



All the contents of this journal, except where otherwise noted, is licensed under a Creative Commons Attribution License

Indirect estimation of infant mortality in small areas*

Ricardo Neupert**
Rogelio Eduardo Fernandez Menjivar***
Rogelio Eduardo Fernandez Castilla****

The Brass-type indirect methods of early-age mortality estimation have been used for more than four decades, providing very robust estimates for countries without reliable vital registration systems. However, when estimation areas become smaller, the number of dead children could be very small, especially among those born to young women, who provide essential information to estimate recent mortality. In these cases, estimates could be affected by random errors and unexpected annual fluctuations. At the same time, although it is very unlikely that demographic trends in a small area would follow patterns very different from those prevailing in the broader environment they belong to, it is possible that some local events may become relevant to small areas, causing some deviations from the assumptions that may hold true to the larger area. The objective of this paper is to propose an adaptation of the indirect estimation approach, which would allow obtaining infant and child mortality estimates for small areas. As such, this proposal belongs to the realm of indirect estimation methods, sharing the limitations and advantages that characterize this type of estimation procedures. The method is illustrated with data from the 2014 Population and Housing Census of Myanmar. The results indicate that the method proposed here provides reliable and consistent infant mortality estimates, compared to the original Brass' method, even in very small areas.

Keywords: Infant mortality. Indirect methods. Census. Small areas. Myanmar.

* The authors wish to thank the Department of Population, Ministry of Immigration and Population of Myanmar, specially Dr. U Nyi Nyi and Dr. Daw Khaing Khaing Soe, for their support during our work in Nay Pyi Taw, and facilitating the data utilized in this article. We thank also anonymous referees from REBEP for their useful comments on previous versions. The contents of this article are the sole responsibility of the authors.

** Independent consultant, Santiago, Chile (rneupert@hotmail.com; <https://orcid.org/0000-0002-5171-5255>).

*** Centro de Desenvolvimento e Planejamento Regional (Cedeplar), Universidade Federal de Minas Gerais (UFMG), Belo Horizonte-MG, Brazil (edufedz@gmail.com; <https://orcid.org/0000-0002-3956-1281>).

**** Independent consultant, Tucuman, Argentina (r.fernandez.castilla@gmail.com; <https://orcid.org/0000-0002-6138-0401>).

Introduction: nature of the proposed method and data sources

Mortality during childhood has received substantial attention especially in less developed countries, where a considerable number of deaths still occur in early life. Saving children less than 5 years of age, and especially less than 1 year old, is a major public health issue, calling for accurate statistics and rigorous measures to identify the extent of the problem and monitor progress (ROWLAND, 2003). Particularly important is the estimate of infant and child mortality¹ in small areas (or local areas). This term is used to define a small administrative area, such as a district, municipality or county. The demand for small area mortality statistics has grown enormously in recent decades, especially in the context of decentralized approach to health planning, service provision and resource allocation. However, early-age mortality calculation in small areas posits substantial challenges, particularly in less developed countries, where data sources are most frequently restricted to information from population censuses, and estimation methodology consists on William Brass' indirect methods.

Demographic literature is abundant on methodologies to derive estimates for small areas, most often applied to data from vital statistic records or analogous data sources, seldom available in statistically underdeveloped countries. Contributions in this field of estimation methodology are not discussed here. This paper addresses indirect estimation methods based on population census data. We refer to some features of the former methodologies only briefly, just to acknowledge this line of work, and proceed to focus on the main objective of this article.

The obvious data sources to calculate infant and child mortality are vital registration systems, particularly the numbers of live births and deaths classified by age. Unfortunately, in less developed countries, vital registration systems are incomplete or non-existent. Under constraint data environments childhood mortality is typically estimated using data from population and housing censuses with the so-called *indirect methods*.² Methods developed or adapted for application to data from vital registration systems, or analogous data sources like surveillance systems or some administrative records, most frequently consist of statistical modeling, sometimes incorporating certain a priori information (or reasonable assumptions about this) that would allow Bayesian inference. Stoner *et al.* (2019) provide an interesting framework for correcting underreporting of TB in Brazil, which can be very

