Estimating Pregnancies, Abortions and Pregnancy Intentions using a Bayesian Accounting Model

Jonathan Bearak
Guttmacher Institute

Leontine Alkema
University of Massachusetts, Amherst

Estimates of pregnancies, abortions and pregnancy intentions can help assess how effectively women and couples are able to fulfil their childbearing aspirations. Abortion incidence estimates are also a necessary foundation for research on the safety of abortions performed and the consequences of unsafe abortion. Estimates can additionally inform policy and programmes, such as by highlighting the importance of access to safe, legal abortion care, a critical reproductive health service.

However, estimating the distribution of pregnancies by intention and outcome is challenging. Data requirements include information on the proportion of births that are intended and on the incidence of abortion. Countries may lack data on one or both of these variables, for some or all time periods in question. Additionally, the availability and reliability of data may vary non-randomly.

Rigorous methodologies are needed for the estimation of these imperfectly measured outcomes. Therefore, we propose a theoretically grounded Bayesian statistical model that jointly estimates unintended pregnancy and abortion incidence using all available data on the proximate determinates of fertility. We additionally developed a data classification process applicable to all available data. Data sources and the data classification process are described in detail in the study protocol, available elsewhere.[1]

Our model is grounded in a theoretical framework in which the incidence of unintended pregnancy is a function of the numbers of women with an unmet need for contraception and women using a contraceptive method who experience a method or user failure, separately by marital status, and the risk of pregnancy in each of these population groups *(see Figure 1)*. Similarly, the incidence of intended pregnancy is a function of the number of women with no need for contraception, separately by marital status, and their risk of pregnancy.

Thus, the number of pregnancies Ω to occur in country *c* during five-year time period *t* is equal to the sum of pregnancies across all population groups. Algebraically, where Ω^f is the number of pregnancies to occur in population group *f*, $\Omega_{ct} = \sum_f \Omega_{ct}^f$.

The number of pregnancies to occur in a population group is in turn a function of the number of women in that group, w_{fct}, and their risk of pregnancy, ω_{fct}:

$$\Omega_{ct}^f = w_{fct}\ \omega_{fct}.$$

[1] Bearak J, Popinchalk A, Sedgh G, Alkema L.. "Pregnancies, abortions, and pregnancy intentions: a protocol for modeling and reporting global, regional and country estimates." *Reproductive Health* 2019: https://doi.org/10.1186/s12978-019-0682-0

The incidence of abortion within a population group, Ψ^f, is a function of the numbers of pregnancies in that group and the group-specific probability that a pregnancy will end in an abortion, α_f:

$$\Psi_{ct}^{f} = \Omega_{ct}^{f}\, \alpha_{fct}.$$

The incidence of abortion in a country-period is in turn the sum of the numbers of abortions across population groups, $\Psi_{ct} = \sum_f \Psi_{ct}^f$. Alternatively, replacing Ψ^f with the above equations, the incidence of abortion can be expressed as the summation across all population groups of the product of the number of women, the risk of pregnancy, and the probability that a pregnancy ends in abortion,

$$\Psi_{ct} = \sum_f w_{fct}\, \omega_{fct}\, \alpha_{fct}.$$

Pregnancy outcomes are given by abortions, live births, or miscarriages. In our model framework, live births (figure 1, 4th column) are given by UNPD estimates.[2] Consistent with previous pregnancy estimates,[3,4] we estimate miscarriages using an approach derived from life tables of pregnancy loss by gestational age in which there is, on average, one miscarriage for every ten abortions, and one for every five live births.[5,6,7].

Marital status, contraceptive need and use, and abortion are key proximate determinants of pregnancy rates and fertility.[8] However, the sizes of these population groups will not explain all differences between time periods or between countries. The risk of pregnancy in these population groups can be influenced by women's fecundity and the timing and frequency of their sexual activity.[8] Additionally, the percent of unintended pregnancies which end in abortion may vary according to differences in women's motivation to avoid an unintended birth, social and personal stigma, and concrete obstacles to abortion access.[9] Therefore, we will consider covariates which may proxy these factors. Available covariates are unlikely to be able to explain all variability across countries and within countries over time in pregnancy rates and probabilities of aborting an unintended pregnancy for two main reasons. First, information on determinants is limited, i.e. available covariates will be proxy covariates at best. Second, covariates may be estimated imperfectly and are subject to uncertainty. As a result, there will be unexplained heterogeneity across countries and within countries over time.

---

[2] Alkema L, Kantorova V, Menozzi C, Biddlecom A. 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*. 2013: https://doi.org/10.1016/S0140-6736(12)62204-1

[3] Sedgh G, Bearak J, Singh S, Bankole A, Popinchalk A, Ganatra B, et al. Abortion incidence between 1990 and 2014: global, regional, and subregional levels and trends. *The Lancet* 2016: https://doi.org/10.1016/S0140-6736(16)30380-4

[4] Bearak J, Popinchalk A, Alkema L, Sedgh G. Global, regional, and subregional trends in unintended pregnancy and its outcomes. *The Lancet Global Health* 2018: https://doi.org/10.1016/S2214-109X(18)30029-9

[5] Dellicour S, Aol G, Ouma P, Yan N, Bigogo G, Hamel MJ, et al. Weekly miscarriage rates in a community-based prospective cohort study in rural western Kenya. BMJ Open. 2016: https://bmjopen.bmj.com/content/6/4/e011088

[6] Bongaarts J. Modeling the fertility impact of the proximate determinants: Time for a tune-up. *Demographic Research*. 2015: https://dx.doi.org/10.4054/DemRes.2015.33.19

[7] Bongaarts J, Potter R. Fertility, Biology, and Behavior: An Analysis of the Proximate Determinants. New York: Academic Press; 1983.

[8] Bongaarts J. A Framework for Analyzing the Proximate Determinants of Fertility. *Population and Development Review* 1978: https://www.jstor.org/stable/1972…

[9] Rossier C, Michelot F, Bajos N, COCON Group. Modeling the process leading to abortion: an application to French survey data. *Studies in Family Planning*. 2007: https://www.ncbi.nlm.nih.gov/pubmed/17933290

We will address the issue of unexplained heterogeneity in our outcomes—subgroup estimates of pregnancy rates and propensities to abort— with a Bayesian hierarchical time series model. After accounting for covariates, we expect temporal correlations in the unexplained fluctuations. This will be captured through a time series model on subgroup outcomes. Similarly, we expect similarities across countries within subregions in the unexplained fluctuations. We will use a hierarchical model to estimate country parameters, such that information is exchanged across countries within the same group. Countries in which the statistical relationships are expected to be similar will be grouped together, and these may differ from geographic subregions.

We will use a Bayesian framework to (i) implement the modeling strategy for the unknown outcomes as explained above, and (ii) incorporate all available data, as well as the uncertainty associated with each datum. Estimates for pregnancies will be consistent with information on pregnancy outcomes, i.e. the total of abortions, live births, and associated miscarriages. The model will include data on abortion incidence, the percent of live births that were intended, and data on the distribution of outcomes by population group to calibrate the group-specific rates. The Bayesian approach will produce point estimates that combine information directly from data for the respective country-period with information from other periods and countries. Uncertainty intervals around each of our estimates account for the quantity and quality of all available data, as well as the unexplained heterogeneity across countries and periods.

Given the data limitations, a standard regression-based approach to estimation may produce questionable results. However, it remains important to assess the performance of our theoretically grounded approach. We will conduct simulation and validation exercises to explore this.

Figure 1. Theoretical Framework