

How Informative are Vital Registration Data for Estimating Maternal Mortality?

A Bayesian Analysis of WHO Adjustment Data and Parameters

Fengqing Chao and Leontine Alkema *

Extended abstract for XXVII IUSSP International Population Conference 2013, August 14, 2013

Abstract

Monitoring maternal mortality is challenging due to fragmented data of varying quality. The maternal mortality estimates published by the WHO in 2012 included data adjustment parameters to account for these data quality issues, but there was a discrepancy between the WHO assumption about, and the observed variability in, misclassification errors in vital registration (VR) observations.

We developed a Bayesian hierarchical time series model to estimate the extent of VR misclassification errors and to provide a plausible assessment of the uncertainty associated with VR observations for countries with and without external information on VR adjustment parameters. The resulting Bayesian distribution for VR adjustments was more comparable to the observed biases than the WHO expert distribution and the model allows for estimation of VR adjustment values for any period of interest for countries with partial information on such adjustments. We also illustrated that a fully Bayesian modeling approach for estimating maternal mortality can provide more data-driven insights into maternal mortality estimates and data adjustment parameters. However, given the paucity of, and the issues with, maternal mortality data, validation of modeling assumptions and findings is challenging; more data collection and research on measuring maternal mortality and assessing data quality are needed.

Keywords: Autoregressive time series model; Maternal mortality; Multilevel model; Proportion of maternal deaths among all deaths to women of reproductive ages (PM); Vital registration (VR); World Health Organization (WHO).

*Department of Statistics and Applied Probability and Saw Swee Hock School of Public Health, National University of Singapore, Singapore. The authors are grateful to John Wilmoth, Alex Cook, Nasheen Naidoo, Jin Rou New and Alison Gemmill for feedback on earlier versions of this manuscript, and to Doris Chou, Ann-Beth Moller and Lale Say for assistance with the data and discussions of this study.

1 Introduction

Maternal mortality is widely considered as a sentinel indicator of the quality of a health care delivery system and as a key indicator of population health and social and economic development (Wilmoth et al., 2012). Millennium Development Goal 5 (MDG 5) calls for a reduction in the maternal mortality ratio (MMR) by three quarters between 1990 and 2015. To measure progress, the World Health Organization (WHO), the United Nations Children’s Fund (UNICEF), the United Nations Population Fund (UNFPA) and the World Bank published estimates of maternal mortality in 2012, referred to hereafter as the WHO estimates (WHO et al., 2012).

It is challenging to estimate maternal mortality due to the paucity of accurate data, especially in developing countries where maternal mortality is high and such estimates are most needed. The WHO estimates were based on limited data on the proportion of maternal deaths among all deaths of women of reproductive age, adjusted to account for data issues such as under-reporting, misclassification and inconsistent definitions. The adjustments and the uncertainty associated with the adjustment were based on external data and/or expert opinion.

In this paper, we assessed the accuracy of the estimates and probability distributions used for the vital registration (VR) misclassification parameters used in the WHO modeling approach. We identified two issues: (1) it is not clear how information on varying periods for countries with information should be summarized into the periods needed for the maternal mortality estimation, and (2) for countries without additional information on the extent of misclassification of maternal deaths in the VR, the WHO expert distribution may understate uncertainty therein. To improve upon the current WHO modeling approach, we developed a Bayesian estimation approach for the VR misclassification adjustments for countries with or without additional information on the quality of the VR data. The resulting probability distribution for VR adjustments was compared to those from the WHO and adjustments published by Naghavi et al. (2010), used by the Institute for Health Metrics and Evaluation for constructing global estimates of maternal mortality (Lozano et al., 2013). We also incorporated the Bayesian VR adjustment into the current WHO estimation approach and implemented a fully Bayesian maternal mortality estimation model to examine the effect of the differences in VR adjustment parameters on the maternal mortality estimates for selected countries.

The paper is organized as follows. We first briefly summarize the WHO estimation method, and introduce our alternative estimation approach in Section 2. We then present results in Section 3 and end with a discussion of findings.

2 Methods

2.1 Summary of WHO estimation method

The WHO maternal mortality estimation methods used are described in detail elsewhere (Wilmoth et al., 2012, WHO et al., 2012). We summarize the method here.

The key indicator in the WHO estimation approach is the proportion of maternal deaths (PM)

among all deaths of women of reproductive ages. Estimation of PM is complicated in countries with HIV/AIDS epidemics because of the difficulties in determining whether a death of a woman who was HIV-positive and died during the maternal risk period should be counted as a maternal death. In the WHO method, the “total” PM is estimated as $PM = (1 - a)PM^{na} + aPM^a$, where a refers to the proportion of AIDS deaths among all deaths to women of reproductive ages, PM^{na} is the non-AIDS PM (the proportion of non-AIDS maternal deaths among the total number of non-AIDS deaths of women of reproductive ages), and PM^a is the AIDS PM (the proportion of AIDS maternal deaths among the total number of AIDS deaths to women of reproductive ages). This paper focused on the estimation of PM^{na} , whereas PM^a and a are estimated from other sources and not the subject of this paper.

Let z_i denote the observed total PM for observation i . Data are available from VR systems and other sources such as household surveys (described in detail in the 2012 WHO report (WHO et al., 2012)). The WHO analysis started by carrying out an adjustment procedure resulting in an adjusted PM, denoted by y_i , which represents the proportion of non-AIDS maternal deaths among all deaths of women of reproductive ages, given by:

$$y_i = (z_i \cdot \gamma_i - s_i)q_i, \quad (1)$$

where s_i refers to an adjustment related to AIDS deaths and q_i refers to an adjustment for observations that reported to be pregnancy-related deaths (as opposed to maternal deaths). Parameter γ_i is an under-reporting or misclassification parameter and is determined by the data source of observation i . For VR observations, γ_i quantifies the extent to which maternal deaths have been misclassified. For example, if the observed proportion of maternal deaths among all deaths to women of reproductive ages is underestimated by 50%, $\gamma_i = 1.5$. Adjustment parameters were informed by external studies.

In the WHO approach, for countries with sufficient information from vital registration systems from 1990 to 2012, PM^{na} estimates were based on adjusted PM data. For the countries without sufficient VR information, estimates were obtained from a multi-level model which was fitted to the adjusted data from all countries combined. The multi-level model was given by

$$\begin{aligned} \log(y_i) &\sim N(\phi_i, \sigma_i^2), \\ \phi_i &= \log(1 - a_i) + \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \beta_3 x_{3,i} + \alpha_{c[i]}^C + \alpha_{r[i]}^R, \\ \alpha_j^C &\sim N(0, \sigma_{ac}^2), \\ \alpha_k^R &\sim N(0, \sigma_{ar}^2), \end{aligned} \quad (2)$$

where $\log(1 - a_i)$ is an offset to remove the AIDS deaths from the denominator of the adjusted PM, x_1 , x_2 and x_3 are predictors for the PM (referring to the log of the general fertility rate, the log of GDP per capita and the proportion of births with a skilled birth attendant), and $\alpha_{c[i]}^C$ and $\alpha_{r[i]}^R$ refer to the country and region-specific intercepts respectively. The uncertainty assessment of the resulting non-AIDS PM estimates was based on repeated draws from probability distributions

on the adjustment parameters, and repeated fitting of the multilevel model to the resulting sets of adjusted PM values.

The probability distribution for the VR adjustment parameters was given by:

$$\log(\gamma_i) \sim N(\log(V_i) - 0.5 \cdot 0.05^2, 0.05^2), \quad (3)$$

where the mean value V_i was a country-specific value for a subset of countries where country-specific studies were available, while for the remaining 63 countries with VR data but without external information on the VR adjustment, $V_i = 1.5$, based on the median of reported VR adjustments. The standard deviation in Eq.(3) was based on expert opinion.

Figure 1 illustrates the distribution of 35 reported VR misclassification errors, or in other words, reported VR adjustment parameters, obtained from various studies in 19 countries. The data set, based on Appendix 1 from the WHO 2012 report and additional information provided by WHO analysts, is given in Table 2 in the Appendix. The observed adjustments range from 0.95 (from New Zealand with observation period 2006-2008) to 2.03 (from Australia with observation period 2003-2006). The VR adjustment data are from 19 countries and 9.7 country-years of data are available for each country on average. Figure 2 illustrates the observed VR adjustments for all 19 countries.