¹ *Infant mortality* is the probability of dying between birth and exact age one (in life table notation is represented by $1q_0$) while *child mortality* corresponds to the probability of dying between exact ages 1 and 5 ($4q_1$). The subscript number to the right of "q" indicates the exact age at starting exposure and the subscript to the left indicates the length of exposure in exact years. The probability of dying between birth and exact age 5 ($5q_0$) is denoted as *under-five* or *early-age mortality*. Infant mortality and under-five or early-age mortality are usually expressed in terms of per 1,000 live births and child mortality is expressed as per 1,000 children alive at exact age of one year (see, for example, KINTNER, 2004).

² Another source of under-five mortality statistics are the demographic surveys. Yet, most often surveys are not suitable for small area estimation because of few -sometimes none- observations in small area sample size. Estimation methods that borrow information across related areas through implicit or explicit linking models and auxiliary information such as census data have been developed (see RAO, 2008). Although it is recognized that those methods are suitable and valid approaches this paper does not deal with them. The attention here is centered on indirect estimation methods based on census data.

useful to adapt for infant and child mortality. When sample surveys collect information on vital event histories, estimation methodologies applied to vital statistics can be usefully adapted for these analyses; Mercer *et al.* (2015) developed a very useful application to Tanzanian data, combining information from Demographic and Health Surveys and some surveillance systems; Swanson (2015) and Swanson *et al.* (2019) conducted estimations of infant mortality for small areas (counties in Estonia and counties in California) through statistical modeling approaches based on a Beta Binomial distribution. Readers interested in this line of work can find abundant bibliographical references in the work of these authors.

When data sources consist of population censuses, the most frequently used indirect methods are the Brass-type methods. This approach was pioneered by William Brass (1964) and improved by others (HILL, 2013; POPOFF; JUDSON, 2004; UNITED NATION, 1983). It requires information of women's age within five-year age groups, the number of live-born children they have ever had, and the number of those children who have died. The source of data for this method is usually a census. Two basic questions are posited to all women 15 years and older: *How many children have you had?*, followed by: *How many of those children are still alive?* The specific wording in the census questionnaires may vary in order to adapt to cultural contexts, with the aim to improve the quality of the responses.

The proportion of children dead by age of their mother is obviously a mortality indicator, although not a conventional life table one. Brass' method consists in transforming this proportion into a regular life table indicator. The key challenge in this procedure is to determine the length of time that children born alive by mother's age group have been exposed to the risk of dying. This time exposure varies according to the mother's age *cohort*. For each 5-year age group of mothers, the time exposure estimation is obtained on the basis of: the age pattern of fertility, which provides the proportion of children born alive by age of the mother at birth of the child (then the time exposure is deducted through reference to the mother's age at the time of the survey); and the age pattern of mortality, which indicates the distribution of children dead by length of time after their birth. Based on model age fertility patterns and model life tables, these time exposures and distribution of deaths are estimated, which allows for the transformation of the proportions of children dead into probabilities of dying from birth up to specific children's ages. Moreover, through these procedures, the estimated time location corresponding to the age specific probabilities of dying is obtained; this is done by establishing time series that provide the evolution of the mortality levels in a period of time preceding the survey or census. The results of the modeling described above are presented in the form of multipliers, tabulated according to fertility and mortality patterns, or through equations based on fertility indicators and model life tables, according to the estimation procedure adopted.

The Trussell or the Palloni and Heligman variants of the original Brass method provide a series of seven estimates of the probability of dying from birth to age x : ${}_xq_0$ ($x = 1, 2, 3, 5, 10, 15, 20$), one for each five-year age group of women in reproductive age (15-19, 20-24, ... 45-49 years). The Trussell version makes use of the Coale-Demeny model life table

system (UNITED NATION, 1983) and the Palloni-Heligman uses the United Nations model system (PALLONI; HELIGMAN, 1985).

