Global estimation of child mortality using a Bayesian B-spline bias-reduction method

Leontine Alkema and Jin Rou New
Department of Statistics and Applied Probability, National University of Singapore, Singapore
Contact: alkema@nus.edu.sg

Extended abstract for PAA 2014
September 2013

Abstract

For the great majority of developing countries without well-functioning vital registration systems, estimating levels and trends in child mortality is challenging, not only because of limited data availability but also because of issues with data quality. We developed a Bayesian penalized B-spline regression model for assessing levels and trends in the under-five mortality rate (U5MR) for all countries in the world, whereby biases in data series are estimated through the inclusion of a multilevel model. This model was recently accepted by the United Nations Inter-agency Group for Child Mortality Estimation to measure countries' progress in reducing U5MR. We present the resulting estimates of the U5MR for selected countries. Given the increasing share of neonatal and infant deaths among under-five deaths, we are extending the model to also provide estimates of neonatal and infant mortality, with uncertainty bounds, which account for data quality issues.

Introduction

For the great majority of developing countries without well-functioning vital registration systems, estimating levels and trends in the under-5 mortality rate (U5MR) is challenging, not only because of limited data availability but also because of issues with data quality. Every year, the United Nations Inter-agency Group for Child Mortality Estimation (UN IGME, including the United Nations Children's Fund, the World Health Organization, the World Bank, and the United Nations Population Division) produces and publishes estimates of child mortality comparable across countries and years for 195 countries.

In the most recent UN IGME publication in 2013 (UN IGME 2013), estimates of the U5MR were based on a revised estimation method as compared to the 2012 publication. The new method, the Bayesian B-spline bias-reduction (B3) method, was chosen to replace the previously-used method because it better accounts for data errors (including biases and sampling and non-sampling errors in the data), provided a more flexible trend fitting method and resulted in improved model validation (Alkema and New 2013).

The first objective of this paper is to summarize the IGME 2013 estimation method and present resulting country estimates of the U5MR. Given the increasing share of neonatal and infant deaths among all under-five deaths, the second objective is to extend the model to also provide estimates of neonatal and infant mortality, with uncertainty bounds, which account for data quality issues.

In this extended extract, we summarize the B3 method for the U5MR and give illustrative country examples of resulting estimates. Further details on the B3 method for estimating U5MR are given in Alkema and New (2013) and estimates of the U5MR for all countries can be found on www.childmortality.org and in UN IGME (2013).

The Bayesian B-spline bias-reduction (B3) method for U5MR

The UN IGME compiles data on U5MR annually from various sources, typically vital registration (VR) systems, surveys and censuses. These data sources either record recent births and deaths on an ongoing basis or collect retrospective information on child mortality in the form of full or summary birth histories of women. From such data, it is possible to construct observations of U5MR directly from the reported births and deaths below age 5 or indirectly via models applied to the information from summary birth histories (Hill et al. 2012). The UN IGME 2013 database, which contains the underlying data used for estimation, is publicly available on CME Info (http://www.childmortality.org).

Different data sources may yield varying estimates of U5MR for a given time period because of differences in data errors, such as random errors in sample surveys or systematic errors due to misreporting. Additionally, data may not be available for all (recent) years of interest. To obtain country-specific U5MR estimates which are comparable across time within countries, as well as across countries, trend fitting procedures are used.

For the UN IGME 2013 estimates, a penalized B-spline regression model was used for U5MR trend fitting as a flexible alternative to the previously-used Loess smoother to better capture recent trends in data. The B-spline method is illustrated in Figure 1 for Norway. B-splines are smooth curves, here placed 2.5 years apart, which add up to unity at any point in time. For any year, the estimated U5MR (on the natural logarithm scale) is the sum of the non-zero splines in that year multiplied by the corresponding spline coefficient. When estimating the spline coefficients, a flexible yet reasonably smooth U5MR curve was obtained by including a penalization of changes in spline coefficients. In the resulting spline fit, the difference between two adjacent coefficients is given by the difference between the previous two coefficients with an estimated data-driven “distortion term” added to it. For example, in Norway during the early 1980s, these distortion terms are estimated to be around zero when U5MR did not change much, but they are negative in the late 1980s when the U5MR started to decline again. The variance of the distortion terms plays the role of the smoothing parameter in the splines fit: larger variance allows for greater fluctuations in the distortion terms, thus greater variations in the trend from one period to the next. The country-specific smoothing parameter was estimated through a hierarchical model, and default settings were used for a predefined subset of countries with VR data and other small countries. U5MR projections, after the end of the most recent observation period, were obtained by projecting forward the differences between adjacent spline coefficients. These projected differences were based on the combination of country-specific projected differences in spline coefficients, based on the country's recent past, and a global distribution of observed past differences.

