Sex Differences in U5MR: Estimation and identification of countries with outlying levels or trends *

Leontine Alkema
National University of Singapore

Cheryl C. Sawyer
United Nations

Extended abstract for PAA 2013,
September 21, 2012

Abstract

Under natural circumstances, the ratio of the male to female under-five mortality rate is greater than one. However, deprivation of girls access to health care or proper nutrition could lead to distorted ratios of under-five mortality. Monitoring of mortality by sex is challenging because of issues with data availability and quality. Moreover, the sex ratio is expected to vary with under-five mortality, which makes it challenging to define "expected levels". We present a Bayesian hierarchical model to estimate the sex ratio of under-five mortality for all countries. In addition, we estimate the relative difference between national sex ratios and expected sex ratios based on the global relation between under-five mortality and the sex ratio. All estimates include an uncertainty assessment to enable assessments of whether differences between countries or within countries over time are significant or highly uncertain. Preliminary results suggest that girls may be disadvantaged in several countries.

*Leontine Alkema, Department of Statistics and Applied Probability and Saw Swee Hock School of Public Health, National University of Singapore, Singapore 117546; Email: alkema@nus.edu.sg. Cheryl C. Sawyer, Mortality Section, United Nations Population Division, New York, NY 10017; Email: sawyerc@un.org.

1 Introduction

After birth, girls have biological advantages over boys with respect to survival up to age five. Under natural circumstances, the ratio of the male to female under-five mortality rate (U5MR, the probability of dying before age 5), is greater than one. However, deprivation of girls access to health care or proper nutrition could lead to distorted ratios of under-five mortality. A recent assessment found that a number of countries in southern and western Asia, as well a number of developing regions in Africa, experience higher mortality for girls compared to boys from one to four years old (Sawyer 2012).

Obviously, gender discrimination is in conflict with human right laws. Additionally, increased mortality for girls as compared to boys (especially in combination with pre-natal gender discrimination) may lead to "missing girls"; an under-representation of women in the population. A spectrum of societal consequences may arise from a sex imbalance. At the least, a "marriage squeeze", whereby a surplus of single men (potentially the most vulnerable men from low socio-economic standing) will not be able to find a wife, is one of the unavoidable consequences (UNFPA 2012). The full spectrum of societal consequences remains to be seen and is likely to depend on the extent of the imbalance.

Monitoring trends in sex differentials in under-five mortality is of key importance to safeguard girls' rights and avoid or modulate the societal consequences of gender imbalance. Unfortunately, monitoring is challenging because of issues with data availability and quality. This is illustrated in Figure 1(a), which shows observations on the ratio of male to female U5MR in Jordan (denoted by SR5). The data do not provide a coherent story on levels and trends.

The monitoring of trends in sex differentials is also complicated by the lack of a specific "expected SR5 level". The SR5 tends to vary with U5MR; as under-five mortality decreases, the share of causes of death that more typically affect boys tends to increase and girls' advantage tend to increase as well. The empirical relation between SR5 and U5MR is illustrated in Figure 1(b). To pinpoint to countries where SR5 levels or trends divert from "what is expected", the expected levels first need to be defined, for example based on the regional or global relation between SR5 and U5MR.

Sawyer (2012) presented statistical techniques to construct SR5 estimates for all countries. While the study provided a big step forward in international monitoring of SR5s, there were some drawbacks to the estimation method related to the choice of the estimation method and the extrapolation beyond the most recent data point. Additionally, there was no uncertainty assessment of the estimates and the focus of the article was on regional and global trends and not on country comparisons to a reference level. Lozano and colleagues (2011) constructed estimates of sex-specific under-five mortality with an uncertainty assessment but the approach did not provide insights into countries with outlying trends or levels either; results were presented in plots only and levels or trends were not discussed. Moreover, a validation of the estimation method was lacking and potential issues with data quality were not accounted for.

