**The Global Burden of Disease fertility forecasts:
Summary of the approach used and associated statistical concerns**

Leontine Alkema¹
Department of Biostatistics and Epidemiology
University of Massachusetts Amherst

August 17, 2020

Abstract

BACKGROUND

The Global Burden of Disease (GBD) project's forecasts up to 2100 suggest fertility drops will be even greater in sub-Saharan Africa than the UN Population Division (UNPD) has predicted.

OBJECTIVE

This reflection summarizes the main assumptions used in the GBD fertility forecasts. I assess the methods used, focusing on high fertility countries and the use of met need for contraceptives as a predictor.

RESULTS

Based on GBD's forecasting method, I draw two conclusions. Firstly, GBD fertility forecasts are based on unvalidated assumptions about increasing met need for contraception and may overestimate decreases in fertility in countries with low levels of modern contraceptive use. Secondly, the GBD forecast model for fertility is *not* a causal model for predicting changes. Claims GBD researchers make about the effect of changing access to family planning on fertility are not informative for guiding policy. Based on the GBD validation exercise, I conclude that the GBD study did not check the performance of the method for predicting left-out fertility data. Also the approach used to compare the predictive performance of UNPD and GBD forecasting methods may give the GBD method an inherent advantage.

CONCLUSIONS

Communication regarding the GBD method and its findings must avoid causal language and acknowledge the method's limitations. Future research should examine the performance of the method, especially for countries with low modern contraceptive use.

CONTRIBUTION

This paper summarizes the GBD fertility forecasting method and indicates three areas of concern about it and its use.

¹ Contact: lalkema@umass.edu. This work was funded by the Bill and Melinda Gates Foundation. I thank Patrick Gerland for helpful comments.

Introduction

The Global Burden of Disease (GBD) project recently published population forecasts up to 2100 (Vollset et al., 2020). These forecasts suggest faster declines in sub-Saharan fertility than found in the forecasts by the UN Population Division (UNPD) in the World Population Prospects (United Nations, Department of Economic and Social Affairs, Population Division, 2019). An important difference between these forecasts is that the UNPD did not use covariates for predicting fertility changes, while GBD used forecasts of educational attainment and met need for contraceptives.

In this paper, I summarize the main assumptions used in the GBD fertility forecasts and identify three concerns associated with the forecasting approach and its validation. I focus on the approach used for forecasting completed cohort fertility (CCF) and one of its drivers, met need for contraceptives. My expertise to examine the GBD method stems from prior experience with fertility and met need forecasting. I have contributed to model development for fertility projections (Alkema et al., 2011; Raftery et al., 2014) and estimation and projections of family planning indicators, including met need for contraceptives (Alkema et al., 2013, Cahill et al., 2018). This experience means that I am familiar with issues that arise when projecting met need and fertility. I also was fortunate to work with and learn from substantive experts in these content areas. My concerns about the validity and use of the GBD fertility forecast are based on dispassionate review and commitment to scientific integrity.

A summary of the GBD's completed cohort fertility forecast method

The GBD forecasts of fertility start in the year 2018, based on GBD fertility estimates up to the year 2017 (Murray et al., 2018). The fertility forecast is centered around forecasts of CCF, referring to cumulative fertility from age 15 to 50. CCF forecasts start with birth cohorts born in 2003.

Fertility for 2100 is obtained from the CCF for all cohorts who are in their reproductive ages during that year. The reference forecast of future CCFs is obtained in five steps:

1. Fit a linear regression model to obtain the global relation between CCF and cohort-based summaries of education and met need for contraceptives, using past estimates of these three indicators. The regression model is given by
$$CCF = \beta_0 + \beta_1 * \text{met need} + \text{non-linear function of education.}$$
Past estimates refer to fertility estimates of birth cohorts born between 1969 and 2002 from Murray et al. (2018).
2. For each country and birth cohort considered, calculate the difference between (on a logit-scale): (i) the GBD-based CCF estimate and (ii) estimate based on regression function, using country values of education and met need and estimates of regression coefficients. I refer to this difference as the residual, i.e. residual = logit-transformed estimate – logit-transformed “regression-based prediction.”
3. Forecast education for birth cohorts up to 2085 (not discussed here).
4. Forecast met need for birth cohorts up to 2085 for each country through the following steps:
 - a. Calculate the estimated rate of change (RoC) in logit-transformed age-specific met need in the past.
 - b. Use that country-specific rate of change to project met need to 2100. Two exceptions are applied: past RoCs that are extreme are replaced with less extreme limits, and a maximum to projected met need was set based on maxima observed so far.