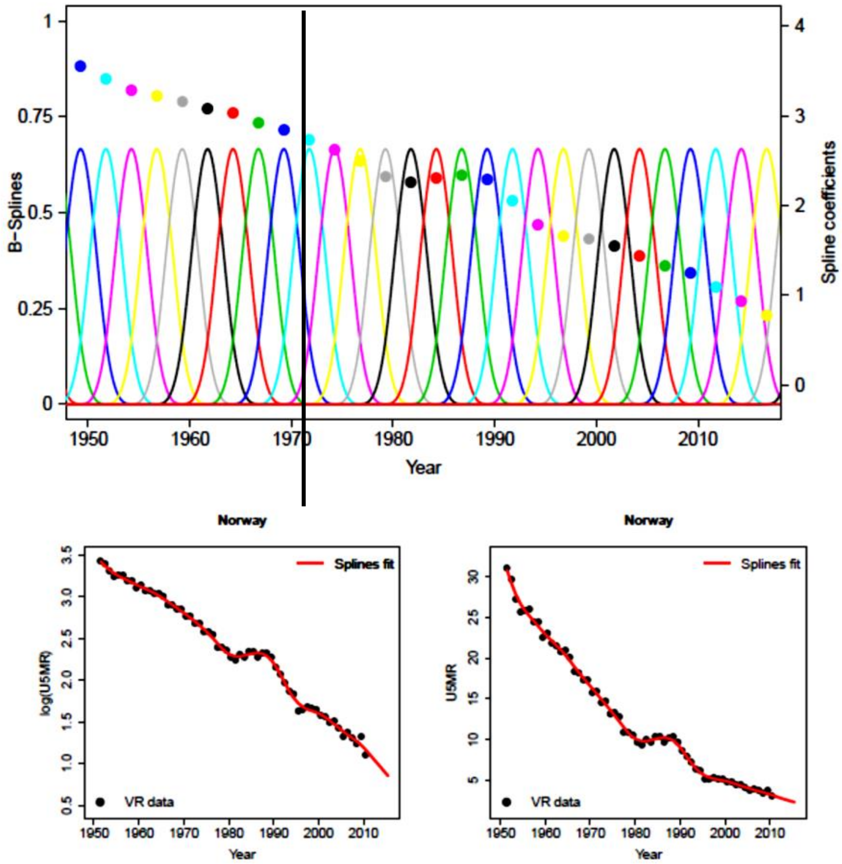

Figure 1: Illustration of the B-splines regression model for Norway. Top row: B-splines and the estimated spline coefficients. Bottom row: Observed log(U5MR) and U5MR (black dots) plotted against time, together with the spline estimates (red line).

In the UN IGME 2013 approach, various properties of the errors that provide information about the quality of the observation and the extent of expected error were taken into account to improve the accuracy of the resulting U5MR estimates. These properties include: the standard or stochastic error of the observation; its source type (e.g. whether the observation is obtained from a Demographic Health Survey or a census) and if the observation is part of a data series from a specific survey (and how far the data series is from other series with overlapping observation periods). These properties were summarized in the so-called data model. When estimating the U5MR, the data model adjusts for the errors in the observations, including the average systematic biases associated with different types of data sources, using information on data quality for different source types from all countries in the world.