Limitations of the current WHO adjustment method Country-specific observation periods for VR adjustments vary and generally do not coincide with the observation periods that are used to calculate the observed PM in the WHO multilevel model (which are generally 5-year periods). Additionally, the adjustments are often not available for the entire VR observation period. As a consequence, the WHO VR adjustments for countries with external information (the V_i s in Eq.(3)) are often based on partial information. The procedure to impute adjustments for missing observation years has not been formalized.

The WHO expert distribution for countries without external information on VR adjustment parameters (Eq.(3) with $V_i = 1.5$) is added to Figure 1. A comparison of observed adjustments and the WHO expert distribution reveals that the expert distribution understates the variability of observed adjustment factors: if this expert distribution were used for countries where information *is* available, the variability in adjustments would be underestimated.

2.2 A Bayesian model for VR misclassification parameters

To overcome the limitations of the current WHO adjustment procedure, we developed an alternative model for VR misclassification parameters that provides annual estimates of the VR adjustment parameter for all countries, based on the available data on such adjustments, to (i) impute VR adjustments for countries with external information for a subset of observation years, and (ii) provide a more plausible representation of the extent of VR misclassification bias for countries without external information.

A Bayesian hierarchical time series model is used to model the underlying true (but most often unknown) VR misclassification parameters. This type of model set-up was motivated by the need

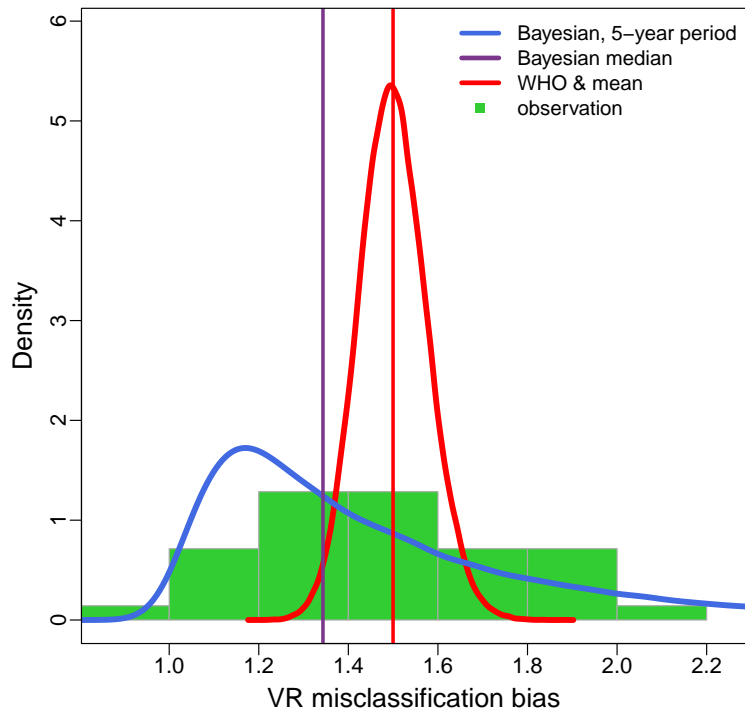


Figure 1: **Observed and estimated VR misclassification biases.** Histogram of observed VR misclassification biases and density functions for the WHO and Bayesian VR adjustment for a 5-year observation in a country with no external information on the VR misclassification error.

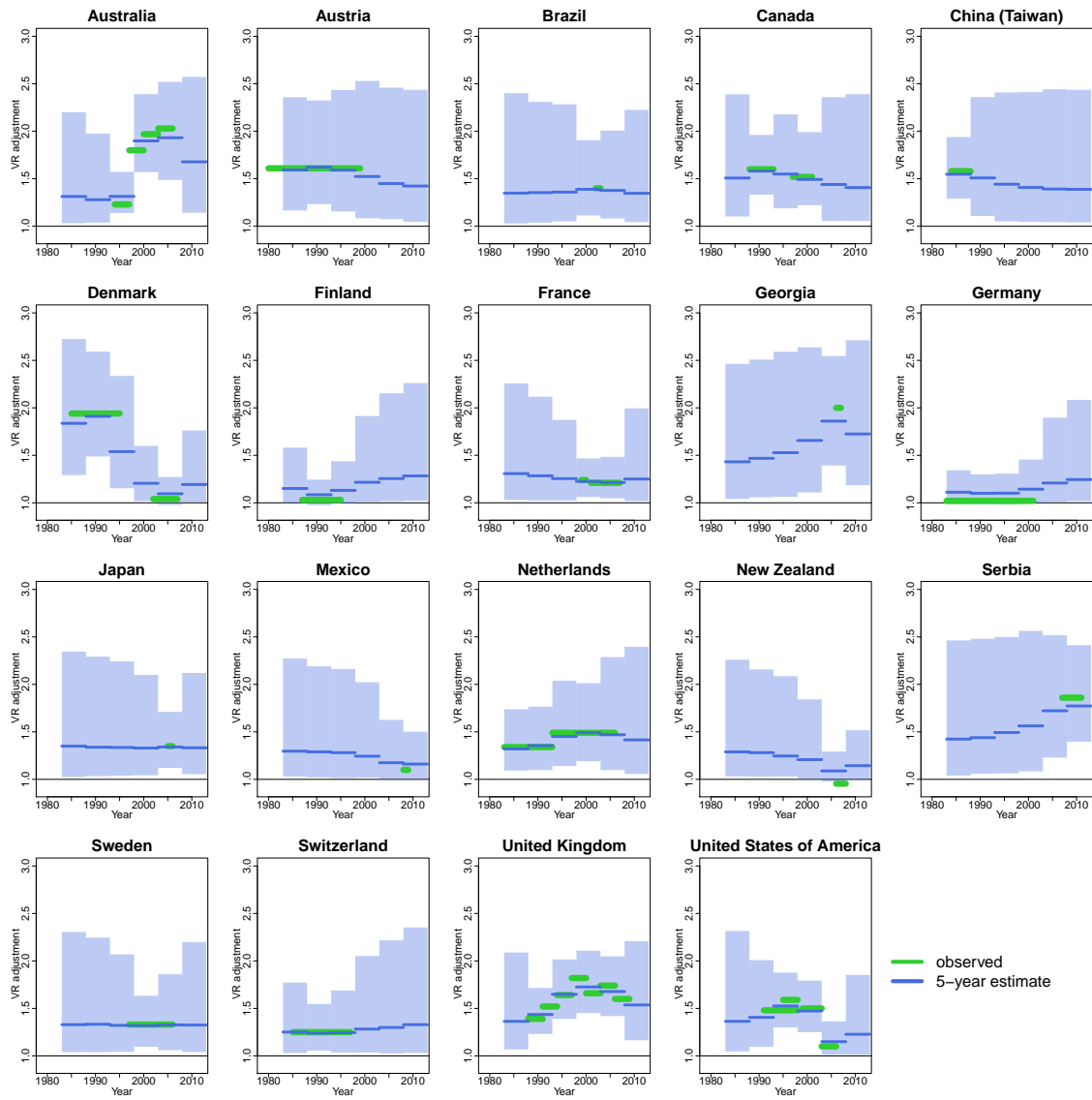


Figure 2: **Observed VR adjustments and Bayesian estimates for countries with external information on VR misclassification biases.** The green line segments represent the observed VR adjustments. Bayesian posterior median estimates for the VR adjustments for 5-year periods (the default periods used in the WHO modeling approach) are added in blue and blue shades represent 95% credible bounds. The line span 10 years corresponds to the observation period.

(i) to deal with observation periods of varying lengths, (ii) to allow for variation in misclassification errors between countries and within countries over time (e.g., the observations in Figure 2 suggest that the average adjustment in the United Kingdom may be higher than in Finland or Germany, and that the adjustment has changed over time in several countries), (iii) to minimize the number of model parameters given the limited number of observed misclassification outcomes. Technical details of the model specification are discussed in the remainder of this subsection.