There are questions about how the estimated RoC in logit-transformed age-specific met need in the past was calculated. The text does not say what past period was used, but only that all past estimates are equally weighted. Hence, forecasts do not account for possible stalls or accelerations in a country's recent past.

5. Calculate the point forecasts for 2100 using point forecasts for birth cohorts up to 2085, as follows:
 - a. Point forecast for CCF = prediction based on regression function + residual.
 - b. The prediction is obtained by plugging the 2100 forecasts into the regression equation.
 - c. The residual is given by the residual for the 2002 birth cohort from step 2.

Based on this approach, the difference between a country's cohort-age-specific fertility and its fertility expected based on its met need and education covariates remains constant until the end of the century.

In order to construct the uncertainty intervals, the GBD makes the assumption that the residuals from step 2 above follow a random walk, where $\text{residual}(c,t+1) = \text{residual}(c,t) + \text{random error term}$. The variance of the random error terms is estimated using the residuals from step 2. Uncertainty intervals are obtained based on trajectories of the CCF. The trajectories incorporate two or three sources of uncertainty:

1. The uncertainty in the CCF estimate for the 2002 birth cohort.
2. The uncertainty due to deviations away from the regression function, as captured by the random walk.
3. Less clearly, the uncertainty in education forecasts. The paper states that uncertainty in all covariates (independent drivers) is included but there is no uncertainty in the GBD met need forecast. I did not check if uncertainty in education forecasts is assessed and included.

In addition to the reference forecast, forecast scenarios are constructed. These scenarios are obtained by changing the future trajectories of met need for contraceptives and education.

Three concerns regarding the GBD fertility forecasts

My three main concerns are as follows. First, the fertility reference forecasts for countries with low levels of modern use may be biased downwards. Second, the GBD forecast model for fertility is not a causal model for predicting changes. Hence claims that the GBD model can be used to explore the impact of changes in family planning access on fertility through forecast scenarios are invalid. Third, the GBD forecasting method has been validated to check its performance for predicting left-out GBD fertility estimates only; it has not been validated to check how well it predicts left-out fertility data, and the approach used to check the predictive performance of UNPD and GBD forecasting methods may give the GBD method an inherent advantage over the UNPD method. I discuss each concern in greater detail below.

Concern 1: Upward bias of met need forecasts leading to downward bias of fertility forecasts.

The GBD based its forecasts of met need for contraceptives in countries with low levels of modern contraceptive use on the estimated rate of change in logit-transformed age-specific met need in the past, as described in step 4 above.

I am concerned that the met need forecasts are biased upwards for countries at low levels of modern contraceptive use. This concern reflects my analysis of differences across countries in observed rates of change using the population of women aged 15-49 who are married or in a union, based on estimates obtained from the UNPD (United Nations, Department of Economic and Social Affairs, Population Division, 2020). Figure 1 illustrates that for the populations considered and comparing across countries, the RoC tends to decrease as modern use increases. If indeed the RoC decreases with the level of modern use, using a constant ROC in forecasts regardless of the level of modern contraceptive use results in overprediction of the rate of change and thus, overprediction of met need.

Granted, my analyses are not based on the exact indicators used in the GBD study. In particular, my analysis concerns across-country comparisons and a different summary measure. Hence, the assumption of a constant rate of change over the duration of a contraceptive transition may be invalid in my analysis but valid in the GBD study. Further analysis of the specific data used is therefore needed. GBD could have addressed this concern through validation of met need forecasts, but the paper does not include them and the Foreman et al. (2018) paper to which it refers also does not provide them.

If my concern proves founded and met need forecasts are biased upwards in countries with low contraceptive use, GBD has overestimated the future decline in fertility that is associated with met need in short and long term forecasts. In other words, its forecast represents an underprediction of future fertility. Further analysis is needed to assess the performance of GBD fertility forecasts for countries with low contraceptive use and high fertility. While GBD carried out model validation to check the accuracy of the fertility forecasting method, I have some concerns with the approach, as described later in this paper. Moreover, the GBD study did not present results for high fertility countries separately and validation summaries did not include information regarding bias.

If fertility forecasts are biased downwards, population forecasts are biased downwards as well, due to underprediction of the number of births associated with cohorts. Such a downwards bias in population forecasts would have a cumulative effect in future generations. Each cohort that is larger than the GBD suggests will in turn have a greater number of offspring than the GBD predicts. Over time the error in the forecast will grow.

Figure 1: RoC of logit-transformed met need versus proportion of women who use modern methods of contraception. Each dot represents one country. The indicators refer to estimates for the population of women aged 15-49 who are married or in a union. RoC was calculated as the rate of change on logit-scale between 1990 and 2017. Modern contraceptive use refers to the proportion of women who use a modern method of contraception in 1990. The blue line refers to a locally smoothed regression line (loess smoother).

