Global Subnational Distribution of Infant Mortality Rates

Dara Mendeloff
Susana B. Adamo
Marc Levy
Alex de Sherbinin
CIESIN, Columbia University[1]

Abstract

This paper present the rationale, methodology and data sets used to construct the 2015 Global Subnational Infant Mortality Rates (IMR). The dataset compiles IMR data for 234 countries and territories, 143 of which include subnational units. The data are benchmarked to the year 2015, and are drawn from National Statistical Offices (NSOs), Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), and other sources. Boundary inputs are derived primarily from the Gridded Population of the World, Version 4 (GPWv4). National and subnational data are mapped to grid cells at a spatial resolution of 30 arc-seconds (~1 km) allowing for easy integration with demographic, environmental, and other spatial data. This second version of Global Subnational Infant Mortality Rates has many potential applications not only for the study of the geographic distribution of infant mortality within countries, but also as a proxy for the analysis of poverty and living conditions in general, in conjunction with environmental or other geographical factors.

Introduction

Goal 3 of the SDGs refers to "ensure healthy lives and promote well-being for all at all ages", and target 3.2 calls for ending preventable deaths of newborns and children under 5 years old. Pertinent to all SDGs, the "Leaving none behind" challenge includes tackling 'spatial disadvantage' (Kabeer 2016) and calls for data disaggregation not only by sociodemographic characteristics (age, gender, ethnicity, etc.) but also at higher spatial resolutions, particularly at the subnational level. As Storeygard et al. (2008:209) note, "While national data show a similar pattern for infant deaths, with half concentrated in the same six countries, spatially explicit subnational data would permit further quantification of spatial concentration and identify geographical 'hotspots' or clusters of mortality, some of which would be expected to cut across national boundaries". Country-level figures could mask large subnational heterogeneities and diverging trends (e.g. Golding et al. 2017).

[1] Acknowledgments: Funding for development and dissemination of this data set was provided under the U.S. National Aeronautics and Space Administration (NASA) contract NNG13HQ04C for the continued operation of the Socioeconomic Data and Applications Center (SEDAC), which is operated by the Center for International Earth Science Information Network (CIESIN) of Columbia University.

A georeferenced global subnational dataset of infant mortality rates has many potential applications not only for the study of the geographic distribution of infant mortality within countries, but also as a proxy for the analysis of poverty and living conditions in general, in conjunction with environmental or other geographical factors (e.g. Gonzalez and Gilleskie 2017; Zewdie and Adjiwanou). The infant mortality rate (IMR) for a region or country is of interest to a wide user community in interdisciplinary studies of health, development, sustainability, and the environment. Subnational IMR estimates derived from vital statistics and other sources are frequently used as a proxy for indicators such as poverty and wellbeing since alternative measures, such as gross domestic product or population living on less than one U.S. dollar per day, can be difficult to obtain at sub-national levels for many countries (Dasgupta, 1993).

IMR also has a number of advantages over other metrics. First, its measurement, unlike small area estimates of poverty or income, is highly standardized. Second, when compared with average income metrics, IMRs are less likely to be influenced by the kind of skewed wealth distribution that might make otherwise poor areas appear well-off because a few high income people live there. Third, the data are available for 90% or more of the population in medium and low income countries (Balk et al., 2006). In addition, it could also be a valuable input into other new or existing data sets.

This new version of SEDAC's Global Subnational Infant Mortality Rates data set includes infant mortality rate (defined as the number of children who die before their first birthday for every 1,000 live births) data for the lowest administrative units available for each country as of June 2017. It compiles IMR data for 234 countries and territories, 143 of which include subnational units. The data are benchmarked to the year 2015 (version 1 of the data set was benchmarked to the year 2000 (CIESIN, 2005)). Data are drawn from National Statistical Offices (NSOs), Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), and other sources from 2008 to 2014. Boundary inputs are derived primarily from the Gridded Population of the World, Version 4 (GPWv4). Assigning both national and subnational data to grid cells at a spatial resolution of 30 arc-seconds (~1 km) allows for easy integration with demographic, environmental, and other spatial data. Metadata are included describing methods used and uncertainties related to temporal and spatial scale adjustment, small island state (SIS) status, and conflict regions.