Let $P_{c,t}$ be the proportion of maternal deaths that are correctly reported as maternal deaths in country $c = 1, \dots, C$ during year t , i.e., $P_{c,t} = \frac{R_{c,t}}{M_{c,t}}$, where $R_{c,t}$ is the reported number of maternal deaths for the country-year while $M_{c,t}$ is the true number of maternal deaths. The $P_{c,t}$'s were modeled with an autoregressive time series model of order one (AR(1)) with truncation:

$$P_{c,1978} \sim TN_{[1/3,1]} \left(p_c, \frac{\sigma_{AR}^2}{1 - \rho^2} \right),$$

$$P_{c,t} \sim TN_{[1/3,1]} (p_c + \rho(P_{c,t-1} - p_c), \sigma_{AR}^2), \text{ for } t = 1979, \dots, 2012,$$

where $TN_{[A,B]}(a, b^2)$ denotes a truncated normal distribution with mean a and variance b^2 , truncated to lie between A and B , such that the proportion of correctly reported deaths is assumed to be at least 1/3. The global time series parameters are given by the autoregressive parameter $0 \leq \rho < 1$ and variance σ_{AR}^2 with priors $\rho \sim U(0, 1)$ and $\sigma_{AR} \sim U(0, 0.5)$. The proportion of correctly reported maternal deaths fluctuates around the country-specific mean parameter p_c . This parameter is assumed to be drawn from a common truncated normal distribution,

$$p_c \sim TN_{[1/2,1]} (w, \sigma_p^2),$$

where w refers to the global mean, and σ_p^2 to the variance, with priors $w \sim U(1/2, 1)$ and $\sigma_p \sim U\left(0, \frac{1/2}{\sqrt{12}}\right)$. The truncation on the country-level under-reporting means p_c 's and their global mean w are based on the prior assumption that the average proportion of maternal deaths that are correctly reported in a given year is between 1/2 and 1. Similarly, the upper bound for the prior on σ_p is given by the standard deviation of a $U(1/2, 1)$ distribution, based on the assumption that the country-specific p_c 's are at most as spread out as this distribution.

The model has only four ‘‘global’’ parameters (the autoregressive parameters and the hierarchical mean and variance of the country means) but allows for differences within countries over time through the time series set-up and for differences between countries through the hierarchical model for mean reporting levels.

Inference The AR(1) model parameters were estimated using the 35 observed VR adjustments for various country-periods from Table 2. Let $W_{c[i],k[i]} = V_i$, the observed VR adjustment for country $c[i]$ and period $(t_{c[i],k[i]}, t_{c[i],k[i]} + T_{c[i],k[i]} - 1)$, where $T_{c,k}$ refers to the number of observation years for the k -th observation in country c . The observed VR adjustment relates to the true

number of maternal and reported maternal deaths as follows:

$$W_{c,k} = \frac{\sum_{t=t_{c,k}}^{t_{c,k}+T_{c,k}-1} (M_{c,t} + M_{c,t}^*)}{\sum_{t=t_{c,k}}^{t_{c,k}+T_{c,k}-1} (R_{c,t} + R_{c,t}^*)},$$

where $M_{c,t}^*$ and $R_{c,t}^*$ denote reporting errors, which are assumed to be small as compared to $M_{c,t}$ and $R_{c,t}$ (i.e., the errors represent an increase or decrease in the number of (reported) maternal deaths because of misreporting of the calendar year of death). If the number of maternal deaths does not vary greatly during a VR observation period, $\frac{1}{W_{c,k}} \approx \frac{1}{T_{c,k}} \sum_{t=t_{c,k}}^{t_{c,k}+T_{c,k}-1} P_{c,t}$. Without additional information on the extent of the difference between $\frac{1}{W_{c,k}}$ and $\frac{1}{T_{c,k}} \sum_{t=t_{c,k}}^{t_{c,k}+T_{c,k}-1} P_{c,t}$, we assumed

$$\frac{1}{W_{c,k}} \sim TN_{[0,\infty)} \left(\frac{1}{T_{c,k}} \sum_{t=t_{c,k}}^{t_{c,k}+T_{c,k}-1} P_{c,t}, \sigma_W^2 \right), \quad (4)$$

where $\sigma_W \sim U(0, 0.05)$, i.e., we assume that the standard deviation of the differences is at most 5%.

The effect of the VR adjustment modeling procedure on the WHO maternal mortality estimates from the multi-level model was assessed by replacing the WHO point estimates for the adjustment parameters by the posterior medians from the Bayesian VR adjustment model (obtained via Eq.(4) for the relevant country-periods), and similarly, by replacing the draws of the adjustment parameters from the WHO probability distributions by the posterior samples from the Bayesian VR adjustment model when carrying out the uncertainty assessment.

Validation Validation of the proposed model is challenging because of the paucity of data on VR adjustments. We compared the resulting posterior distribution for 5-year adjustments to the prior distribution to check whether unexpected findings were driven by the prior assumptions and model structure or informed by the data. We also compared the observed adjustments for the countries with external information to the adjustments that would have been obtained if the Bayesian and WHO approaches for countries without external information would have been used.

2.3 A Bayesian estimation model for maternal mortality

In the previous section, we proposed a Bayesian model for the VR adjustment parameters, that was fitted to observed VR adjustments and used to replace the current WHO point estimates and probability distributions to assess the effect of the VR adjustment model on the maternal mortality estimates.

Instead of “plugging in” the Bayesian estimates of the VR adjustments into the WHO maternal mortality model, alternatively, VR adjustments and maternal mortality can be estimated simultaneously. Such a combined Bayesian estimation model could provide more accurate estimates and

uncertainty assessments of both VR adjustments as well as maternal mortality because all relevant information is taken into account simultaneously. For example, suppose that in a country without external information on the extent of misclassification in the VR, the VR data as adjusted by the “default” Bayesian point estimates (the estimates that would result from the model as discussed in the previous section) are far below adjusted data from alternative sources and/or the expected level based on the region and the country’s predictors for PM. Such findings indicate that the extent of under-reporting in the VR may be greater than the default adjustment. When estimating the VR adjustment and PM simultaneously, the posterior distribution of the VR adjustment parameters would reflect this possibility, and higher PM point estimates and associated uncertainty bounds may be obtained.

To illustrate this approach and incorporate the estimation of adjustment parameters in the PM estimation procedure, we fitted a Bayesian model that combined the WHO multi-level model from Section 2.1 with the Bayesian VR adjustment model from Section 2.2. We focused our PM analysis on countries for which the maximum proportion of AIDS deaths among all deaths of women aged 15 to 49 in the population is smaller than 0.05 (102 countries, here referred to as the non-AIDS countries) and did not carry out any AIDS-related adjustments in the estimation procedure. The main reason for focusing on the non-AIDS countries was to avoid the difficulties associated with the AIDS adjustments and to discuss only one type of adjustment. Leaving out s_i from Eq.(1), the data model for observed PM for the non-AIDS countries is simplified as $y_i = z_i \cdot \gamma_i \cdot q_i$ such that

$$\log(z_i) \mid \phi_i, \gamma_i, q_i, \sigma_y \sim N(\eta_i, \sigma_y^2),$$

where $\eta_i = \phi_i - \log(\gamma_i) - \log(q_i)$, and ϕ_i is given by Eq.(2) where $a_i = 0$. Prior distributions for the multi-level model parameters were chosen to be spread out, with the exception of the non-VR adjustment parameters that determine γ_i and q_i for non-VR observations. For these parameters, priors were based on the WHO expert distributions. More details and the full model are provided in the Appendix in Section 5.2. The results from the Bayesian model were compared to the results from a modified WHO model. The modified WHO model is based on the original WHO model, but modified to leave out all AIDS adjustments (as explained for the Bayesian model) and fitted to non-AIDS countries only, such that its results were directly comparable to the results from the Bayesian model. Given that the models were fit to non-AIDS countries only, the results are for illustrative purposes only (to illustrate the insights that can be obtained from this approach w.r.t. VR adjustments).

2.4 Computation

A Markov Chain Monte Carlo (MCMC) algorithm was used to obtain samples from the posterior distributions of the model parameters for the VR adjustment model in Section 2.2 and for the fully Bayesian maternal mortality estimation model in Section 2.3.

In the VR adjustment model and in the Bayesian maternal mortality estimation model, 75,000 iterations were used for each of the three MCMC chains. We thinned for every 10 iterations and

discarded the first 20,000 iterations. Convergence was checked through visual inspection of trace plots and convergence diagnostics of Gelman and Rubin (1992).