Concern 2: Using the GBD fertility forecasting model as a causal model for predicting changes is incorrect

GBD calls its model “causal” and states that it can therefore “provide direct guidance to policy debates on the impact of faster or slower scale-up of educational attainment or access to reproductive health services” (p. 2). While I applaud the focus on access to reproductive health services in policy discussions, the GBD’s fertility model is not a causal model and it cannot provide guidance on the impact of faster or slower scale-up of access. The relation between education, family planning, and fertility is complex and I cannot do justice to the extensive body of work on the topic here. Instead, I will point out that the GBD model does not even capture a simple causal relation between main family planning indicators and fertility reductions.

In its most basic form, fertility reductions are related to family planning through the share of women who have a demand for contraceptives, that is, who *want* to limit or postpone child bearing, and among those who use a contraceptive method to avoid pregnancy. The GBD model includes only the relation between fertility and met need, which is the proportion of women with a demand for family planning who are using a modern method of contraception. By design, the GBD model does not distinguish

fertility reductions that are due to increases in demand for family planning from those that are due to improved access among those in need. Hence, the GBD model cannot be used to explore the impact of changes in family planning access on fertility because the estimated effect of such changes on fertility are confounded with the effect of changes in demand.

Concern 3: Limitations of GBD's model validation exercise

GBD carried out model validation to check predictive performance of its fertility forecast model as well as UNPD's fertility forecast model. In the validation exercise, GBD researchers fitted the fertility forecasting model to GBD fertility estimates up to the year 2007, and used the fitted model to predict the omitted estimates for the years 2008 to 2017. Fertility estimates up to the year 2017 were constructed using the GBD fertility estimation model (Murray et al., 2018). The GBD fertility estimation model was fitted to fertility data, as reported in surveys and registration systems, and uses education as a covariate to inform the estimates. Examples of fertility data and GBD estimates are shown for selected countries in Figure 2. The figure illustrates that in settings where recent fertility data are uncertain or not available at all, the GBD estimation model serves as a fertility forecasting model to predict changes in fertility up to 2017. While the GBD validation exercise assessed how well the forecasting methods predicted the left-out model-based estimates, it did not assess its performance for predicting left-out data.

I argue that prediction errors for left-out estimates in country-periods with limited data assess the internal consistency between the forecasting method and the GBD estimation method, as opposed to the ability of the fertility forecasting model to predict future changes in fertility. For country-periods with limited data, changes in education predict model-based estimates of annual changes in fertility. Hence, prediction errors for left-out model-based estimates for such country-years quantify how well the forecasting method reproduces the education-based forecasts as opposed to actual fertility outcomes. If prediction errors for country-periods with limited data assess internal consistency instead of predictive performance, summary measures based on all country-years are subject to the same limitation.

I also argue that using GBD estimates for checking predictive performance of both the UNPD and GBD forecasting methods may give the GBD method an inherent advantage over the UNPD method. The GBD fertility estimation method and the fertility forecasting method both use education as a covariate but the UNPD method does not. The use of education as a covariate in the two GBD methods is likely to produce greater consistency between the forecasts of those methods than between the forecasts of the UNPD forecasting and GBD estimation method. If the GBD forecasting method produces a forecast that is more consistent with the GBD estimates than the UNPD method, and if, as I argue above, prediction errors assess the internal consistency between the forecasting method and the GBD estimation method, then by design of the validation exercise, prediction errors will be smaller for the GBD method as compared to the UNPD method due to the use of education in GBD estimates and forecasts.

Conclusion

This paper focused on assessing the GBD fertility forecasting method, focusing on high fertility countries and the use of met need as a predictor. Based on the GBD forecasting method used, I have argued that its fertility forecasts for countries with low levels of modern use may be biased downwards, due to upwards bias in its met need forecasts. I am also concerned about the short-term performance of GBD

met need forecasts given that recent stalls or accelerations are not accounted for. I call on GBD to validate its met need forecasts and reconsider its forecasting methodology as needed.

I have also sounded the alarm about using the GBD forecast model for fertility as a causal model. By design, the GBD model does not distinguish fertility reductions that are due to increases in demand for family planning from those that are due to improved access. Hence, the GBD model cannot be used to explore the sole impact of changes in family planning access on fertility because the estimated effect of such changes on fertility is confounded with the effect of changes in demand. Communication regarding the model and its findings needs to avoid incorrect causal language.

