

# Strengthening small area demographic estimates in Latin America

## The importance of small area demographic estimates

- For reliable mortality and fertility estimates to monitor progress in health and population against national targets.
- Local governments have increased responsibilities because of decentralisation and hence require the evidence to support their decision-making.
- The least populous municipalities are commonly in remote areas with disadvantaged populations with poor health outcomes.
- Accurate population data are necessary for making accurate population projections, which are vital for planning purposes across many sectors.

## The need for small area demographic estimation methods

Accurate mortality, fertility and population information are fundamental to program and policy development and monitoring in health and other sectors. The most suitable data source for local death and birth data is a well-functioning Civil Registration and Vital Statistics (CRVS) system, but these are underdeveloped in Latin America, with varying levels of incompleteness between and within countries.<sup>1</sup> Other data sources, such as sample surveys, censuses and administrative data have several positive attributes but also suffer from limitations such as lack of representativeness, sampling error and lack of timeliness.<sup>2</sup>

The **Small Area Demographic Estimation in Latin America Workshop** was held in Rio de Janeiro, Brazil from 19-20 March 2018. The workshop was facilitated by the Bloomberg Philanthropies Data for Health (D4H) Initiative and brought together regional and international demographic experts to discuss methods and approaches to improve the evidence base for mortality, fertility and population at the small area level in Latin America.

The participating countries were Brazil, Colombia, Ecuador and Peru, Australia and the USA.

The objectives of the Workshop were to:

- Describe countries' policy needs for demographic indicators for small areas and challenges
- Examine the strengths and limitations of a range of methods to estimate registration completeness and demographic indicators at the small area level
- Discuss the lessons learned from the application of methods to estimate completeness of registration and demographic indicators at the small area level
- Discuss the applicability of these methods in all countries and potential areas of future development.

## Policy importance of small area estimation

### Definition of 'small area'

Each of the countries defines small area as being municipalities. The heterogeneous nature of each country's population means that there can be huge variations in the populations of the municipalities. For example, in Brazil populations range from 12 million in Sao Paulo to just 800 in the smallest municipalities, and in Colombia 38 per cent of municipalities have populations of less than 10,000 people.

### Challenges in making these estimates

One challenge is generating reliable estimates for very small populations. As mortality is a statistically rare event, in small populations it is important to disentangle real differences and random fluctuation caused by small numbers (ie stochastic variation). There is often a need to combine multiple years of data or to aggregate age groups, sexes or even municipalities, for reliability.

A common issue is with the quality of data for both mortality and fertility, and population, which in turn affects quality of mortality rate estimates. This is because of low registration

1 Mikkelsen L, Phillips DE, AbouZahr C, Setel PW, de Savigny D, Lozano R, Lopez AD: A global assessment of civil registration and vital statistics systems: monitoring data quality and progress. *Lancet* 2015, 386:1395-1406.

2 World Health Organization: Improving mortality statistics through civil registration and vital statistics systems: strategies for country and partner support; outcome of a technical meeting. Geneva, 4-5 November, 2014: World Health Organization; 2014.

completeness, particularly in the smallest municipalities where registration infrastructure is poorest. Population data in large cities are commonly of poor quality because it is difficult to accurately measure migration. In Colombia, the 2005 census was the only census conducted between 1985 and 2018, which severely hampered the ability to make accurate estimates of the population.

Another limitation is a lack of consistency in methods used by national statistical offices to estimate mortality and fertility. Conventional methods to estimate registration completeness and generate mortality estimates are still commonly used in all countries even though they rely on demographic assumptions such as no migration.<sup>3,4,5,6</sup> There is a need to strengthen capacity in statistical offices and ministries of health, especially given the technical challenges of estimating demographic indicators at the small area level with data of suboptimal quality.