Our objective is to develop a robust modeling approach for a comprehensive and replicable analysis of SR5 levels and trends for all countries that includes an uncertainty assessment, takes into account issues with data quality, produces calibrated uncertainty intervals and allows for pinpointing to countries where levels divert from the expected levels. In this

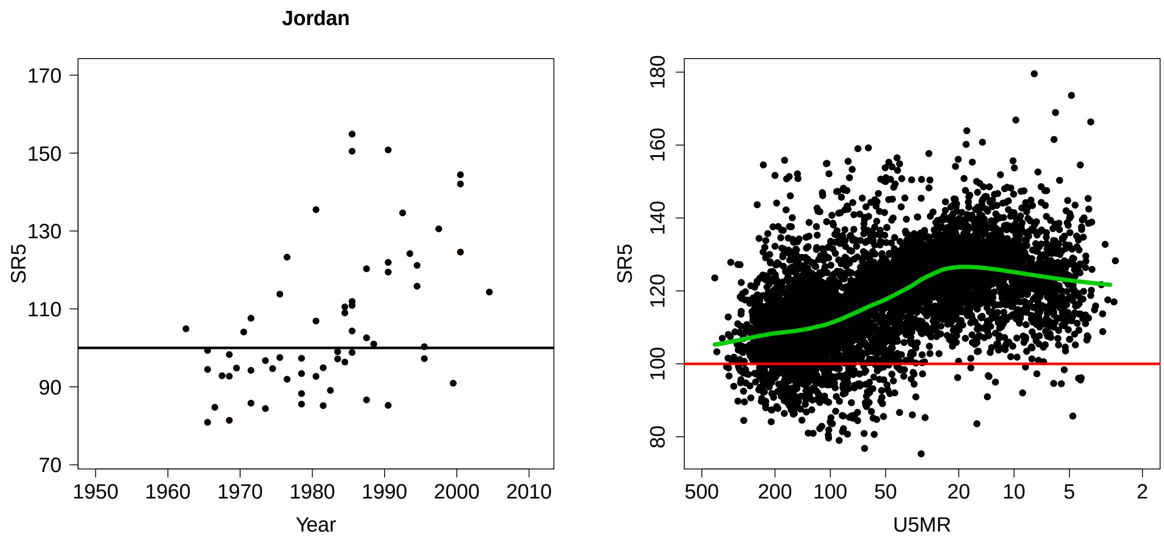

Figure 1: (a) Illustration of SR5 observations for Jordan. (b) Illustration of the empirical relation between SR5 and U5MR. SR5 observations from all countries are plotted in black versus U5MR (on log-scale), a loess smoother is added in green.

extended abstract, we present preliminary results from a Bayesian hierarchical modeling approach. We first discuss available data in Section 2, followed by the presentation of our model. We present preliminary results in Section 4 and discuss the next steps in the Discussion.

2 Data

The data used in this study are observations of under-five mortality by sex from vital registration, surveys, or census. Vital registration systems with complete coverage are considered the gold standard for mortality estimation. Most developing countries have less than complete coverage in their vital registration systems. For this study, however, any available vital registration data were used to compute sex ratios of under-five mortality for analysis, on the assumption that under-registration does not vary significantly by sex. Data were also taken from sample registration systems (India, Bangladesh) or surveillance systems (China). In a few instances estimates come from household deaths by age and sex reported in a census or survey.

Estimates were also taken from full birth histories collected in sample surveys. For Demographic and Health Surveys, World Fertility Surveys and selected surveys from the Pan-Arab Programme on Family Health (PAPFAM), sex-specific estimates from full birth histories for five-year periods preceding the survey were calculated according to the method in Pedersen and Liu (2012). For surveys and censuses in which only summary birth histories were collected, sex-specific indirect estimates of U5MR were calculated using the Brass method either from microdata or from published tabulations of children ever born and children living. In cases where microdata or tabulations were not available, direct or indirect estimates were taken from published survey or census reports.

3 Method

Summary and notation The SR5, denoted by $S_{c,t}$ for country c, year t, is expected to vary with the under-five mortality rate, denoted by $q_{c,t}$ for the corresponding country-year. In our statistical model, we used the global relation between SR5 and the U5MR to find the expected SR5, denoted by $W_{c,t}$, for each country-year (given the U5MR level). The expected SR5 is then multiplied by a second component $P_{c,t}$ that represents the relative advantage or disadvantage of girls to boys compared to other countries at similar levels of U5MR, as indicated by the data in the country. This component is modeled for each country over time by a linear trend combined with a time series model. The coefficients of the linear function are estimated using a Bayesian hierarchical model.