Models were implemented in R 2.15 (R Development Core Team, 2011) and JAGS 3.3.0 (Plummer, 2003). The WHO model used R-package lme4 (Bates and Maechler, 2011) and the Bayesian model used packages rjags (Plummer, 2011) and coda (Plummer et al., 2006).

3 Results

3.1 Bayesian VR adjustment estimates

The results in this section are based on the Bayesian VR adjustment model from Section 2.2. Prior and posteriors for selected model parameters are shown in Figure 3 and Table 1 summarizes the posterior medians and 95% credible intervals (CIs) for the five model parameters in the AR(1) model. The posterior distribution of w , the global mean of the p_c 's (the country-specific mean parameters for the proportion of accurately reported maternal deaths), suggests that values below 0.6 are unlikely (the posterior probability that $w < 0.6$ is 1.5%). The posterior and prior for σ_p (the standard deviation of the p_c 's) are very similar, suggesting that there is little information in the data to estimate this parameter. The posterior for time series parameter ρ indicates high autocorrelation in the time series. The posterior distribution of σ_W suggests the standard deviation is close to its upper bound of 5%.

Table 1: Prior distributions and posterior estimates for the Bayesian VR adjustment model. Posterior estimates are given for the VR adjustment model from Section 2.2 and the PM estimation model from Section 2.3. $W_{c,k}$ refers to the VR adjustment for a 5-year observation in a country with no external information on the VR misclassification error.

Parameter	Prior	Posterior median estimate (95%CI)	
		VR adj. model	PM est. model
w	$U(1/2, 1)$	0.81 (0.63, 0.99)	0.77 (0.53, 0.98)
ρ	$U(0, 1)$	0.94 (0.76, 0.99)	0.97 (0.90, 1.00)
σ_{AR}	$U(0, 0.5)$	0.08 (0.05, 0.15)	0.08 (0.05, 0.12)
σ_p	$U\left(0, \frac{1/2}{\sqrt{12}}\right)$	0.081 (0.004, 0.142)	0.079 (0.005, 0.142)
σ_W	$U(0, 0.05)$	0.045 (0.032, 0.050)	0.046 (0.034, 0.050)
$W_{c,k}$		1.34 (1.03, 2.40)	1.43 (1.03, 2.60)

The bottom-right graph of Figure 3 shows the prior (induced by the hierarchical time series model and the prior on the time series model parameters) and posterior distribution for VR adjustment $W_{c,k}$ for a country without external information, for an observation period of 5 years. The posterior VR adjustment is also added to Figure 1. While the Bayesian posterior distribution is

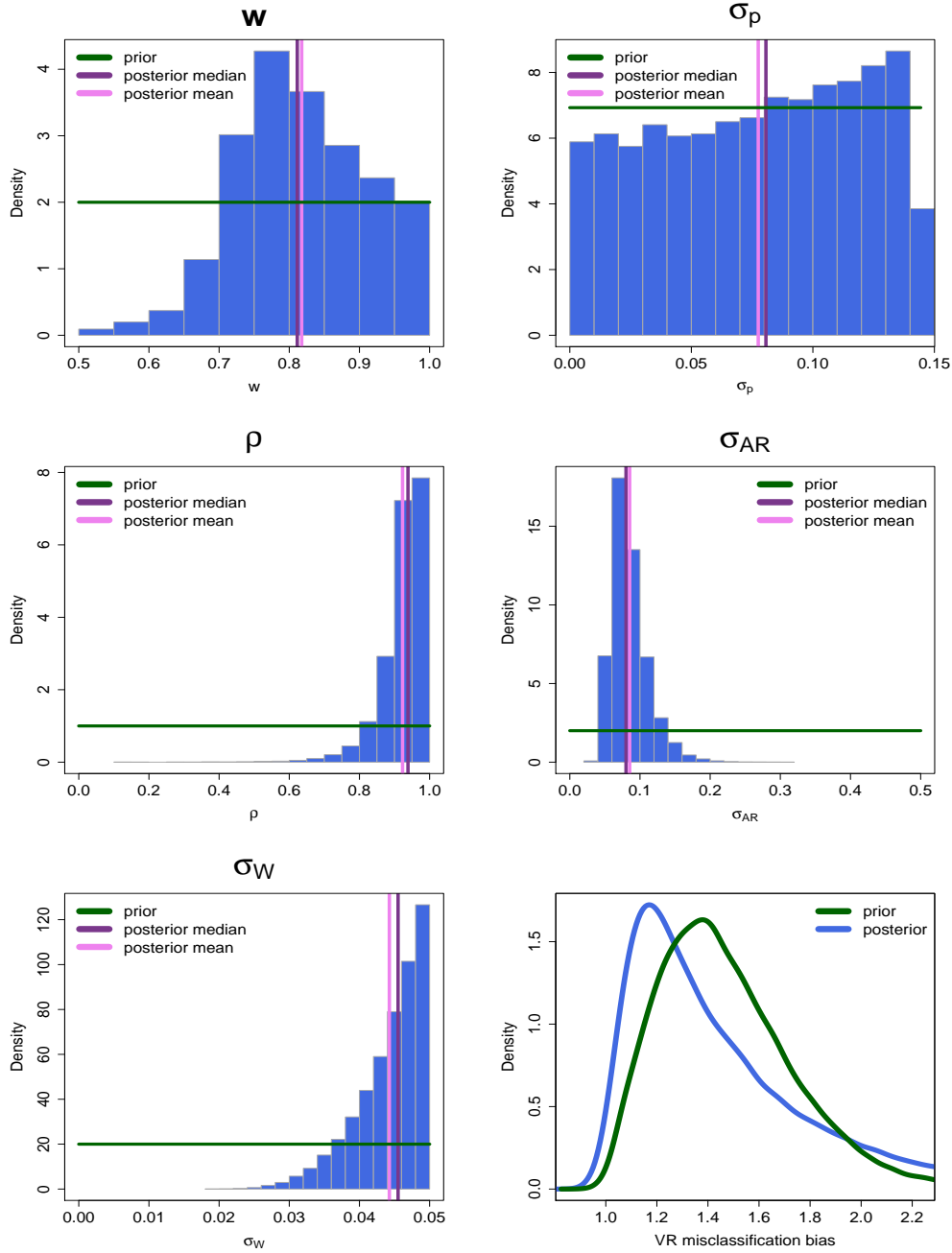


Figure 3: **Prior and posterior distributions of VR adjustment model parameters.** Posterior samples are displayed in the histogram. Priors are denoted by the horizontal green lines. Posterior medians and means are represented in the vertical purple and pink lines respectively. $W_{c,k}$ refers to the VR adjustment for a 5-year observation in a country with no external information on the VR misclassification error.

more weighted towards lower levels of under-reporting than the observed distribution of VR adjustments suggests (see Figure 1), its posterior median is smaller than the prior median (1.34 as compared to 1.43), suggesting that the estimated lower levels of under-reporting are data-driven as opposed to driven by prior model settings. The mismatch between the posterior distribution for the VR adjustment and the observations is explained by the fact that the posterior distribution is not directly comparable to the observed distribution because the observed distribution includes observations for time periods of various durations and some countries contributed multiple observations. In comparison to the WHO expert distribution, the Bayesian median of 1.34 is lower than the WHO point estimate of 1.50. The Bayesian distribution suggests that there is considerable uncertainty in VR adjustment parameters; the 95% credible interval (CI) for the VR adjustment value is given by (1.03, 2.40), as compared to (1.36, 1.65) for the WHO distribution.

Figure 2 shows the observed VR adjustments and 5-year estimates for countries with external information on the VR adjustment. The 5-year estimates refer to $W_{c,k}$ from Eq.(4), obtained from the country-specific proportions $P_{c,t}$ for the standard 5-year periods that are used in the WHO estimation approach. For the majority of countries with partial information, the VR adjustment remains quite uncertain for many country-periods. Figure 4 compares the observed adjustments for the same set of countries (with external VR adjustment information) to the adjustments that would have been obtained if the Bayesian and WHO approaches for countries without external information would have been used. While the observed adjustments are inside the WHO uncertainty bounds only for approximately one out of three observation periods (35% of the observations), the Bayesian uncertainty bounds contain the observed values for approximately four out five observation periods (79%).