For a subset of 10 countries in the regional grouping of CEE/CIS (Armenia, Azerbaijan, Georgia, Kazakhstan, Kyrgyzstan, Moldova, Tajikistan, Turkmenistan, Ukraine and Uzbekistan), VR data are generally considered to be incomplete and mostly excluded from the trend fitting in earlier rounds of UN IGME estimation. In the 2013 estimation approach, two observations in the period 1990-1995 were included per country to inform the trend during that period but not the level estimates. Moreover, for some of these countries where the U5MR extrapolations were below the incomplete VR observations in the recent period from 2005, recent VR observations were included in the model as a minimum bound after accounting for stochastic errors. Furthermore, there were cases where the U5MR extrapolations were too far above VR observations for which there was an assumed minimum level of completeness, hence an upper bound was also included.

Results: Selected country examples for U5MR

B3 estimates for selected illustrative countries are displayed in Figures 2 and 3, together with the estimates that would have been obtained using the default Loess estimation approach used for constructing the IGME 2012 estimates.

Point estimates from the B3 model and the default Loess are almost identical for the Netherlands during the entire observation period, but differ for all or a subset of observation years in the other countries. For Mexico, the trend in the Loess estimates for the late 2000s contradicts the observed trend in VR data. B3 estimates take into account the small stochastic error in the VR and follow the data points closely. For Moldova, the inclusion of the VR observations in the early 1990s with a VR bias parameter for those years results in U5MR estimates that capture the trend in the VR data. The inclusion of VR data for recent years guarantees that the point estimates and credible intervals (CIs) do not cross through the VR. In future revisions for Moldova, a further extension could be to include all incomplete VR observations as a minimum to avoid the situation in the early 1980s, when the lower bound of the CI is below the incomplete VR.

For Ghana, B3 estimates and Loess estimates are similar. Small differences are observed in the years with VR data, where the B3 estimates capture these points while the Loess does not. In more recent years, the extrapolated decline is slightly steeper for the B3 model, as indicated by the decline in the most recent observations. Differences between B3 and Loess estimates are much larger in the other countries in the figure. In Cambodia, the B3 estimates follow the trend as observed in the data series, including the stagnation of child mortality decline in the 1980s and 1990s and the more recent acceleration in the decline of child mortality. The default Loess fit does not capture these fluctuations. In the IGME 2012 method, this country would be a candidate for an expert-based adjustment of the Loess smoothing parameter to better capture the trend. In the B3 penalized spline model approach, such expert adjustments are not necessary.

In Pakistan, the B3 estimates follow the registration data. The DHS from 2006-2007 does not bias the estimates downwards (as observed in the Loess estimates) because of the inclusion of bias parameters for survey data; we estimate that the DHS direct series is biased downwards. Lastly, in Papua New Guinea, B3 estimates suggest a slightly flatter trend in U5MR than the Loess during the 1980s and 1990s based on the lack of downwards trends in all individual series during that period.