To facilitate the analysis of mortality trends, these ${}_xq_0$ estimates have to be expressed in terms of infant mortality, ${}_1q_0$, or under-five mortality ${}_5q_0$, so that they constitute a time series of comparable indicators. Each estimate ${}_xq_0$ is dated somewhere between one and fifteen years before the census (UNITED NATION, 1983). By expressing ${}_xq_0$ in terms of ${}_5q_0$ or ${}_1q_0$, the estimates can be shown as a time series of consistent measurements stretching back about 15 years from the census date. The application of this method requires selecting an appropriate model life table, because mortality variations by age in the selected model should conform, as much as possible, to the age pattern in the studied population; otherwise there may be variations in infant or under-five mortality, when translating the obtained series of ${}_xq_0$ into a unified indicator (${}_1q_0$ or ${}_5q_0$).

These indirect estimation methods have been used for more than four decades. Results are usually accurate if attention is paid to the formal demographic structure of the data and to the logic of the estimation procedure (FEENEY, 1991). The method is based on a set of assumptions, which must be taken into account to avoid erroneous interpretations of the results.³

The method uses cumulative fertility of all women interviewed, so it is usually based on relative larger numbers of births than alternative methods. However, as the estimation areas become smaller, the number of dead children could be very small, especially in the youngest age groups of women. In such circumstances, estimates may be affected by relevant random errors. Furthermore, annual mortality fluctuations may affect small population areas; location-specific factors such as armed conflicts, epidemics or natural disasters are likely to affect infant mortality levels substantially in smaller areas. Random variations may also affect other factors, which influence the results of this method (sex ratios, fertility, and migration). Hence, deviations from the assumptions on which the method is based may become more relevant in results for small areas. Fluctuations and irregular fertility, mortality and migration behavior may bias the results of this method when applied to small area data. Random variations are more relevant when the number of children who died is smaller, as in the case of children born to women in the younger age groups.

As mentioned above, the method renders a series of seven estimates of mortality in the early years of life, which are usually expressed in terms of infant or under-five mortality rates, each dated somewhere from 1 to 15 years before the census. The most recent point

³ One important assumption of this method is that the estimates obtained are presented as time series. However, they are not strictly a time series. Each point is obtained from the data corresponding to an age cohort of women. Hence, the set of seven points are not the early-age mortality of the children born to a unique cohort. Each point estimate represents the mortality experience of children born to individual women cohorts, estimated at a given date location. The assumption underlying this procedure is that children's mortality risk is related only to their age. Other assumptions are that mortality decline has been linear; that the mortality pattern has remained constant throughout the period under consideration; and that the age pattern of fertility has also remained constant. All these assumptions are probably broken in most countries where the method is utilized; yet the technique is quite robust. Changes in mortality patterns and a moderate fertility decline are not very influential on the results. Gradually falling fertility will bias the method toward slight overestimation of mortality. This is not serious because it usually counteracts latent tendencies to under-estimation (SIAP, 1994).

estimate (from women aged 15-19 years) is highly erratic and is usually ignored. The second point, corresponding to women 20-24 years, or the third point, corresponding to women 25-29 years, are considered to represent the most reliable recent estimates, and are frequently presented as the main results of the estimation procedure; that is, they are usually adopted as the infant and child mortality rates estimated from the census data. However, since most women are starting their reproductive life between 20 and 29 years of age, in small areas these younger groups may contain a small number of observations. In addition, these observations may not represent the true early-age mortality experience of the area due to seasonal fluctuations, location-specific features, and eventual issues violating some of the assumptions of the method or just undetermined random factors.

The purpose of this paper is to propose an adaptation of the indirect estimation approach to obtain infant and child mortality estimates for small areas, which would mitigate those limitations. The method is very simple and does not require long or complex calculations. From now on, this method will be referred to as *sub-area mortality estimation method*. This approach does not make corrections for other problems that may affect small area statistics, such as undercounting of the population or events (births, deaths). Therefore, as in all demographic estimates, results should be treated with caution.