Data and Methodology

Input data

IMR data has been collected for 234 countries and territories, 91 include only national units, while 143 include subnational units, representing an improvement of 78 additional countries with subnational data compared to the first version. Distribution of countries by level of administrative unit used and total number of units for each level is displayed in table X.

Table 1:

Adm level	Count of Highest Resolution SNU**	Sum of Total Number of SNUs**
0	91	91
1	113	1868
2	22	3513
3	7	1439
4	1	163
Grand Total	234	7074

Source information by country, including the number of reporting units for each country, is shown in Table 1. Subnational data were drawn from national offices (82 countries), DHS (41 countries), MICS (17 countries), and other sources (3 countries). For the national office source, data are reported by either a national Human Development Report, national statistical office, or ministry of health. Subnational data dates from 2008 to 2014. One exception is the Islamic Republic of Iran, which only has subnational IMR data for the year 1996, but has national IMR for the year 2014.

Spatial boundaries were obtained from the same IMR data source when possible, otherwise GPWv4 input boundaries were used (CIESIN, 2016a), shown in Table 4. Boundaries for countries and territories where no IMR data has been collected (7 countries) are represented with a no data or no population code where applicable. Water features were also obtained from GPWv4 (CIESIN, 2016b) and are represented using the code for no population. Using the GPWv4 boundaries improves interoperability between IMR and the various GPW data such as population count, age, and sex variables (CIESIN, 2017)

Table 4:

Boundary Source	Number of Countries
GPWv4	192*
DHS	42
GADM	6
IGBE	1
Grand Total	241

*Includes 7 countries and territories that do not have IMR data (Åland Islands, British Indian Ocean Territory, Falkland Islands (Malvinas), French Southern Territories, Heard and Mcdonald Islands, South Georgia and the South Sandwich Islands, Svalbard and Jan Mayen Islands)

Methods

The infant mortality rates were either "Acquired" or "Calculated by CIESIN". "Acquired" means that IMRs were collected from vital registration data, surveys, models, or other estimates.

"Calculated by CIESIN" indicates that the rates were estimated using reported live births and infant deaths data. IMRs were calculated by taking the number of infant deaths, dividing them by the number of live births, and multiplying the result by 1,000.

$$IMR = (deaths\ of\ infants\ less\ 1\ year\ old\ /\ live\ births) * 1000$$

Out of all *national* IMRs in the database, 90 countries use "acquired" data. Only 1 (Bermuda) uses estimates calculated by CIESIN. Out of all *subnational* IMRs, 135 countries use "Acquired" data, and eight use estimations "Calculated by CIESIN".

The IMR input data comes from different sources and spans from 2008 to 2014 varying by country and subnational unit. Because of this, the rates were adjusted for year and source to be consistent with the United Nations (UN) Inter-agency Group for Child Mortality Estimation (UN IGME, 2017) national estimates for the year 2015. For those countries not included in UN IGME 2017 report, data from the US Census Bureau International Database was used for imputing the UN IGME missing values. In the end, 195 countries were adjusted to UN IGME national estimates, 33 were imputed using the US Census Bureau International Database, and six countries or territories were not adjusted.

The adjustment approach follows the method used in the first version of the infant mortality data set (Storeygard et al., 2008:216), as follows:

> The IMR value $r_{c,y,s,x}$ for country *c*, year *y* from source *s*, at scale *x* (0 = national; 1 = subnational) were scaled to *y* = 2015 national rates from UN IGME (denoted s = 0).

The highest subnational resolution of IMR was included in the final gridded product. If the unit had an adjusted value, it was used; otherwise an unadjusted value was used.

For those countries were IMR was calculated, administrative units with low number of deaths (less than XX) –usually associated with small island nations or subnational units with small populations—were assigned the IMR of the administrative unit where they were included. This step was not available for the acquired IMRs because the number of infant deaths going into the calculation of the infant mortality rate was generally not known.