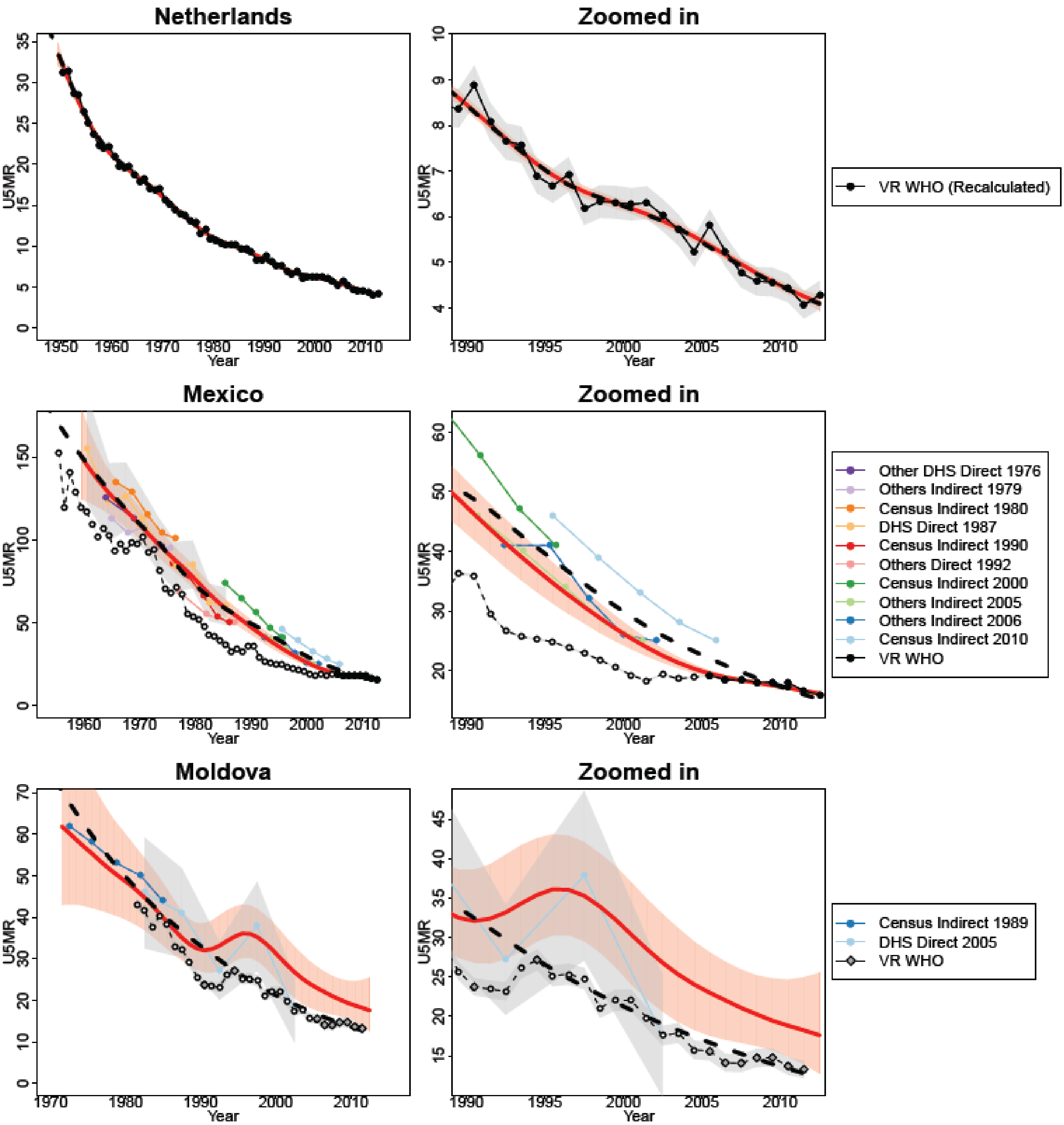

Figure 2: U5MR data series and estimates for the Netherlands, Mexico and Moldova. Connected dots represent data series from the same source, as explained in the legend. B3 estimates are illustrated by the solid red lines and 90% CIs are shown by the red shaded areas. The fittted Loess curve based on UN IGME 2012 methodology is illustrated with the black dashed line.