The rational of the Brass-type sub-area mortality estimation method

As stated above, the principle of Brass-type estimation methods relies on the fact that the proportion of children dead, D_i , among children born to women in age *cohort* i , at the time of the census, is equivalent to a probability of dying from birth to a certain age “ t ”, ${}_tq_0$, which is an average of the risks that affected all children born to women in that *cohort*. The method consists in transforming this age “ t ” implicit in the proportion of children dead, into exact ages x , so estimated risks of dying are expressed in terms of a conventional life table indicator. The idea can be extended to the risk of dying affecting children born to an expanded age *cohort* of women, such as women aged 20-34. In this case, the proportion of children dead, D_{20-34} , would represent an average of the risks of dying that affected the group of children born to the expanded mother’s age *cohort*. By using model age patterns of fertility and model age patterns of mortality, in the same way as in the original method, the implicit time exposure to the risk of dying “ t^* ”, in the proportion D_{20-34} could be estimated, thus translating D_{20-34} into a probability ${}_{t^*}q_0$, then expressing the mortality risk in terms of a conventional life table measurement.

However, for calculating the sub-area mortality estimation under discussion here, the process is simplified, with no need to estimate specific multipliers to adjust for age-fertility and mortality patterns. The central assumption is that the age patterns of mortality are similar in the sub-areas as in the larger, or “parent” area to which the sub-areas belong. Hence, there would be a factor of the scale, or level of mortality, which differentiates the risks of mortality between the mortality level observed in the parent area “ k ” and that of

any sub-area “j” within the area “k” (which will be represented by $\lambda_{j,k}$). That is: $\lambda_{j,k} = {}_xq_0^{j,k} / {}_xq_0^k$, where k indicates the specific area under study, and j indicates the specific sub-area within area k . Thus, $\lambda_{j,k}$ is the factor indicating the differential scale, or mortality level difference between the parent area and the sub-area.

Under these conditions, the ratio between the proportion of children dead to women in cohort 20-34 in the sub-area j (within area k), to the similar proportion in the parent area k , that is $D^{i,k}_{20-34} / D^k_{20-34}$, can be used to estimate the scale factor ($\lambda_{j,k}$ or differential mortality level) for sub-area j . Then, by applying this factor ($\lambda^{i,k}$) to the estimated probabilities of death ${}_xq_0^k$ that have been estimated for area k , we can obtain the corresponding probabilities (${}_xq_0^{j,k}$) for the sub-area j : ${}_xq_0^{j,k} = \lambda^{i,k} {}_xq_0^k$.

The central issue in this method would be to establish the factor of scale, $\lambda_{j,k}$, which links the level of mortality in the sub-area with the level observed in the parent area. This factor of scale can be obtained through the ratios $\lambda_{j,k} = D^{i,k}_{20-34} / D^k_{20-34}$. The proportions $D^{i,k}_{20-34}$ are based on relative large number of observations, as compared to the proportions D_i based on five-year age groups; therefore, they are less affected by random variations, which could introduce significant fluctuations in the proportions of children dead reported by mothers classified into five-year age cohorts. In this manner, it is possible to obtain reasonable estimates of mortality for smaller areas, avoiding the constraints emanating from small numbers of observations and significant random variations that may affect mortality estimates obtained from Brass' method, when applied in small areas.

The sub-area method is illustrated with data from the 2014 Population and Housing Census of Myanmar (MINISTRY OF IMMIGRATION AND POPULATION, 2015). The Myanmar 2014 Census has been extensively analysed with several thematic reports published, revealing good quality data.⁴ The mayor subdivisions of the country are *states* and *regions*, followed by *districts* and *townships*. States appear to have populations large enough to get reliable indirect estimates of infant and child mortality, but some of the districts have smaller populations, where traditional application of a Brass-type method would raise some concerns. Finally, most townships are too small for traditional application of indirect estimation methods.

Hence, a first exercise was done using *districts* data, with the main purpose of validating the sub-area method, that is, to examine whether it can provide reliable and valid results compared to original Brass-type methods. The mortality indicator used in these analyses is the IMR (${}_i q_0$). Most districts have populations large enough to estimate IMR by confidently applying a Brass-type method. In a following exercise, the sub-area method was used to estimate infant mortality among the townships of one region, Kachin. In this second application, only the proposed sub-area method was used, since townships are too small to reasonably apply a Brass-type indirect method.