Two sets of maternal mortality estimates are shown in Figure 5 for the countries with VR data where the multilevel model was used to construct the WHO PM estimates. The estimates were obtained from the original WHO modeling approach, and from the WHO approach with VR adjustments imputed from the Bayesian VR adjustment model. Differences between the WHO and (partially) Bayesian estimates are small, but as expected, the Bayesian estimates are slightly lower and their uncertainty bounds are slightly wider than those from the WHO model.

Comparison to alternative estimates Alternative global estimates of maternal mortality were published in 2011 (Lozano et al., 2011) and in 2012 (Lozano et al., 2013). In these studies, VR misclassification errors were obtained through a cause-of-death reclassification algorithm (Naghavi et al., 2010). The adjustment factors for maternal mortality that were displayed in Figure 6 in Naghavi et al. (2010) are illustrated in Figure 6. The corresponding Bayesian estimates ($W_{c,k}$) for a 1-year period are added for comparison. The two distributions differ slightly: while VR adjustments of around 1.1 are most likely in the Naghavi et al. distribution, the Bayesian mode is slightly higher and larger VR adjustments are more likely. The VR adjustment values from Naghavi et al. are not publicly available so we were not able to verify how the adjustments compare for the country-years that were included in our study.

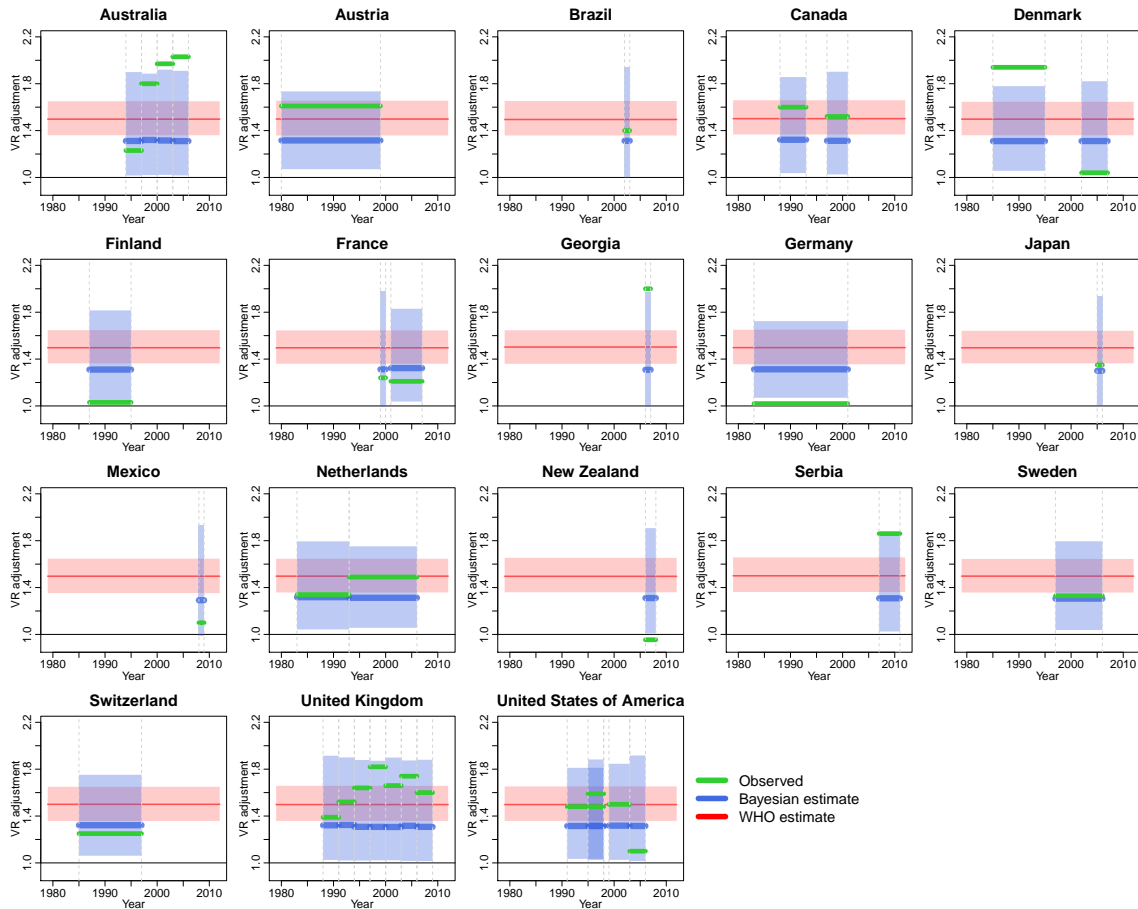


Figure 4: **Point estimates and uncertainty bounds for VR adjustments for countries with external information that would have been obtained if the Bayesian and WHO approaches for countries without external information would have been used.** Light green lines indicate observed VR adjustment factors. Blue lines and shades indicate the corresponding “no-external-information” Bayesian posterior median estimates and 95% credible intervals respectively. The red shades and lines indicate the “no-external-information” WHO mean estimates and 95% uncertainty bounds. The line span corresponds to the observation period.

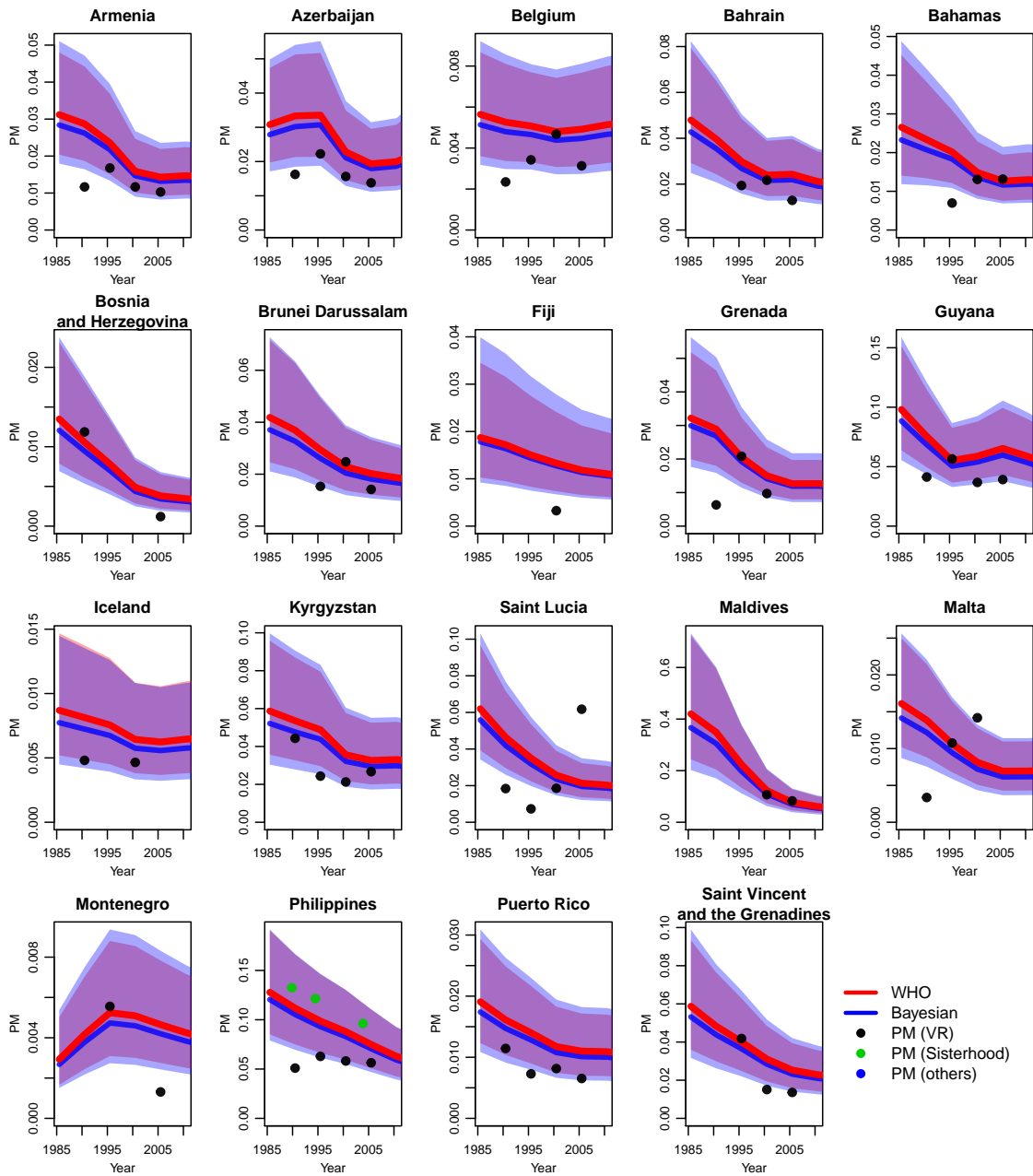


Figure 5: Maternal mortality (PM) estimates and 95% credible intervals for selected countries based on the WHO model (red) and the WHO model with Bayesian VR adjustment estimates (blue). Observations are displayed by source type. The selected countries are the countries with VR data but without external information on the VR misclassification, for which the WHO used a multi-level model for constructing PM estimates. The estimates from the WHO model with Bayesian VR adjustment estimates are referred to as “Bayesian”.