Model details $S_{c,t}$ is modeled as follows:

$$\begin{aligned} S_{c,t} &= W_{c,t} \cdot P_{c,t}, \\ W_{c,t} &= \exp(w_0 + w_1 \tilde{q}_{c,t} + w_2 \tilde{q}^2_{c,t}), \\ P_{c,t} &= \exp(\alpha_c + \beta_c \tilde{t} + \varepsilon_{c,t}), \\ \varepsilon_{c,t} &\sim N(\rho \cdot \varepsilon_{c,t-1}, \sigma^2), \end{aligned}$$

where $\tilde{q}_{c,t}$ is log-transformed (and centered) U5MR $q_{c,t}$ and $\tilde{t}$ the centered version of year t. The coefficients α_c and β_c that model the linear trend in $P_{c,t}$ are estimated with a Bayesian hierarchical model (Lindley and Smith 1972; Gelman, Carlin, Stern, and Rubin 2004):

$$\begin{aligned} \alpha_c &\sim t(\mu = \mu_\alpha, \sigma = \sigma_\alpha, \nu = \nu_\alpha), \\ \beta_c &\sim t(\mu = \mu_\beta, \sigma = \sigma_\beta, \nu = \nu_\beta), \end{aligned}$$

to aid the estimation of the coefficients in countries with limited data. t-distributions are used to allow for countries with outlying trends.

The data model is given by

$$\log(y_i) \sim t(\mu = \log(S_{c[i]),t[i]}, \sigma = \sigma_y/\sqrt{k_i}, \nu = \nu_y),$$

where y_i is the i-th observed ratio of male to female U5MR in country $c[i]$, year $t[i]$, and k_i is the weight assigned to observation i.

Spread out prior distributions are assigned to all hyperparameters. Samples from the posterior distributions were obtained using Markov chain Monte Carlo sampling, implemented in open source software packages `R 2.15` (R Development Core Team 2011) and `JAGS` (Just Another Gibbs Sampler, Plummer (2003)), using `R`-packages **R2jags** (version 0.03-07) and **rjags** (Su and Yajima 2011; Plummer 2011).

4 Preliminary results

Preliminary results illustrate the kind of insights on levels and trends that can be obtained from this type of analysis. Unsurprisingly, Asian countries, notably China and India, top the list of countries where there is a significant "female disadvantage" in the year 2000 (Figure 2), referring to a SR5 that is lower than expected, based on the world level relation between SR5 and the U5MR. The uncertainty intervals allow for pinpointing to additional countries in which females appear to be disadvantaged relative to other countries at the same level of mortality. The list includes several other Asian countries, but also countries from many other world regions

Figure 3 illustrates estimates in Bangladesh, Iran and Jordan over time. The estimates from the United Nations (2011) publication on sex differences are added for comparison and referred to as UNPD estimates. Even though the estimates of the SR5 are above one for Bangladesh and similar to the UNPD estimates, our analysis shows that relative to other countries at similar U5MR levels, the ratio in Bangladesh is significantly lower than expected. Iran is another example of a country where we find lower-than-expected SR5s in 2000. Additionally, while the UNPD estimates for Iran are given by the fit of a straight line, our method provides a closer fit of the trend line to the data, thus providing more insights into trends during the observation period. Lastly, for Jordan, our fit is similar to the linear UNPD fit during the observation period, and tempers the extrapolation compared to the linear fit for more recent SR5. The SR5 is comparable to other countries at similar U5MR levels since the late 1990s.