⁴ Reference materials can be accessed at: http://www.dop.gov.mm/moip/index.php?route=product/category&path=54_52.

Application of the sub-area method to data from Myanmar districts

As stated above, the sub-area method aims to aggregate a larger number of observations, by grouping three mother's five-year age groups, thus reducing the random variation in the number of deaths, as well as other factors that may affect the results. Then, the method identifies the relationship between mortality levels in the smaller area, say district "j" and the level in the "parent area", the state/region "k", to which the smaller area belongs.

The factor $\lambda^{i,k} = D^{i,k}_{20-34} / D^k_{20-34}$ is calculated, which represents the relative risk in district j with respect to region/state k. The district's mortality rates (infant, child or under-five) are obtained by applying (multiplying) this factor $\lambda^{i,k}$ to the respective infant, child or under-five mortality rate of the state/region.

The basic principle of this method is that the expected level of mortality for each parent area as a whole, is the result of the specific levels of each of its sub-areas in such a way that each one of them would contribute with a portion to the mortality level of the parent area, having similar underlying age patterns. The step-by-step procedure to estimate mortality rates in the small areas is as follows:

1. The procedure starts by calculating the desired mortality indicators (infant, child and/or under-five mortality rates) for a given parent area (state/region), by applying the Brass-type indirect methods in the usual manner. In this application, the mortality indicator was IMR, and was determined as equivalent ${}_1q_0$ mortality level corresponding to the estimate derived from the proportion corresponding to age group 20-24, adjusted by the mortality trend for the time series, to correct selective higher mortality of children born to very young mothers.
2. The next step is obtaining the proportion of dead children among all children ever born to women aged 20-34 years, for the respective state/region, as well as for each of the districts within this state/region.
3. Later, the relative risk of each district to the parent state/region is estimated by dividing the proportions of children dead: $\lambda^{i,k} = D^{i,k}_{20-34} / D^k_{20-34}$.
4. Finally, the mortality rates for a given district j, situated in state/region k, say infant mortality rate ${}_1q_0^{j,k}$, is calculated by multiplying the infant mortality estimated for state/region k by the factor $\lambda_{j,k}$: $\lambda_{j,k} \cdot {}_1q_0^k$. $\lambda_{j,k} = {}_1q_0^{j,k}$.
5. Through the same procedure, the child mortality rate, ${}_4q_1$, and under-five mortality, ${}_5q_0$, can be obtained for district j in the region k: ${}_4q_1^k \cdot \lambda_{j,k} = {}_4q_1^{j,k}$; and ${}_5q_0^k \cdot \lambda_{j,k} = {}_5q_0^{j,k}$. In this exercise, only the IMR will be calculated.

By grouping three five-year age cohorts of women it is possible to work with a large number of dead children, diminishing the influence of random variations in the final estimates. As indicated above, the essential assumption is that the age patterns of fertility and mortality would be similar in the districts and in the state/region they belong to, while

the differential factors that determine the overall level of mortality in the region and in the districts would be captured through the relative risks, and reflected in the scale coefficient $\lambda_{j,k}$. It is important to reiterate that the estimate in the larger area's data is supposed to be accurate.

Results

Table 1 shows the results of applying the indirect estimation procedure in the traditional way for all districts, estimating the infant mortality rate (IMR) in the 74 districts in which the 15 states/regions of Myanmar are divided (see Map 1), as well as estimating IMRs for the districts through the proposed sub-area estimation approach. Some of the states/regions have only 2 districts, while others have as many as 14.

The IMRs were estimated using the Palloni-Heligman version of the Brass method; the software MORTPAK (Version 4.3) was used for the calculations (UNPD, 2013). The alternative rates were calculated according to the proposed sub-area approach. The table also includes the size of relevant populations. It is particularly important to observe the small number of dead children born to women 20-24. In some districts, the number is less than 100. As expected, when women 20-34 years are considered, the size increases substantially.