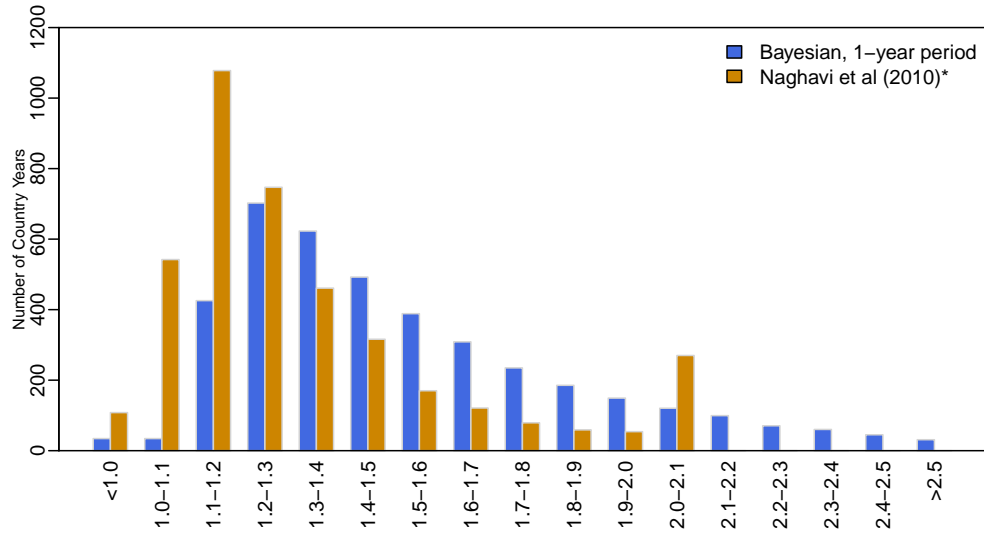


Figure 6: **Comparison of VR misclassification values from Naghavi et al. (2010) and the Bayesian VR adjustment model.** The Bayesian estimates are based on 1-year observation periods. *Values greater than 2 were aggregated in the Naghavi et al. results.

3.2 Bayesian maternal mortality estimates

The posteriors for the VR adjustment parameters resulting from the Bayesian maternal mortality model are similar to those resulting from fitting the VR adjustment model to the VR adjustment data only (see Table 1), except that the posterior for w , the global mean of the p_c 's (the country-specific mean parameters for the proportion of accurately reported maternal deaths), assigns greater probability mass to lower values of under-reporting, and the resulting distribution of $W_{c,k}$ for a 5-year period in a country without information suggests greater VR adjustments as compared to the estimates obtained from the VR adjustment model fitted to the VR adjustment data only. The finding that posteriors for the VR adjustment parameters resulting from the Bayesian maternal mortality model suggest more uncertainty towards greater VR adjustments may be caused by a selection bias of countries that provide information about VR adjustments; the VR adjustment may be lower in countries that have the resources to provide follow-up studies to investigate the accuracy of VR reporting as compared to countries without additional information.

Figure 7 illustrates the differences in estimated PM between the modified WHO and the fully Bayesian model for selected countries. The comparison shows that for about half of the countries, the Bayesian estimates are comparable to or slightly lower than the WHO estimates. Among the remaining countries, the Bayesian PM estimates are notably higher than the modified WHO estimates for Fiji, Philippines and Brunei Darussalam. For example, Fiji is the country with the largest adjustment (for under-reporting of maternal deaths) in the most recent observation period;

the Bayesian point estimate for the adjustment is 2.30 (95% CI given by 1.44 to 3.43), and the resulting PM estimate for the year 2000 is 20 per 1,000 (95% CI given by 9 to 39 per 1,000) as compared to 11 per 1,000 (95% CI given by 6 to 21 per 1,000) as given by the modified WHO model. Given that no VR adjustment data are available in Fiji, the Bayesian VR adjustment value is driven by the expected PM from the covariates in the multilevel model, the regional intercept and the estimated variability in country intercepts. External information is necessary to verify the accuracy of this VR adjustment.

Figure 8 shows the point estimates and 95% CIs for the VR adjustment parameters in the most recent observation period for all non-AIDS countries with VR data but without external information on VR adjustments, estimated in the fully Bayesian model. The Bayesian posterior estimates for VR adjustment parameters vary from 1.14 (for Saint Lucia) to 2.30 (for Fiji) and have much wider credible bounds as compared to the WHO expert distribution. The comparison in Figure 8 allows for identification of countries where recent VR adjustments are higher or lower than expected, as compared to the WHO adjustment of 1.5, that may warrant further investigation, as discussed for Fiji. For Saint Lucia, an increasing trend in VR data on PM is observed (see Figure 7), which contradicts the decrease in PM that would be expected based on changes in the predictors of PM. This results in a downwards adjustment of recent VR data in the Bayesian model. Again, country information is required to understand the specific situation in Saint Lucia and the accuracy of VR data.

4 Discussion

In this paper, we aimed to provide a plausible assessment of the extent of misclassification of maternal deaths in VR data and the uncertainty therein for countries where no external quantification of misreporting is available, and to capture time trends within countries with partial external information. We used a Bayesian hierarchical time series model to assess the extent of VR misclassification errors, which resulted in a distribution that is more comparable to the observed biases and increases the uncertainty that is associated with maternal mortality rates. A comparison of the Bayesian estimates, the WHO estimates and alternative estimates published by Naghavi et al. (2010) suggested that the Bayesian estimates are in between the two sets of estimates.

The inclusion of the Bayesian VR adjustment model in the WHO model for the proportion of maternal deaths illustrated in which country-periods greater or smaller misclassification errors are expected based on additional information on PM and model assumptions. It also showed that posteriors for the VR adjustment parameters resulting from the Bayesian maternal mortality model suggested slightly larger VR adjustments. This is possibly explained by a selection bias of countries that provide information about VR adjustments; the VR adjustment may be lower in countries that have the resources to provide follow-up studies to investigate the accuracy of VR reporting as compared to countries without additional information. These findings warrant further investigation to avoid underestimation of maternal deaths.