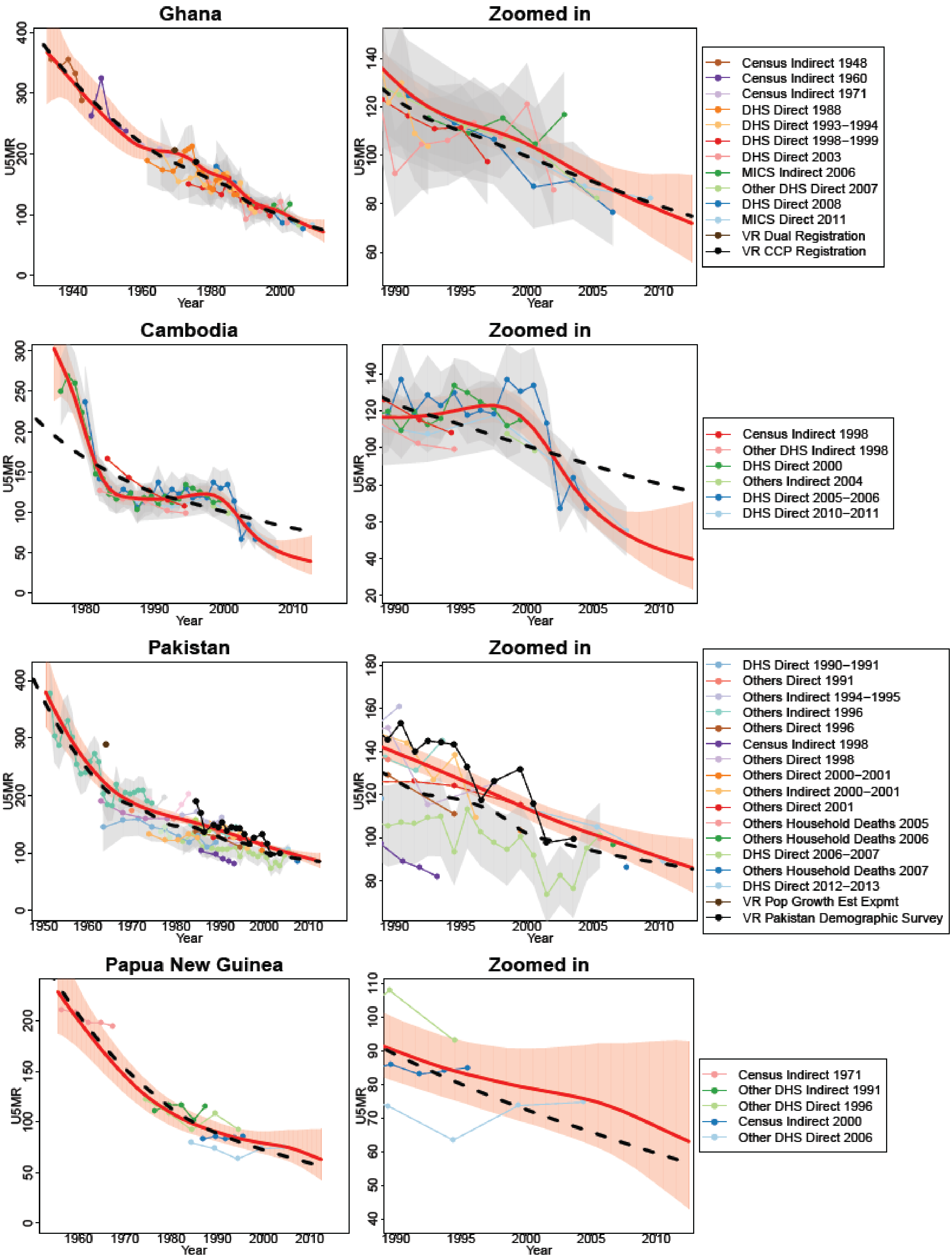

Figure 3: U5MR data series and estimates for Ghana, Cambodia, Pakistan and Papua New Guinea. Connected dots represent data series from the same source, as explained in the legend. B3 estimates are illustrated by the solid red lines and 90% CIs are shown by the red shaded areas. The fittted Loess curve based on UN IGME 2012 methodology is illustrated with the black dashed line.

Acknowledgements

The authors are grateful to all members of the United Nations Inter-agency Group for Child Mortality Estimation and its Technical Advisory Group for very helpful discussions, suggestions and comments.

This research was funded by a research grant at the National University of Singapore and the United Nations Children's Fund. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the United Nations Children's Fund.

References

1. Alkema, L. and J. R. New (2013). Global estimation of child mortality using a Bayesian B-spline bias-reduction method. arXiv:1309.1602 [stat.AP]. Available at http://arxiv.org/abs/1309.1602.
2. The United Nations Inter-agency Group for Child Mortality Estimation (2013) Levels & Trends in Child Mortality: Report 2013. UNICEF.
3. Hill, K., You, D., Inoue, M., & Oestergaard, M. Z. (2012). Child mortality estimation: accelerated progress in reducing global child mortality, 1990-2010. *PLoS Medicine*, *9*(8), e1001303. doi:10.1371/journal.pmed.1001303