Future research should examine the performance of the GBD fertility forecast method, especially for countries with low modern contraceptive use, and compare predictive performance between UNPD and GBD methods using standard approaches. While the GBD validation exercise assessed how well the forecasting methods predicted the left-out model-based estimates, it did not assess its performance for predicting left-out fertility data. Given that a subset of model-based estimates are forecasts based on the GBD fertility estimation model, prediction errors for such forecasts assess internal consistency between forecasting methods, as opposed to the model's ability to predict changes in fertility outcomes. In addition, the GBD study did not present information regarding bias or calibration of uncertainty intervals.

References

Alkema, Leontine, Adrian E. Raftery, Patrick Gerland, Samuel J. Clark, François Pelletier, Thomas Buettner, and Gerhard K. Heilig (2011). "Probabilistic Projections of the Total Fertility Rate for All Countries." *Demography* 48, no. 3: 815–39. <https://doi.org/10.1007/s13524-011-0040-5>.

Alkema, Leontine, Vladimira Kantorova, Clare Menozzi, and Ann Biddlecom (2013). "National, Regional, and Global Rates and Trends in Contraceptive Prevalence and Unmet Need for Family Planning between 1990 and 2015: A Systematic and Comprehensive Analysis." *The Lancet* 381, no. 9878: 1642–52. [https://doi.org/10.1016/S0140-6736\(12\)62204-1](https://doi.org/10.1016/S0140-6736(12)62204-1).

Cahill, Niamh, Emily Sonneveldt, John Stover, Michelle Weinberger, Jessica Williamson, Chuchu Wei, Win Brown, and Leontine Alkema (2018). "Modern Contraceptive Use, Unmet Need, and Demand Satisfied among Women of Reproductive Age Who Are Married or in a Union in the Focus Countries of the Family Planning 2020 Initiative: A Systematic Analysis Using the Family Planning Estimation Tool." *The Lancet* 391, no. 10123. [https://doi.org/10.1016/S0140-6736\(17\)33104-5](https://doi.org/10.1016/S0140-6736(17)33104-5).

Foreman, Kyle J., Neal Marquez, Andrew Dolgert, Kai Fukutaki, Nancy Fullman, Madeline McGaughey, Martin A. Pletcher, et al. (2018). "Forecasting Life Expectancy, Years of Life Lost, and All-Cause and Cause-Specific Mortality for 250 Causes of Death: Reference and Alternative Scenarios for 2016–40 for 195 Countries and Territories." *The Lancet* 392, no. 10159: 2052–90. [https://doi.org/10.1016/S0140-6736\(18\)31694-5](https://doi.org/10.1016/S0140-6736(18)31694-5).

Murray, Christopher J L, Charlton S K H Callender, Xie Rachel Kulikoff, Vinay Srinivasan, Degu Abate, Kalkidan Hassen Abate, Solomon M Abay, et al. (2018). "Population and Fertility by Age and Sex for 195 Countries and Territories, 1950–2017: A Systematic Analysis for the Global Burden of Disease Study 2017." *The Lancet* 392, no. 10159: 1995–2051. [https://doi.org/10.1016/S0140-6736\(18\)32278-5](https://doi.org/10.1016/S0140-6736(18)32278-5).

Raftery, Adrian E., Leontine Alkema, and Patrick Gerland (2014). "Bayesian Population Projections for the United Nations." *Statistical Science* 29, no. 1: 58–68.

United Nations, Department of Economic and Social Affairs, Population Division (2020). Estimates and Projections of Family Planning Indicators 2020. New York: United Nations.

United Nations, Department of Economic and Social Affairs, Population Division (2019). World population prospects 2019: volume I: comprehensive tables. New York: United Nations.

Vollset, Stein Emil, Emily Goren, Chun-Wei Yuan, Jackie Cao, Amanda E. Smith, Thomas Hsiao, Catherine Bisignano, et al. (2020). "Fertility, Mortality, Migration, and Population Scenarios for 195 Countries and Territories from 2017 to 2100: A Forecasting Analysis for the Global Burden of Disease Study." *The Lancet*. [https://doi.org/10.1016/S0140-6736\(20\)30677-2](https://doi.org/10.1016/S0140-6736(20)30677-2).

Figure 2: GBD 2017 study data (symbols) and estimates (red line) of the total fertility rate (TFR) from 1970 to 2017 for Burkina Faso, Mozambique, Tanzania, and Rwanda. Figures taken from Murray et al. (2018), licensed under [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/).

Age-specific fertility fitted trend lines, 1950–2017: Rwanda, TFR
Data points enclosed with a diamond are reference sources

Age-specific fertility fitted trend lines, 1950–2017: Tanzania, TFR
Data points enclosed with a diamond are reference sources