mortalities" [Tw] OR ((Pregnancy[mesh] OR "pregnancy complications" [Mesh] or “pregnant women” or parturition[mesh] or mothers[mesh] or "maternal health services"[mesh] or pregnancy or pregnant or parturition or mother* or gestation or gestational or childbirth or childbirths or maternal or maternity) AND (mortality OR mortalities OR Death OR deceased OR fatality OR fatalities))	116919
3	#1 AND #2	1977
4	"mothers"[MeSH Terms] OR "mothers"[All Fields] OR "mother"[All Fields] OR "mothers"[MeSH Terms] OR "mothers"[All Fields] OR "maternal"[All Fields] OR "pregnancy"[MeSH Terms] OR "pregnancy"[All Fields] OR "parturition"[MeSH Terms] OR "parturition"[All Fields] OR "postpartum period"[MeSH Terms] OR "postpartum"[All Fields] AND "period"[All Fields] OR "postpartum period"[All Fields] OR "postpartum"[All Fields] OR antepartum[All Fields] OR intrapartum[All Fields] OR "parturition"[MeSH Terms] OR "parturition"[All Fields] OR "childbirth"[All Fields] OR "delivery, obstetric"[MeSH Terms] OR ("delivery"[All Fields] AND "obstetric"[All Fields]) OR "obstetric delivery"[All Fields] OR "parturition"[MeSH Terms] OR "parturition"[All Fields] OR "birth"[All Fields] OR termination[All Fields] OR "abortion, induced"[MeSH Terms] OR ("abortion"[All Fields] AND "induced"[All Fields]) OR "induced abortion"[All Fields] OR "abortion"[All Fields] OR "abortion, spontaneous"[MeSH Terms] OR ("abortion"[All Fields] AND "spontaneous"[All Fields]) OR "spontaneous abortion"[All Fields] OR "miscarriage"[All Fields]	566626
5	"death"[MeSH Terms] OR "death"[All Fields] OR fatal[All Fields] OR fatality[All Fields] OR "mortality"[Subheading] OR "mortality"[All Fields] OR "mortality"[MeSH Terms]	1543760
6	#4 AND #5 AND #1	1715
7	#6 OR #3	2516
8	#6 OR #3Publication date from 1990/01/01 to 2050/12/31	2285

Global Index Medicus
http://www.globalhealthlibrary.net
No AGE, HUMAN or YEAR limits applied.
Options : Regional Indexes searched

#	Searches (LILACS)	Results
1	((underreporting) OR (under reporting) OR underreported OR (under reported) OR (data quality) OR (official figures) OR (record linkage) OR (quality of information) OR (officially reported) OR (multiple sources) OR (linkage) OR (under registered) OR (under registration) OR (under registering) OR underregistered OR underregistration OR (under registering) OR (source of error) OR (misclassification) OR (misclassified) OR (errors AND registration) OR (late maternal mortality) OR (confidential enquiries) OR (confidential enquiry)) AND ((MOTHERS AND Mortality) OR (Maternal Mortality) OR (Maternal Death) OR (maternal mortality) OR (maternal deaths) OR (pregnancy related deaths) OR (pregnancy related deaths))	910
	IMEMR (same as above)	163
	WPRIM (same as above)	33
	IMSEAR(same as above)	20
	AIM(same as above)	16

EBSCO
http://www.ebsco.com
No AGE, HUMAN or
YEAR limits applied 1990 – 2013
Searched in SUBJECTS, ABSTRACT and TITLE fields only across databases suite.

#	Searches	Results
1	(underreporting OR under reporting OR underreported OR under reported OR data quality OR official figures OR official national figures OR record linkage OR quality of information OR officially reported OR multiple sources OR linkage OR under registered OR under registration OR under registering OR underregistered OR underregistration OR under registering OR sources of error OR misclassification OR misclassified OR (errors AND registration) OR late maternal mortality OR confidential enquiries OR confidential enquiry OR (data collection AND (methods OR standards)) OR audit OR (population surveillance AND (methods OR standards))) AND (maternal mortality OR pregnancy related deaths OR maternal deaths)	2831 citations found. EBSCO system Duplicates removed

		remaining : 1019 See below for results per database
	Academic Search Complete	550
	Academic Search Premier	535
	CINAHL Complete	457
	CINAHL Plus with Full Text	428
	Health Source: Nursing/Academic Edition	121
	Women's Studies International	117
	Gender Studies Database	93
	Consumer Health Complete - EBSCOhost	81
	PsycINFO	80
	Food Science Source	61
	SocINDEX with Full Text	31
	MasterFILE Premier	30
	Business Source Complete	29
	Public Affairs Index	27
	Business Source Premier	26
	Environment Complete	18
	Psychology and Behavioral Sciences Collection	16
	Vocational and Career Collection	12
	MedicLatina	11
	Middle Eastern & Central Asian Studies	11
	Health Source - Consumer Edition	9
	Education Research Complete	9
	Agricola	8