With the analysis, we illustrated that Bayesian modeling approaches can be used to provide

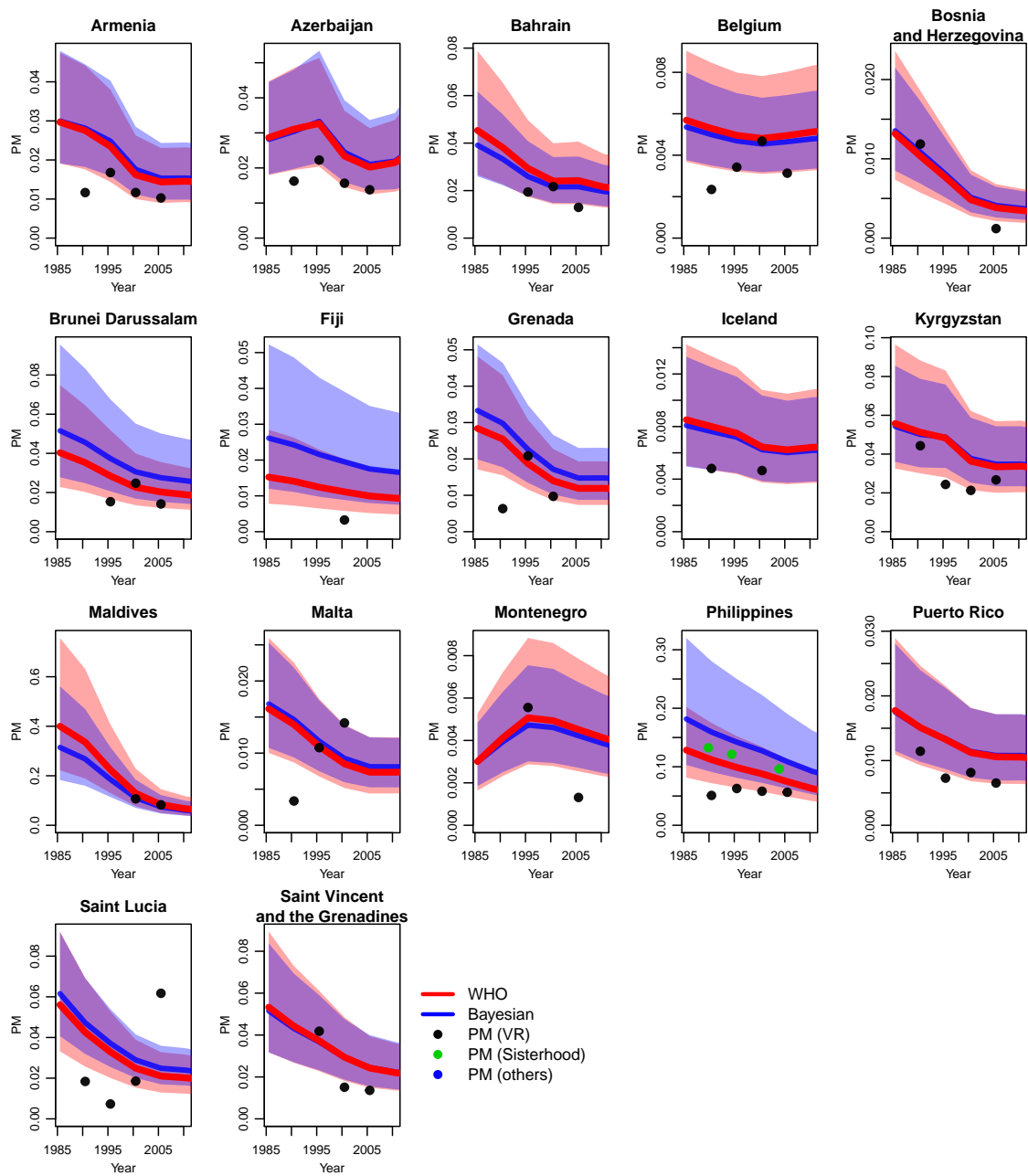


Figure 7: Maternal mortality (PM) estimates and 95% credible intervals for selected countries based on the modified WHO model (red) and the fully Bayesian model (blue). Observations are displayed by source type. The selected countries are the non-AIDS countries with VR data but without external information on the VR misclassification, for which the WHO used a multi-level model for constructing PM estimates.

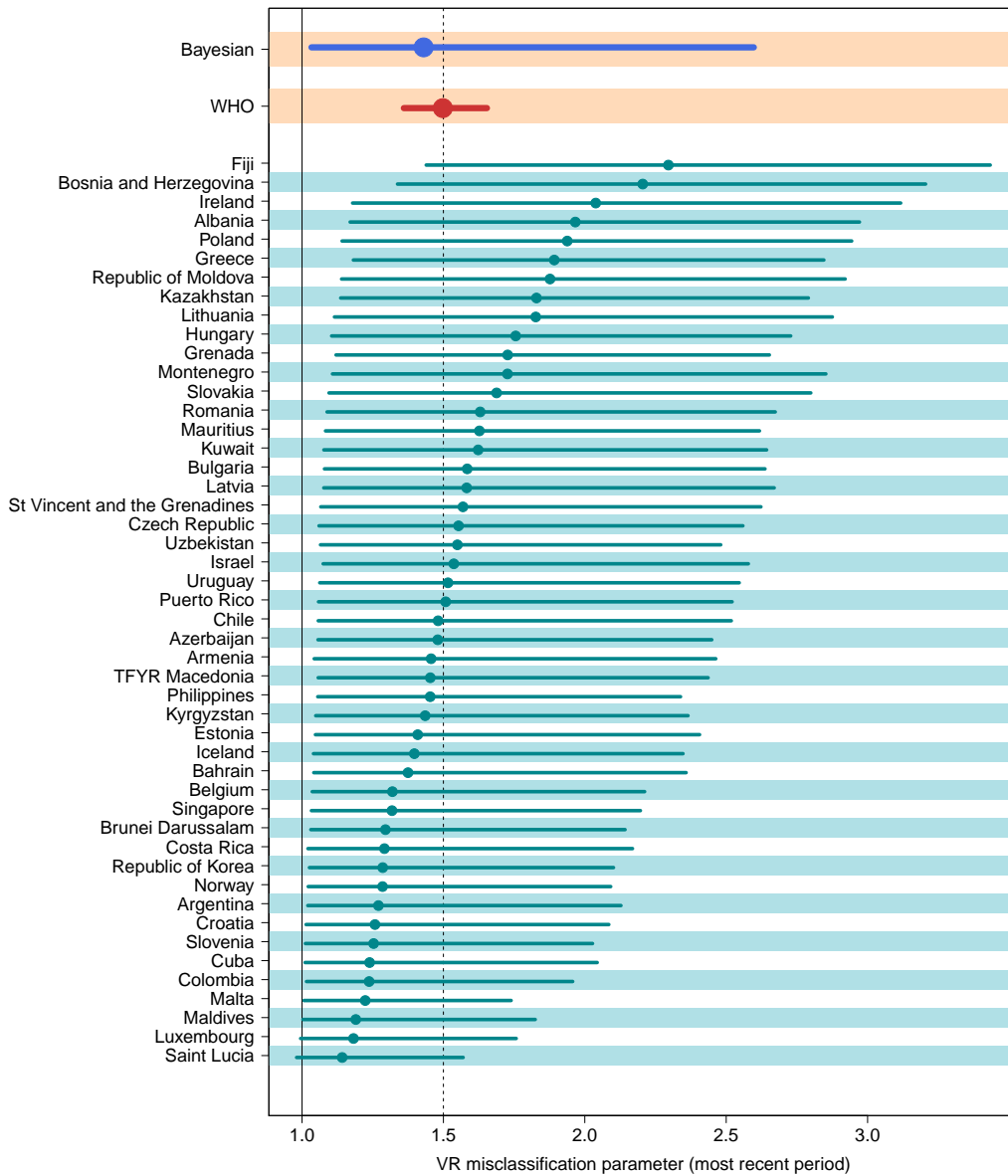


Figure 8: **Bayesian posterior VR adjustment parameters in the most recent observation period for selected countries.** Point estimates (dots) and 95% credible intervals (lines). The selected countries are all non-AIDS countries with VR data but without out external information on the VR misclassification. The posterior Bayesian adjustment for a “new country” without information on p_c and the WHO adjustment are added for comparison.

more objective and data-driven insights into maternal mortality estimates and data adjustment parameters. We verified that the Bayesian VR adjustment model would provide more plausible adjustment estimates for countries with external information on VR quality than the WHO model, if the information available in those countries would not be used to construct the estimates. Also, prior and posterior distributions in the Bayesian model were compared to communicate which parameters were most influenced by prior assumptions. Unfortunately, given the limited number of data points, formal cross-validation exercises could not be carried out. More data collection to assess VR data quality is needed to truly validate any VR adjustment modeling approach.

In this paper, we focused on the challenges in using VR data for estimating maternal mortality. Similar, or potentially greater challenges exist for using data from other sources. Instead of focusing more attention on statistical models for estimating maternal mortality, we call for more data collection and research to measure maternal mortality and assess data quality.