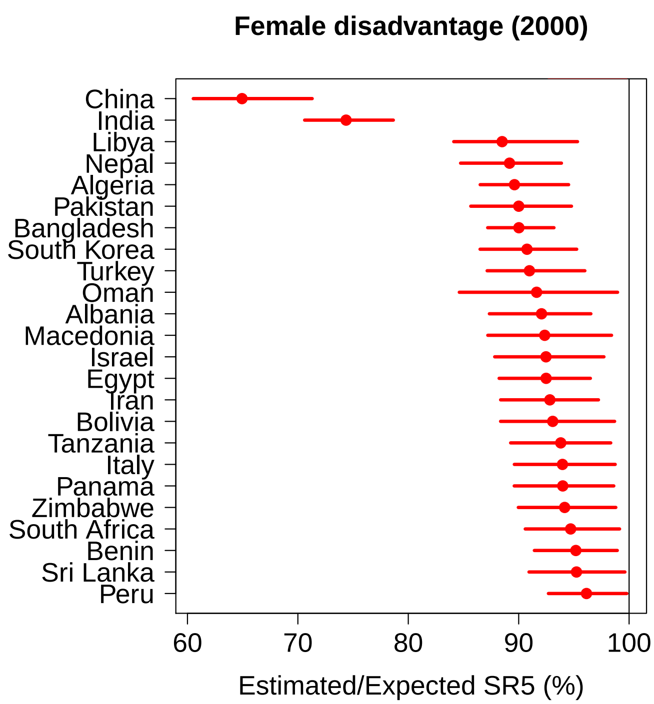

Figure 2: **Preliminary estimates of $P_{c,t}$, the ratio of the estimated SR5 in a country and its expected level based on its U5MR ($\cdot$100%) for the year 2000.** Ratios smaller than 100% suggest a "female disadvantage", referring to a SR5 that is lower than expected, based on the global relation between SR5 and the U5MR.