	Peace Research Abstracts	8
	Alt HealthWatch	7
	Professional Development Collection	6
	Education Full Text (H.W. Wilson)	6
	International Security & Counter Terrorism Reference Center	5
	SPORTDiscus with Full Text	5
	Risk Management Reference Center	3
	Historical Abstracts	3
	Political Science Complete	3
	ERIC	2
	Computer Source	2
	Communication & Mass Media Complete	2
	Library, Information Science & Technology Abstracts with Full Text	2
	Computers & Applied Sciences Complete	2
	Associates Programs Source	2
	Vocational Studies Premier	2
	Caribbean Search	2
	Criminal Justice Abstracts with Full Text	2
	Biological & Agricultural Index Plus (H.W. Wilson)	2
	Legal Collection	1
	Bibliography of Native North Americans	1
	National Criminal Justice Reference Service Abstracts	1
	Central & Eastern European Academic Source	1
	Humanities Abstracts (H.W. Wilson)	1
	Humanities Full Text (H.W. Wilson)	1
	Total of citations found after duplicates removed	1019

Web of Science
http://www.webofknowledge.com
No AGE, HUMAN or YEAR limits applied.

#	Searches (TOPIC FIELD)	Results
1	(underreporting OR “under reporting” OR underreported OR “under reported” OR “data quality” OR “official figures” OR “official national figures” OR “record linkage” OR “quality of information” OR “officially reported “ OR “multiple sources” OR linkage OR “under registered” OR “under registration” OR “under registering” OR underregistered OR underregistration OR “under registering” OR “sources of error” OR misclassification OR misclassified OR (errors AND registration) OR “late maternal mortality” OR “confidential enquiries” OR “confidential enquiry” OR (“data collection” AND (methods OR standards)) OR audit OR (“population surveillance” AND (methods OR standards))) AND (“maternal mortality” OR “pregnancy related deaths” OR “maternal deaths”)	688

Popline
http://www.popline.org
.

#	Searches (TOPIC FIELD)	Results
1	(underreporting OR “under reporting” OR underreported OR “under reported” OR “data quality” OR “official figures” OR “official national figures” OR “record linkage” OR “quality of information” OR “officially reported “ OR “multiple sources” OR linkage OR “under registered” OR “under registration” OR “under registering” OR underregistered OR underregistration OR “under registering” OR “sources of error” OR misclassification OR misclassified OR (errors AND registration) OR “late maternal mortality” OR “confidential enquiries” OR “confidential enquiry” OR (“data collection” AND (methods OR standards)) OR audit OR (“population surveillance” AND (methods OR standards))) AND (“maternal mortality” OR “pregnancy related deaths” OR “maternal deaths”)	142

Web of Science
http://www.webofknowledge.com
No AGE, HUMAN or YEAR limits applied.

#	Searches (Web of Science – Russian Index)	Results
1	(«недостаток информации» OR «утерянные данные» OR «дефекты сбора данных» “несообщение” OR «сокрытие» OR «несообщённый» OR «сокрытый» OR «несообщённая» OR «сокрытая» OR «скрытый» OR «скрытая» OR «сокрытые» OR «скрытые» OR «качество информации» OR	