Topology was applied such that there are no gaps or overlaps between country or subnational unit boundaries. The data were then gridded to a spatial resolution of 30 arc-seconds which is about 1 kilometer at the equator. The final gridded product represents the estimated IMR at the highest subnational resolution available for the year 2015.

Coded values are included for no data (-9999), and for no population or uninhabited areas including bodies of water (-7777).

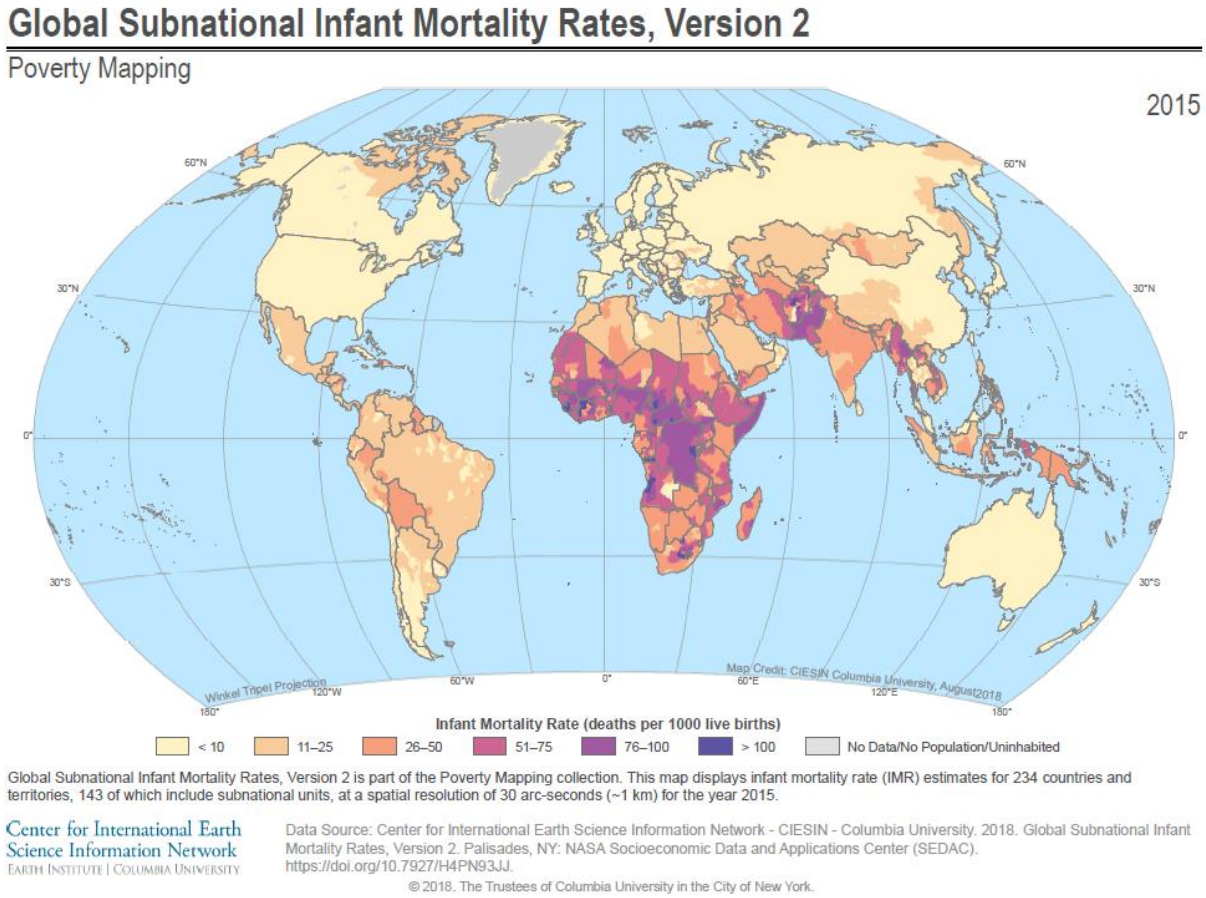
Preliminary Output

Data Set Description

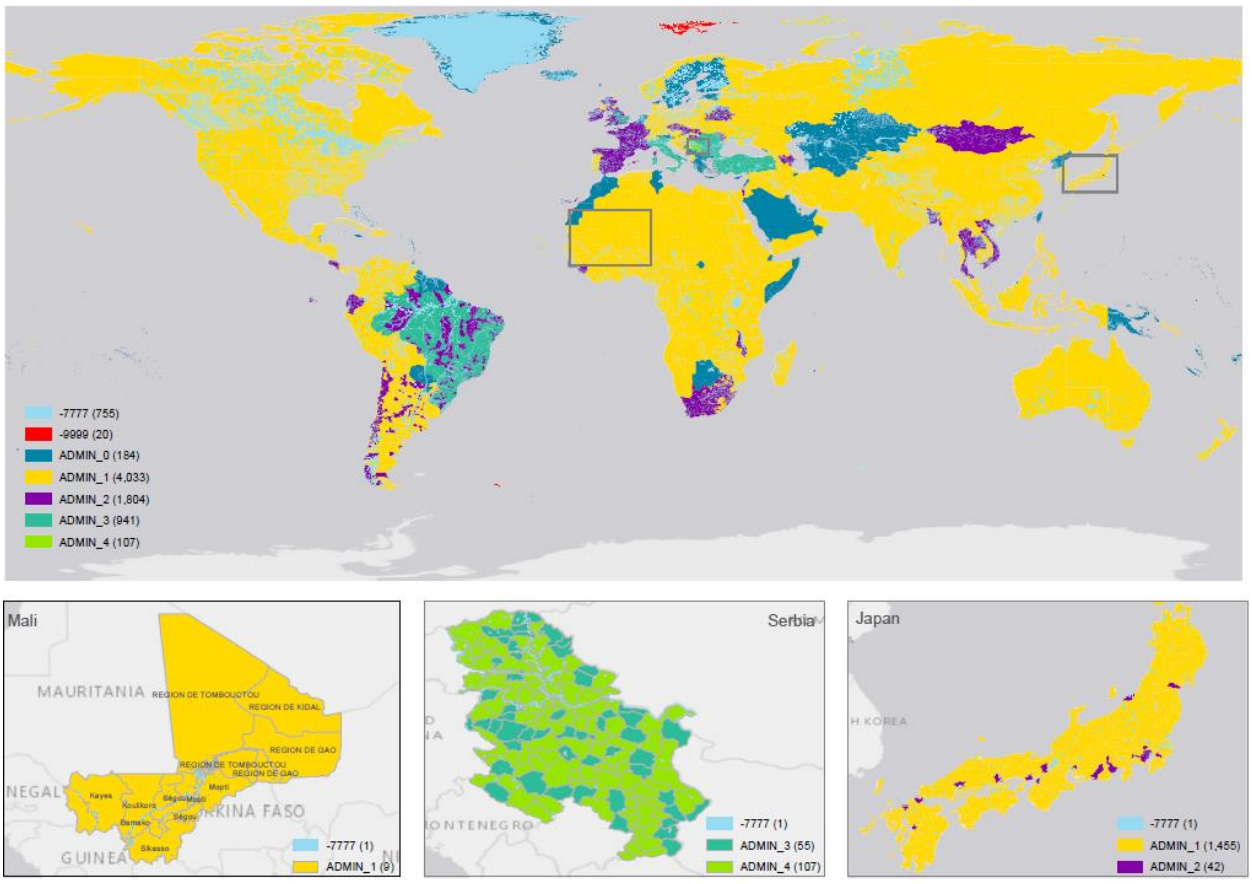
The Global Subnational Infant Mortality Rates, Version 2 consists of estimates of infant mortality rates for the year 2014. The data product is a grid (raster data) of IMR rates at 30 arc second (~1 km) resolution. The nominal resolution of the vector input layers varies from country to country. The data are available as a global grid in GeoTIFF format. The downloadable is a compressed zip file, which contains: 1) Global GeoTiff for the year of estimate, and 2) PDF documentation. The dataset is expected to be available in November 2018 from the SEDAC website[2]. The raster data are in GeoTIFF format and can be used directly in mapping and geospatial analysis.

Map 1 illustrates the global distribution of IMR and Map 2 illustrates the data quality.