## Advanced demographic and statistical methods for small area estimation

### Estimating mortality in small areas

One method for estimating small area mortality uses a Bayesian poisson model using death, population and covariate data to estimate life expectancy. The method by Dwyer-Lindgren et al has been applied at the county level in the United States.<sup>7,8</sup> The method does not need multiple counties or years and can describe trends and spatial patterns and produces less bias than other similar methods. The method also facilitates the estimation of cause-specific mortality where garbage causes of death (ie those causes with limited policy value) are redistributed to other causes.

Another approach for small area mortality estimation has been used in Brazil.<sup>9,10</sup> The method uses a Bayesian model with a prior of local registration completeness based on published studies. It also uses the TOPALS regression, a relational model developed for smoothing and projecting age-specific probabilities of death for small samples/variance to produce complete life tables. This model was found to be insensitive to the standard used (eg from Human Mortality Database). This approach has a strong advantage in being able to adjust for incompleteness of registration and has been demonstrated in a Latin American country.

### Measuring death registration completeness

A new method to estimate death registration completeness was discussed which avoids the pitfalls of older methods such as being reliant on implausible demographic assumptions and having poor accuracy for subnational areas. The method by Adair et al uses the inputs registered crude death rate, under-five mortality rate, population age structure and under-five death registration completeness.<sup>4</sup> The method was demonstrated for use in Brazilian municipalities in Amazonas (where completeness ranged from 50% to 96%), for all Brazilian states, Colombian departments and Peruvian regions. The method's utility is diminished where the adult mortality rate is unusually high for a given under-five mortality rate, or where there is considerable uncertainty in the under-five mortality rate. The method is a useful tool for generating mortality statistics and for local CRVS managers to monitor temporal and spatial differences in completeness.

### Estimating fertility in small areas

A range of methods for estimating fertility at the national and sub-national level from limited data was discussed, and their application to Brazil were presented. These methods included the Synthetic Relational Gompertz method, Brass P/F method, modified Brass P/F method and Own Children method. The most plausible estimates of fertility were found to be from a combination of the Synthetic Relational Gompertz model and Brass P/F methods. Other methods such as the Own Children modified Brass P/F method also provided robust estimates.

3 Adair T, Lopez AD: Estimating the completeness of death registration: An empirical method. *PLoS One* 2018, 13:e0197047

4 Dorrington R: The Generalized Growth Balance method. In *Tools for demographic estimation*. Edited by Moultrie T DR, Hill A, Hill K, Timaeus I, Zaba B. Paris: IUSSP; 2013: 258-274

5 Dorrington R: The Brass Growth Balance method. In *Tools for demographic estimation*. Edited by Moultrie T DR, Hill A, Hill K, Timaeus I, Zaba B. Paris: IUSSP; 2013: 196-208

6 Murray CJ, Rajaratnam JK, Marcus J, Laakso T, Lopez AD: What can we conclude from death registration? Improved methods for evaluating completeness. *PLoS Med* 2010, 7:e1000262.

7 Dwyer-Lindgren L, Bertozzi-Villa A, Stubbs RW, Morozoff C, Kutz MJ, Huynh C, Barber RM, Shackelford KA, Mackenbach JP, van Lenthe FJ, et al: US County-Level Trends in Mortality Rates for Major Causes of Death, 1980-2014. *JAMA* 2016, 316:2385-2401.

8 Dwyer-Lindgren L, Stubbs RW, Bertozzi-Villa A, Morozoff C, Callender C, Finegold SB, Shirude S, Flaxman AD, Laurent A, Kern E, et al: Variation in life expectancy and mortality by cause among neighbourhoods in King County, WA, USA, 1990-2014: a census tract-level analysis for the Global Burden of Disease Study 2015. *Lancet Public Health* 2017, 2:e400-e410.

9 Gonzaga MR, Schmertmann CP: Estimating age- and sex-specific mortality rates for small areas with TOPALS regression: an application to Brazil in 2010. *Revista Brasileira de Estudos de População* 2016, 33.

10 Schmertmann CP, Gonzaga MR: Bayesian Estimation of Age-Specific Mortality and Life Expectancy for Small Areas with Defective Vital Records. *Demography* 2018, 55:1363-1388.