	«качество данных» OR «официальные данные» OR «официальные цифры» OR «официальная статистика» OR «национальная статистика» OR «национальные данные» OR «многочисленные источники» OR «множественные источники» OR «сцепленные данные» OR «связанные данные» OR «незарегистрированные» OR «незарегистрированный» OR «незарегистрированная» OR «не зарегистрированный» OR «не зарегистрированная» OR «не зарегистрированные» OR «отказ от регистрации» OR «регистрация не проводилась» OR «не регистрировалась» OR «не регистрировался» OR «не регистрировались» OR «причина ошибки» OR «источник ошибки» OR «причины ошибки» OR «источники ошибки» OR «причина ошибок» OR «источник ошибок» OR «причины ошибок» OR «источники ошибок» OR «ошибочная классификация» OR «ошибка классификации» OR «ошибка в классификации» OR «неправильная классификация» OR «неверная классификация» OR «неправильная группировка» OR «неверная группировка» OR «ошибка в группировке» OR «ошибочная группировка» OR «поздняя материнская смертность» OR «поздней материнской смертности» OR «позднюю материнскую смертность» OR «конфиденциальный запрос» OR «конфиденциальное расследование» OR «закрытая информация» OR «закрытые сведения» OR «закрыть информацию» OR «утаить информацию» OR «утаённая информация» OR («сбор информации» OR «сбора информации» AND («методы» OR «стандарты» OR «механизм» OR «техника» OR «алгоритм» OR «методика») OR «аудит» OR («надзор» OR «популяционный надзор» OR «здоровье населения» OR «состояние здоровья населения» OR «здоровье популяции»))	
2	(«материнская смертность» OR «акушерская смертность» OR «акушерско-гинекологическая смертность» OR «послеродовая смертность» OR «смерть в родах» OR «родовая смертность» OR «гибель рожениц» OR «гибель родильниц» OR «смертность рожениц» OR «смертность родильниц»)	
	1 AND 2	18

End of Box 2

APPENDIX B

BIVARIATE VECTOR AUTOREGRESSIVE PROCESS

A bivariate VAR process of lag 1, denoted VAR(1) is given by the following:

$$\boldsymbol{y}_t = \boldsymbol{c} + A\boldsymbol{y}_{t-1} + \boldsymbol{u}_t \tag{B.1}$$

in which $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is a 2×2 coefficient matrix representing the correlation between $\boldsymbol{y}_t$ and $\boldsymbol{y}_{t-1}$, and $\boldsymbol{u}_t$ is a 2×1 unobservable zero-mean innovation process (serially uncorrelated, ie. with time invariance covariance matrix Σ_u. In a zero-mean VAR(1) process the intercept terms are set to 0, ie $\boldsymbol{c} = (0,0)^T$.

$$\Sigma_u = \begin{bmatrix} \sigma_1^2 & \phi \cdot \sigma_1 \cdot \sigma_2 \\ \phi \cdot \sigma_1 \cdot \sigma_2 & \sigma_2^2 \end{bmatrix}$$

Written more simply, we have

$$\begin{aligned} y_{1_t} &= a_{11} y_{1_{t-1}} + a_{12} y_{2_{t-1}} + u_{1_t} \\ y_{2_t} &= a_{21} y_{1_{t-1}} + a_{22} y_{2_{t-1}} + u_{2_t} \\ E(u_t) &= 0 \\ E(u_t u_t^{'}) &= \Sigma_u \\ E(u_t u_s^{'}) &= 0 \text{ for } s \neq t \end{aligned}$$

If this iterative process starts at some time , ie. $t = 1$, we get

$$\begin{aligned}
\boldsymbol{y}_1 &= \boldsymbol{c} + A\boldsymbol{y}_0 + \boldsymbol{u}_1 \\
\boldsymbol{y}_2 &= \boldsymbol{c} + A\boldsymbol{y}_1 + \boldsymbol{u}_2 = \boldsymbol{c} + A(\boldsymbol{c} + A\boldsymbol{y}_0 + \boldsymbol{u}_1) + \boldsymbol{u}_2 \\
&= (I_2 + A)\boldsymbol{c} + A^2\boldsymbol{y}_0 + A\boldsymbol{u}_1 + \boldsymbol{u}_2, \\
&\vdots \\
\boldsymbol{y}_t &= (I_2 + A + \ldots + A^{(t-1)})c + A^t\boldsymbol{y}_0 + \sum_{i=0}^{t-1} A^i\boldsymbol{u}_{t-i}
\end{aligned}$$

Hence the vectors $\boldsymbol{y}_1, ..., \boldsymbol{y}_t$ are uniquely determined by $\boldsymbol{y}_0, \boldsymbol{u}_1, ..., \boldsymbol{u}_t$. The joint distribution of $\boldsymbol{y}_1, ..., \boldsymbol{y}_t$ is determined by the joint distribution of $\boldsymbol{y}_0, \boldsymbol{u}_1, ..., \boldsymbol{u}_t$.

