Methodologies for Adjustments to Infant Mortality Estimates

Submission for 2007 Annual Meeting of the Population Association of America

Dennis Donahue
U.S. Census Bureau
Population Division
International Programs Center

The purpose of this paper is to encourage discussion about useful tools for estimating and adjusting infant mortality. Obtaining accurate estimates of infant mortality is essential, not only because it is such a reactive gauge for other demographic trends (e.g., life expectancy), but also because policymakers and the general public have begun to recognize infant mortality as a key indicator of social development and general population health. Unfortunately, our understanding of infant mortality at the global level continues to be limited by two major problems, underreporting and incomparability.

Many regions of the world lack the resources to monitor and record vital events such as infant deaths adequately, as vital registration systems remain underdeveloped or nonexistent. As such, many births are not registered until many months or years later (if at all), and any deaths that may have occurred to infants also remain unregistered. Despite efforts by the World Health Organization (WHO) and others, standards and definitions are not consistent (or consistently followed), resulting in further uncertainty about our ability to estimate, much less compare, the level of infant mortality across regions.

Great effort has been expended to address these issues using censuses and surveys to gather retrospective reports of deaths. In recent years, additional efforts have focused on large-scale surveys like the Demographic Health Surveys program (DHS), the Reproductive Health Surveys (RHS), and Multiple Indicator Cluster Surveys (MICS). These, and other surveys, have done much to make up for the underreporting due to inadequate national vital registration systems. Still, comparisons between the different types of surveys, as well as limitations on sample sizes and sampling frames have left a wide range of uncertainty about the true level of infant mortality in many countries. Our capacity to compare estimates produced by national statistical agencies with estimates from surveys is limited by the scarcity of countries that have both types of data available. In virtually every case where we have both, the nationally produced infant mortality rate (IMR) is considerably lower than that estimated from the survey results.

Several statistical methods have been developed over the years to attempt to model the true level of infant mortality from partial and sometimes conflicting data. Some of the models are designed to deal with issues relating to underreporting, while others were developed specifically to resolve differences in definitions or standards between national statistical agencies and organizations such as WHO (predominantly in Eastern Europe and the Former Soviet Union). This paper reviews several models to evaluate whether we have an improved understanding of infant mortality today than we had previously. The figures on the following pages illustrate the range of estimates produced by the adjustment methods under evaluation.
Depending on the chosen data source and particular adjustment methodology, the estimated number of infant deaths in Ukraine for the years 1989 to 2002 varies widely. Note that the RHS estimates nearly match the official estimates from the Ukraine Ministry of Statistics, and that the Kingkade and Sawyer model suggests a widening gap between reported and actual infant deaths. Given that the Kingkade and Sawyer adjustments were designed to specifically address misreporting, and Ukraine’s vital statistics system is believed to have improved to the point that completeness and coverage are nearly 100 percent (WHO, 2006), the discrepancy between the model and the reported data is troubling. Limited data from surveys are available for comparison or corroboration, as the only available survey, the 1999 Ukraine Reproductive Health Survey, provides only one point estimate for a broad time period.
Calculated infant mortality rates for Moldova are shown for the period 1989 through 2003. In this example, the survey results (DHS 2005) correspond to the reported IMRs in the early 1990s and in recent years, but the survey suggests the IMR was much higher in the late 1990s than reported. The higher survey estimate nearly matches the IMR estimate from Kingkade and Sawyer. Other experimental models show near-universal agreement with the reported data.
Another example from the same region of the world, with relatively similar experiences in recent social and statistical events as the previous two examples, shows that trends, whether in actual experience or only in reporting, do not follow predictable patterns. Survey results from UNICEF (United Nations Children’s Fund) suggest a progressive worsening of reported data (i.e., the survey estimate diverges further from reported estimates over time). Contrarily, the Kingkade and Sawyer method, which had suggested a high level of misreporting in the two previous examples, now shows estimates that are nearly identical to the reported data.
One advantage of the Hogan Method is that it allows the analyst to evaluate whether the proposed adjustment factor is being applied to the appropriate age category. In cases of misreporting, where it is believed that infant deaths are being misreported as stillbirths or miscarriages, results similar to those shown in the figure above are expected. Most of the misreported events are likely to have occurred to very young infants. In a country where underreporting is suspected to be widespread (e.g., in a population where children often remain unreported for years), the adjustment factors could be expected to be above 1.0 for all age groups. The disadvantage of this method, though to a lesser extent than the Kingkade and Sawyer method, is the requirement for detailed age of death data for at least one year. However, this level of detail is often available from surveys and censuses that collect mortality information, though most surveys would have an insufficient sample for this purpose.