## Estimating population in small areas

A method for population estimation in small areas in Colombia using a Bayesian hierarchical model was discussed. Due to small population sizes and lack of data availability the estimates for small areas were challenging to compute. Bayesian methods are an alternative to population estimation and can use prior information from multiple administrative sources. The method was demonstrated for the Colombian municipality of Jamundi and showed it can be replicated in other countries.

## Areas for future development

The following were discussed as areas for improvement in methodological development for small area demographic estimation:

- Statistical agencies such as IBGE should strongly consider the use of Bayesian models
- Any method used needs to correct for poor quality registration and population data
- Improved measurement and modelling of immigration, which significantly impacts on small area population estimates
- In more remote areas, satellite imagery could be a means of estimating population
- More partnerships between academia and government
- Methodological developments should precede the policy need for them to prevent a lag in reliable evidence being available.

## Currently used methods for small area estimation in Latin America

### Fertility and mortality

In Peru, fertility and the completeness of birth registration are currently estimated using the P/F ratio based on data from the census. The Brass methods are used to estimate infant mortality, completeness of death registration and to generate complete life tables and life expectancy. These are traditional methods to estimate mortality and have been superseded by more accurate methods.<sup>11,12</sup> In particular, their estimate of completeness is producing biased measures as it assumes a stable population (i.e. closed to migration and with constant birth and death rates). Completeness is also likely much higher.

In Ecuador surveys are used for early age mortality and have resulted in completeness of death registration of 88 per cent for men and 82 per cent for women using intercensal methods.

### Estimating birth registration in Columbia

In Columbia birth registration completeness in small areas was conducted in a novel manner by comparison with the Expanded Program on Immunisation (EPI) administrative data. Birth registration completeness was 97 per cent in 2014 having increased from 91 per cent in 2010. The number of registered births was less than the number of BCG vaccinations received. Adjusted registration data were used to estimate fertility.

### Innovations in Brazil

Brazil is somewhat unique in having two sources of data, civil registry data from IBGE and deaths (SIM) and births (SINASC) data from Ministry of Health, that both have high levels of completeness. When linked the data showed 98 per cent of IBGE deaths captured by SIM and 95 per cent of SIM deaths captured by IBGE; and 98 per cent of IBGE births captured by SINASC and 97 per cent of SINASC births captured by IBGE. By municipality the highest levels of incompleteness was in northern Brazil. A drawback of this however is that there potentially is dependence between the sources, which may bias estimates upwards.

11 Wang H, Dwyer-Lindgren L, Lofgren KT, Rajaratnam JK, Marcus JR, Levin-Rector A, Levitz CE, Lopez AD, Murray CJ: Age-specific and sex-specific mortality in 187 countries, 1970-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012, 380:2071-2094.

12 Rajaratnam JK, Tran LN, Lopez AD, Murray CJ: Measuring under-five mortality: validation of new low-cost methods. *PLoS Med* 2010, 7:e1000253.

An innovative approach to improve registration involved a proactive search of deaths and live births in areas where registration completeness is particularly low. Births and deaths that were not reported to the Ministry of Health were actively identified from various sources such as community health agents, registry offices, official and non-official cemeteries and funeral homes.<sup>13</sup> From the data mortality correction factors were calculated, which have been applied to all of Brazil and has since been an important tool for calculating correction factors of vital statistics. It is a labour-intensive process but valuable for countries like Brazil with high completeness. Additionally, the search for uninformed events to the official sources became part of the routine of capturing vital events in some municipalities, contributing to the improvement of data quality.

To maintain high levels of registration municipalities are required to reach at least 80 per cent of expected events within 60 days of the end of the month or are penalised. This approach has resulted in a reduction of days from event to reporting. A complimentary approach would be to use the empirical completeness method, which estimates completeness by allowing for different levels of mortality (as measured by the under-five mortality rate) and different population size of the municipality.