B.0.0.1 Stationary Processes

The bivariate VAR(1) process is characterized as a stationary process, defined by time-invariance of the first and second moments, ie a stochastic process y_t is stationary if the following conditions are met:

$$E(y_t) = \mu \text{ for all } t, \tag{B.2}$$

and

$$E[(y_t - \mu)(y_{t-h} - \mu)'] = \Sigma_y(h) = \Sigma_y(-h)' \text{ for all } t \text{ and } h = 0, 1, 2, ... \tag{B.3}$$

Condition B.2 means that all y_t have the same mean μ, and condition B.3 means that the autocovariances of the process do not depend on t, but do depend on the time lag h. By this definition, the innovation process u_t is an example of a stationary process.

B.0.0.2 Unconditional expectation and autocovariances

(1) The unconditional expectation of a zero-mean stationary bivariate VAR(1):

$$\begin{aligned}
\boldsymbol{y}_t &= A\boldsymbol{y}_{t-1} + \boldsymbol{u}_t \\
E(\boldsymbol{y}_t) &= E(A\boldsymbol{y}_{t-1} + \boldsymbol{u}_t) \\
E(\boldsymbol{y}_t) &= AE(\boldsymbol{y}_{t-1}) + E(\boldsymbol{u}_t) \\
E(\boldsymbol{y}_t) &= AE(\boldsymbol{y}_{t-1}) + 0 \text{ zero-mean innovation process} \\
E(\boldsymbol{y}_t) &= E(\boldsymbol{y}_{t-1}) = \mu \text{ by defn of stationarity} \\
\mu &= 0 \text{ by zero-mean VAR(1) definition}
\end{aligned}$$

In more general form:

$$\mu = (I - A)^{-1}\boldsymbol{c}$$

(2) The unconditional autocovariance of VAR(1) process:

Let $\boldsymbol{y}_t = A\boldsymbol{y}_{t-1} + \boldsymbol{u}_t$ represent a stationary VAR(1) process for variable y with white noise covariance matrix $E(u_t u_t^{'}) = \Sigma_u$. The unconditional autocovariance is derived as follows:

$$E[(y_t - \mu)(y_{t-h} - \mu)^{'}] = AE[(y_{t-1} - \mu)(y_{t-h} - \mu)^{'}] + E[u_t(y_{t-h} - \mu)^{'}] \tag{B.4}$$

$$\Gamma_y(0) = A\Gamma_y(-1) + \Sigma_u = A\Gamma_y(1)^{'} + \Sigma_u \text{ for } h = 0 \tag{B.5}$$

$$\Gamma_y(h) = A\Gamma_y(h-1) \text{ for } h > 0 \tag{B.6}$$

For $h = 1$, we get from B.4, $\Gamma_y(1) = A\Gamma_y(0)$. Substitution $A\Gamma_y(0)$ for $\Gamma_y(1)$ gives:

$$\Gamma_y(0) = A\Gamma_y(0)A^{'} + \Sigma_u \tag{B.7}$$

Using the vector function, this can be written as

$$vec\Gamma_y(0) = vec(A\Gamma_y(0)A^{'}) + vec\Sigma_u \tag{B.8}$$

$$= (A \otimes A)vec\Gamma_y(0) + vec\Sigma_u \tag{B.9}$$

$$vec\Gamma_y(0) = (I - A \otimes A)^{-1} vec\Sigma_u \tag{B.10}$$

The conditional distribution given by B.1 is applied to the stationary time series for $t = 2, ..., T$. However, at initial time t_0, ie. $t_0 = 1$, we use the unconditional expectation and covariance to apply stochastic error.

$$\boldsymbol{y}_t | \boldsymbol{y}_{t-1} \sim N_2(A\boldsymbol{y}_{t-1}, \Sigma_u, \text{ for } t = 2, ..., T$$

$$\boldsymbol{y}_1 \sim N_2(\boldsymbol{\mu}, \Gamma_y(0))$$

...

BIBLIOGRAPHY

[1] L. Alkema, S. Zhang, D. Chou, A. Gemmill, A.B. Moller, D. Ma Fat, L. Say, C. Mathers, D. Hogan (2017). "A Bayesian Approach to the Global Estimation of Maternal Mortality." *The Annals of Applied Statistics* 11(3):1245-74. doi:10.1214/16-aoas1014.