References

- Bates, D. and M. Maechler (2011). *lme4: Linear mixed-effects models using Eigen and Eigenfaces*. R package version 0.999375-42.
- Gelman, A. and D. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* 7, 457–511.
- Lozano, R., M. Naghavi, K. Foreman, S. Lim, K. Shibuya, V. Aboyans, J. Abraham, et al. (2013). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the global burden of disease study 2010. *The Lancet* 380(9859), 2095–2128.
- Lozano, R., H. Wang, K. J. Foreman, J. K. Rajaratnam, M. Naghavi, J. R. Marcus, L. Dwyer-Lindgren, K. T. Lofgren, D. Phillips, C. Atkinson, A. D. Lopez, and C. J. L. Murray (2011). Progress towards Millennium Development Goals 4 and 5 on maternal and child mortality: an updated systematic analysis. *Lancet* 378(9797), 1139–65.
- Naghavi, M., S. Makela, K. Foreman, J. O’Brien, F. Pourmalek, and R. Lozano (2010). Algorithms for enhancing public health utility of national causes-of-death data. *Population Health Metrics* 8 Issue 1.
- Plummer, M. (2003). JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), March 20-22, Vienna, Austria. ISSN 1609-395X. <http://mcmc-jags.sourceforge.net/>*.
- Plummer, M. (2011). *rjags: Bayesian graphical models using MCMC*. R package version 3-5.

- Plummer, M., N. Best, K. Cowles, and K. Vines (2006). Coda: Convergence diagnosis and output analysis for mcmc. *R News* 6(1), 7–11.
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.
- WHO, UNICEF, UNFPA, and T. W. Bank (2012). Trends in maternal mortality 1990-2010: Estimates developed by WHO, UNICEF, UNFPA and The World Bank. ISBN 978-92-4-150363-1.
- Wilmoth, J. R., N. Mizoguchi, M. Z. Oestergaard, L. Say, C. D. Mathers, S. Zureick-Brown, M. Inoue, and D. Chou (2012). A New Method for Deriving Global Estimates of Maternal Mortality. *Statistics, Politics, and Policy* 3(2).

5 Appendix

5.1 Bayesian VR adjustment model

The Bayesian VR adjustment model is specified as follows:

$$\begin{aligned}
P_{c,1978} &\sim TN_{[1/3,1]} \left(p_c, \frac{\sigma_{AR}^2}{1-\rho^2} \right), \text{ for } c = 1, \dots, C, \\
P_{c,t} &\sim TN_{[1/3,1]} (p_c + \rho(P_{c,t-1} - p_c), \sigma_{AR}^2), \text{ for } t = 1979, \dots, 2012; c = 1, \dots, C \\
p_c &\sim TN_{[1/2,1]} (w, \sigma_p^2), \text{ for } c = 1, \dots, C, \\
w &\sim U(1/2, 1), \\
\sigma_p &\sim U \left(0, \frac{1/2}{\sqrt{12}} \right), \\
\frac{1}{W_{c,k}} &\sim TN_{[0,\infty)} \left(\frac{1}{T_{c,k}} \sum_{t=t_{c,k}}^{t_{c,k}+T_{c,k}-1} P_{c,t}, \sigma_W^2 \right), \\
\sigma_W &\sim U(0, 0.05).
\end{aligned}$$

5.2 Bayesian maternal mortality estimation model

For the non-AIDS countries, we assume that the adjustment for AIDS deaths, s_i is negligible such that Eq.(1) simplified to $y_i = z_i \cdot \gamma_i \cdot q_i$ with:

$$\begin{aligned}
\log(z_i) \mid \phi_i, \gamma_i, q_i, \sigma_y &\sim N(\eta_i, \sigma_y^2), \\
\eta_i &= \phi_i - \log(\gamma_i) - \log(q_i), \\
\phi_i &= \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \beta_3 x_{3,i} + \alpha_{c[i]}^C + \alpha_{r[i]}^R, \\
\alpha_j^C &\sim N(0, \sigma_{ac}^2), \\
\alpha_k^R &\sim N(0, \sigma_{ar}^2).
\end{aligned}$$

VR adjustment parameters γ_i are modeled with the Bayesian VR adjustment model. Adjustment parameter $\gamma_i = \theta_s$ for observations obtained from information on sisterhood survival, and $\gamma_i = \theta_o$ for observation from other (non-VR and non-sisterhood) sources. The WHO distributions for these parameters are identical:

$$\begin{aligned}
\log(\theta_s) &\sim N(\log(1.1) - 0.5 \cdot 0.05^2, 0.05^2), \\
\log(\theta_o) &\sim N(\log(1.1) - 0.5 \cdot 0.05^2, 0.05^2).
\end{aligned}$$

In the Bayesian model, the WHO distributions are used as priors, but truncated at 1 to avoid high posterior probabilities of values less than 1, given that those values are deemed extremely unlikely. For parameters relating to q_i , the WHO expert distributions are used as priors. All prior distributions used in the fully Bayesian model are listed in Table 3, together with their posterior median

Table 2: **VR adjustment data set.** Sources: Appendix 1 from the WHO 2012 report (WHO et al., 2012) and additional information provided by WHO.

Country	Period/year	VR Adjustment
Australia	1994-1996	1.23
Australia	1997-1999	1.80
Australia	2000-2002	1.97
Australia	2003-2005	2.03
Austria	1980-1998	1.61
Brazil	2002-2002	1.40
Canada	1988-1992	1.60
Canada	1997-2000	1.52
China (Taiwan)	1984-1987	1.58
Denmark	1985-1994	1.94
Denmark	2002-2006	1.04
Finland	1987-1994	1.03
France	1999-1999	1.24
France	2001-2006	1.21
Georgia	2006-2006	2.00
Germany	1983-2000	1.02
Japan	2005-2005	1.35
Mexico	2008-2008	1.10
Netherlands	1983-1992	1.34
Netherlands	1993-2005	1.49
New Zealand	2006-2007	0.95
Serbia	2007-2010	1.86
Sweden	1997-2005	1.33
Switzerland	1985-1996	1.25
United Kingdom	1988-1990	1.39
United Kingdom	1991-1993	1.52
United Kingdom	1994-1996	1.64
United Kingdom	1997-1999	1.82
United Kingdom	2000-2002	1.66
United Kingdom	2003-2005	1.74
United Kingdom	2006-2008	1.60
United States of America	1991-1997	1.48
United States of America	1995-1997	1.59
United States of America	1999-2002	1.50
United States of America	2003-2005	1.10

estimate and 95% CI. The input data are observed PM (the z_i 's) for the non-AIDS countries and the observed VR adjustments for all countries (to estimate the global model parameters in the VR adjustment model).

Table 3: Posterior estimates for the Bayesian mortality estimation model parameters.

Parameter	Prior	Posterior median estimate (95% CI)
w	$U(1/2, 1)$	0.77 (0.53, 0.98)
ρ	$U(0, 1)$	0.97 (0.90, 1.00)
σ_{AR}	$U(0, 0.5)$	0.08 (0.05, 0.12)
σ_p	$U\left(0, \frac{1/2}{\sqrt{12}}\right)$	0.079 (0.005, 0.142)
σ_W	$U(0, 0.05)$	0.046 (0.034, 0.050)
θ_s	$TN_{[1,\infty)}(\log(1.1) - 0.5 \cdot 0.05^2, 0.05^2)$	1.10 (1.01, 1.20)
θ_o	$TN_{[1,\infty)}(\log(1.1) - 0.5 \cdot 0.05^2, 0.05^2)$	1.06 (1.00, 1.15)
$\hat{\beta}_0$	$N(\hat{\beta}_0, 100^2)$	2.68 (1.56, 3.74)
β_1	$N(0, 100^2)$	-0.31 (-0.44, -0.17)
β_2	$N(0, 100^2)$	0.94 (0.67, 1.24)
β_3	$N(0, 100^2)$	-1.20 (-1.82, -0.52)
q_S	Beta(mean=0.10, sd=0.04)	0.10 (0.04, 0.19)
q_o	Beta(mean=0.15, sd=0.06)	0.13 (0.05, 0.25)
σ_{ac}	$U(0, 100)$	0.41 (0.32, 0.51)
σ_{ar}	$U(0, 100)$	0.48 (0.29, 0.84)
σ_y	$U(0, 100)$	0.31 (0.27, 0.35)
Notes:		
$\hat{\beta}_0 = 2.83$, obtained from WHO multilevel regression model fit.		
$q_i = q_S$ for i in Sub-Saharan Africa, and $q_i = q_o$ otherwise.		