[2] Disclaimer: CIESIN follows procedures designed to ensure that data disseminated by CIESIN are of reasonable quality. If, despite these procedures, users encounter apparent errors or misstatements in the data, they should contact SEDAC User Services at ciesin.info@ciesin.columbia.edu. Neither CIESIN nor NASA verifies or guarantees the accuracy, reliability, or completeness of any data provided. CIESIN provides this data without warranty of any kind whatsoever, either expressed or implied. CIESIN shall not be liable for incidental, consequential, or special damages arising out of the use of any data provided by CIESIN.



Map 1: 2015 Global Subnational Infant Mortality Rates



Map 2: Administrative level used for IMR (data quality)

Limitations

The Global Subnational Infant Mortality Rates data set has a few limitations. The quality of IMR data varies globally in terms of defining, reporting live births and infant deaths, as well as methods of calculating rates. Sampling errors, data omission, and misreporting are common examples of how the input data may be skewed (Storeygard et al., 2008).

Some sources are unclear as to whether births are reporting stillborn and live births or live births alone. Also, methods of calculation of the infant mortality rates differs between sources, and from the approach followed in those calculated by CIESIN in those cases where the rates was not available directly from the source.

DHS and MICS survey-based data may have sampling errors and bias related to the survey design, sample size and temporal range, as both report IMRs for five (national) or ten (subnational) year periods.

References

Balk, D., Deane, G.D., Levy, M.A., Storeygard, A., & Ahamed, S. 2006. The Biophysical Determinants of Global Poverty: Insights from an Analysis of Spatially Explicit Data. Paper presented at the 2006 Annual Meeting of the Population Association of America, Los Angeles, CA, 30 March-1 April 2006.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2017. Gridded Population of the World, Version 4 (GPWv4): Basic Demographic Characteristics, Revision 10. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H45H7D7F.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2016a. Gridded Population of the World, Version 4 (GPWv4): Population Count. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4X63JVC.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2016b. Gridded Population of the World, Version 4 (GPWv4): Land and Water Area. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H45M63M9.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2005. Poverty Mapping Project: Global Subnational Infant Mortality Rates. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://doi.org/10.7927/H4PZ56R2.

Dasgupta, P. 1993. An Inquiry into Well-Being and Destitution. Clarendon Press, Oxford.

Golding, N., Burstein, R., Longbottom, et al. (2017). Mapping under-5 and neonatal mortality in Africa, 2000–15: a baseline analysis for the Sustainable Development Goals. *The Lancet, 390*(10108), 2171-2182. doi:https://doi.org/10.1016/S0140-6736(17)31758-0

Gonzalez, R. M., & Gilleskie, D. (2017). Infant Mortality Rate as a Measure of a Country's Health: A Robust Method to Improve Reliability and Comparability. *Demography, 54*(2), 701-720. doi:10.1007/s13524-017-0553-7

Kabeer, Naila. 2016. 'Leaving no one behind': the challenge of intersecting inequalities. In ISSC, IDS and UNESCO (2016),

Storeygard, A., D. Balk, M. Levy, and G. Deane. 2008. The Global Distribution of Infant Mortality: A Subnational Spatial View. *Population, Space and Place* 14:209–229. https://doi.org/10.1002/psp.484.

U.S. Census Bureau. International Programs. International Data Base. Revised: December 22, 2017. Version: Data:17.0810 Code:12.0321

UN IGME. 2015. Levels and Trends in Child Mortality Report 2015. Estimates Developed by the UN Interagency Group for Child Mortality Estimation. Accessed 9 Sept. 2015. http://www.unicef.org/media/files/IGME_report_2015_child_mortality_final.pdf

World Social Science Report 2016, Challenging Inequalities: Pathways to a Just World, UNESCO Publishing, Paris. Pp. 55-58. http://unesdoc.unesco.org/images/0024/002459/245935e.pdf

Zewdie, S. and V. Adjiwanou. 2017. Multilevel analysis of infant mortality and its risk factors in South Africa. International Journal of Population Studies. 3(2):43-56. ojs.whioce.com/index.php/ijps/article/download/330/354