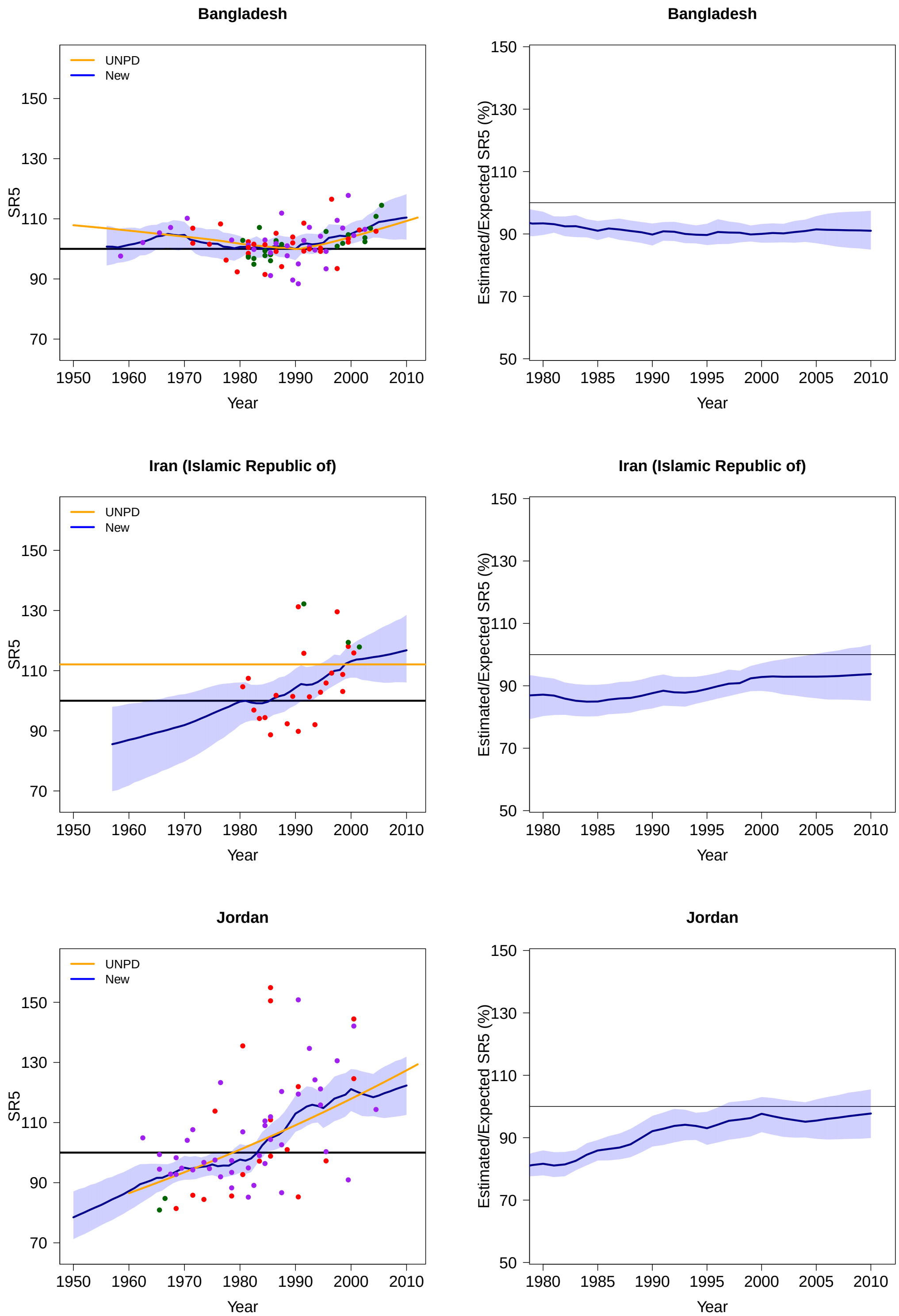

Figure 3: **Examples of estimates over time for Bangladesh, Iran and Jordan.** Left column: SR5 ($\cdot 100\%$). Right column: $P_{c,t}$, the ratio of estimated SR5 over the SR5 that is expected based on the U5MR ($\cdot 100\%$). Estimates from the United Nations (2011) publication for SR5 are displayed in orange, estimates with the Bayesian hierarchical time series model in blue. The shaded area represents the 95% uncertainty range. SR5 observations are represented by colored dots.

5 Discussion and future work

We discussed a modeling approach to estimate the difference the under-five mortality rate between girls and boys. We used the world level relation between SR5 and the U5MR to find the expected SR5 for each country-year. A second "relative difference" component $P_{c,t}$ was included in the model to represent the relative advantage or disadvantage of girls to boys compared to other countries at similar levels of U5MR, as indicated by the data in the country. This relative difference was modeled for each country over time by a linear trend combined with a time series model. The coefficients of the linear function were estimated using a Bayesian hierarchical model.

The preliminary results provide interesting insights into levels and trends in the SR5, and point to directions for model improvement. In the next steps in this project, we will include an improved data model to account for differences in data quality and explore different estimation methods to improve the extrapolation of the SR5. For example for Iran in Figure 3, the back-projections before 1980 seem to be driven only by the mean trend during the observation period and might not be realistic at all. We plan to carry out validation exercises by leaving out data series and checking how well the model predicts these left out observations and whether additional data change the estimates significantly. Lastly, we plan to investigate the relation between SR5 and U5MR in more detail and to analyze regional differences.

Acknowledgements

This work was supported by a research grant from the National University of Singapore. The views expressed in this article are those of the authors and do not necessarily reflect the views of the United Nations. Its contents have not been formally edited and cleared by the United Nations. The designations employed and the presentation of material in this article do not imply the expression of any opinion whatsoever on the part of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. The authors are grateful to Stella Ting Hai Yin for research assistance and the members of (the technical advisory group of) the Inter-Agency Group for Child Mortality Estimation for helpful comments and discussions.

References

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2004). *Bayesian Data Analysis* (2nd ed.). Boca Raton, Fl.: Chapman & Hall/CRC.

Lindley, D. V. and A. F. M. Smith (1972). Bayes estimates for the linear model. *Journal of the Royal Statistical Society, Series B 34*, 1–41.

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.

Pedersen, J. and J. Liu (2012). Child Mortality Estimation: Appropriate Time Periods for Child Mortality Estimates from Full Birth Histories. *PLoS Medicine 9*(8), e1001289.

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.

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.

Sawyer, C. C. (2012). Child Mortality Estimation: Estimating Sex Differences in Childhood Mortality since the 1970s. *PLoS Medicine 9*(8), e1001287.

Su, Y. S. and M. Yajima (2011). *R2jags: A Package for Running jags from R*. R package version 0.02-17.

UNFPA (2012). Sex Imbalances at birth: Current Trends, Consequences and Policy Implications. Technical report.

United Nations, Department of Economic and Social Affairs, Population Division (2011). *Sex Differentials in Childhood Mortality*. United Nations publication, New York.