## Improving the quality of demographic evidence at the small area level in Latin America

The workshop outlined the strong need for reliable demographic evidence for small areas in Latin America. However, there are several challenges to attaining this objective, including the very small size of populations in some municipalities, the poor quality of registration and population data, the methods applied by national statistical agencies and human resource capacity to apply more advanced methods. More specifically:

- The empirical completeness method can estimate completeness of registration using relatively limited data available at the subnational level.
- The Bayesian methods have the advantage of being able to utilise data from a range of sources, measure uncertainty in resultant rates, to smooth data across age, time and space, and correct mortality data for incompleteness.

- The Bayesian approaches would benefit from incorporating the empirical completeness methods in their modelling process.
- The Bayesian population projection methods presented for Colombia demonstrate how multiple sources can be used to make demographic estimates for small areas.
- The use of administrative data to assess registration completeness and adjust rates should be encouraged where data can be readily linked.
- Importantly, these methods used should be responsive to needs of national and local governments.

Investments in human resource capacity are being done through D4HI Fellowships. Fellows from Ecuador (INEC), Brazil (IBGE and Ministry of Health) and Peru (INEI) have spent time at University of Melbourne to learn methods of completeness estimation. A training course in completeness methods was conducted for DANE and Ministry of Health staff in Colombia in November 2018, while further training programs in Bayesian population estimation were attended by a DANE consultant. The regular sharing of methods and experiences between statistical agencies would also be of much benefit. There should be improved partnerships between academia and government to facilitate the sharing of new approaches.

The most effective way to improve small area demographic evidence is to strengthen both CRVS and population data. The D4H Initiative is a significant investment in these improvements, the fruits of which should continue be borne in years to come. Continued national government investments will also be necessary to support this objective. Improvements in routine reporting systems will help provide policymakers with the required evidence to make informed decisions.

<sup>13</sup> Szwarcwald CL, de Frias PG, Junior PR, da Silva de Almeida W, Neto OL: Correction of vital statistics based on a proactive search of deaths and live births: evidence from a study of the North and Northeast regions of Brazil. *Popul Health Metr* 2014, 12:16.



## Summary

In small, sub-national areas reliable data on fertility, mortality and population are essential for policy and planning and for monitoring progress and development against national standards. In Latin America well-functioning civil registration and vital statistics (CRVS) systems are still under development, making these demographic data difficult to obtain. Small area demographic estimation methods can be used in the interim, as discussed at the Small Area Demographic Estimation in Latin America Workshop, held in Rio de Janeiro, Brazil from 19-20 March 2018.

Advanced methods for estimation of demographic factors in small areas were discussed including the use of various Bayesian approaches and empirical completeness methods. Currently used methods for fertility and mortality estimation are often traditional methods that have been superseded by newer approaches and could be updated. Some innovative approaches were presented, including linking with an immunisation program to update birth registration and implementing active processes to identify unregistered birth and death events. There is a strong need for investment in human resource capacity at the national statistical office level, one that is being met through training and fellowship programs by the Data for Health initiative and the University of Washington. Importantly, the application of advanced methods for small area demographic estimation should be responsive to the needs of local governments and would benefit from closer links between academia and government.

The workshop was organised as part of the Data for Health Initiative, which is funded by Bloomberg Philanthropies and the Australian Department of Foreign Affairs and Trade. Participants for the Workshop came from four D4H countries – Brazil, Colombia, Ecuador and Peru – as well as the United States and Australia.

The workshop was co-convened by Dr Tim Adair at the University of Melbourne, Dr Fatima Marinho and Dr Juan Cortez of the Ministry of Health, Brazil. Workshop organisation and support for participants was provided by IBGE, Ministry of Health Brazil, Vital Strategies-CRVS and Vital Strategies-Data Impact.

## For more information contact:

CRVS-info@unimelb.edu.au  
crvsgateway.info

CRICOS Provider Code: 00116K

Version: 0319-01