[2] L. Alkema, D. Chou, D. Hogan, S. Zhang, A.B. Moller, A. Gemmill, D.M. Fat, T. Boerma, M. Temmerman, C.D. Mathers, L. Say (2015). "Global, regional, and national levels and trends in maternal mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Inter-agency Group for Maternal Mortality Estimation". *The Lancet* 387(10017): 462 - 474.

[3] S. R. Asmussen (2003). "Markov Chains". Applied Probability and Queues. Stochastic Modelling and Applied Probability. 51. pp. 3–8. $doi : 10.1007/0 - 387 - 21525 - 5\_1.ISBN978 - 0 - 387 - 00211 - 8$.

[4] J. Bennett and J. Wakefield (2001). "Errors-in-variables in Joint Population Pharmacokinetic/Pharmacodynamic Modeling". *Biometrics* 57(3): 803-812.

[5] Caswell, H. 2001. Matrix Population models, Second Edition. Sinauer Associates, Inc. Sunderland, MA.

[6] F. Chao and L. Alkema (2013). "How informative are vital registration data for estimating maternal mortality? A Bayesian Analysis of WHO adjustment data and parameters". *Statistics and Public Policy* 3(2) 6-18.

[7] H. Chu, Z. Wang, S. Cole, S. Greenland (2006). "Sensitivity analysis of misclassification: A graphical and a Bayesian Approach". *The Annals of Epidemiology* (16) 834-841.

[8] Core Team, R. (2014). R: *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

[9] R. Eisinga, M. Grotenhuis, B.Pelzer. "Saddlepoint approximations for the sum of independent non-identically distributed binomial random variables". Statistica Neerlandica, 67(2):190–201, May 2013.

[10] A Gelman (2006)."Prior distributions for variance parameters in hierarchical models".*Bayesian Analysis* 1(3):515-533

[11] S. Helleringer, G. Pison, A. Kante, G. Duthe, A. Andro (2014). "Reporting errors in siblings' survival histories and their impact on adult mortality estimates: Results from a record linkage study in senegal". *Demography* (51): 387-411.

[12] K. Hill, Y. Choi, I. Timaeus (2005).“Unconventional approached to mortality estimation”. Demographic Research. (13) 281-300. doi: 10.4054/DemRs.2005.13.12

[13] Institute of Health Metrics and Evaluation (2014).“Global Burden of Disease: Massive shifts reshape the health landscape worldwide”.

[14] Leslie, P. H. 1945. On the use of matrices in certain population mathematics. Biometrika 33: 183–212.

[15] B. Liu, T. Quertermous.“Approximating the sum of independent non-identical binomial random variables”. *arXiv*

[16] H. Lütkepohl(2005). New Introduction to Multiple Time Series Analysis. Berlin: Springer. ISBN 3-540-40172-5.

[17] B. Masquelier (2013). “Adult mortality from sibling survival data: A reapprasial of selection biases”. Demographuc, 50, 207-228.

[18] E. Peterson, D. Chou, AB. Moller, A. Gemmill, L. Say, L. Alkema (2019). “Estimating maternal mortality using data from national civil registration vital statistics systems: A Bayesian hierarchical bivariate random walk model to estimate sensitivity and specificity of reporting.” *arXiv*

[19] M. Plummer (2014). “Cuts in Bayesian graphical models”. *Statistical Computing* 25: 37-43.

[20] S. Preston, P Heuveline, M Guillot.“Demography: Measuring and Modeling Population Processes”. Blackwell Publishing Ltd. (2001)

[21] United Nations Maternal Mortality Estimation Inter-Agency Group (2015). Trends in maternal mortality: 1990 to 2015. Estimates by WHO, UNICEF, UNFPA, The World Bank and the United Nations Population Division. Geneva, World Health Organization, 2015.

[22] United Nations Maternal Mortality Estimation Inter-Agency Group (2019). Trends in maternal mortality: 2000 to 2017. Estimates by WHO, UNICEF, UNFPA, The World Bank and the United Nations Population Division. Geneva, World Health Organization, 2019.

[23] J. Wilmoth, N. Mizoguchi, M. Oestergaard, L. Say, C. Mathers, S. Zureick-Brown, M. Inoue, D. Chou (2012) ”A New Method for Deriving Global Estimates of Maternal Mortality,” *Statistics, Politics, and Policy*: Vol.3: Is. 2, Article 3. DOI: 10.1515/2151-7509.1038

[24] World Health Organization. (2010). International classification of diseases for mortality and morbidity statistics (10th Revision).