TABLE 1
Estimates of infant mortality rates (IMR), per 1,000 live births, in state/regions and districts

State/ Region	District	Population	Female reproductive age	Dead children from women 20-24	Dead children from women 20-34	Children ever born from women 20-34	IMR Province/ Region	IMR according to a Brass- type method	IMR according the sub-area method
Union		47,929,999	13,410,743	52,254	482,722	6,397,264	65.6		
Kachin		1,370,748	364,566	1,455	14,250	208,781	52.8		
	Myitkyina	478,482	128,659	527	4,914	72,037		54.8	52.7
	Mohntin	491,436	131,445	440	4,644	71,921		48.1	49.9
	Bhamo	312,769	83,453	347	3,316	49,481		51.5	51.8
	Putao	88,061	21,009	141	1,376	15,342		68.0	69.4
Kayah		272,730	72,267	367	3,273	46,870	62.2		
	Loikaw	235,159	62,442	313	2,742	39,708		61.8	59.4
	Bawlakhe	37,571	9,825	54	531	7,162		56.3	66.0
Kayin		1,454,264	356,505	1,448	13,977	202,122	53.6		
	Hpa-An	759,550	179,552	658	6,875	97,031		55.3	54.9
	Pharpon	32,719	8,308	56	430	5,680		56.0	58.7
	Myawady	201,246	53,850	179	1,734	32,686		35.9	41.1
	Kawkareik	460,749	114,795	555	4,938	66,725		59.7	57.4
Chin		469,109	113,577	1,020	8,304	92,920	75.5		
	Haka	96,468	23,567	88	550	16,921		29.8	27.5
	Falam	164,569	39,089	286	2,015	31,379		61.4	54.3
	Mindat	208,072	50,921	646	5,739	44,620		109.9	108.7

(continue)

(continued)

State/ Region	District	Population	Female reproductive age	Dead children from women 20-24	Dead children from women 20-34	Children ever born from women 20-34	IMR Province/ Region	IMR according to a Brass- type method	IMR according the sub-area method
Sagaing		5,076,326	1,444,947	5,003	50,680	666,205	60.0		
	Sagaing	487,473	143,439	218	2,484	47,040		42.4	41.6
	Shwebo	1,397,765	408,698	989	10,866	168,730		49.6	50.8
	Moniwa	723,580	218,849	482	4,591	76,180		56.1	47.5
	Katha	822,308	225,691	1,161	11,400	127,123		65.9	70.7
	Kalay	496,345	138,989	653	5,763	74,250		63.7	61.2
	Tamu	111,986	28,866	143	1,319	19,651		51.0	52.9
	Mawlaik	160,136	41,657	269	2,493	25,064		79.3	78.4
	Hkamti	348,512	85,841	611	6,471	65,560		74.3	77.8
	Yinmarpin	528,221	152,917	477	5,293	62,607		64.3	66.6
Tanintharyi		1,352,283	342,392	1,945	16,608	203,167	70.8		
	Dewei	470,015	116,614	337	3,175	53,012		51.9	51.9
	Myeik	672,898	172,708	1,177	10,022	112,393		77.9	77.2
	Kawthaung	209,370	53,070	431	3,411	37,762		73.3	78.2
Bago		4,743,808	1,312,133	4,606	45,127	597,471	61.9		
	Bago	1,723,854	472,546	1,681	16,839	238,511		56.3	57.9
	Taungoo	1,091,902	300,880	1,269	12,401	148,371		67.9	68.5
	Pyay	881,665	253,978	584	6,066	87,706		57.8	56.7
	Thayawady	1,046,387	284,729	1,072	9,821	122,883		68.3	65.5
Magway		3,786,538	1,090,638	4,714	49,243	468,025	83.9		
	Magway	1,199,577	343,826	1,295	14,294	143,480		78.9	79.5
	Minbu	660,868	190,840	865	8,727	83,768		82.7	83.1
	Thayet	700,691	203,485	786	8,374	85,547		77.8	78.1
	Pakokku	982,252	282,959	1,509	15,185	126,432		99.2	95.8
	Gangaw	243,150	69,528	259	2,663	28,798		65.3	73.8
Mandalay		5,843,424	1,730,326	4,070	43,820	670,656	50.3		
	Mandalay	1,577,437	490,233	829	7,587	171,774		38.7	34.0
	Pyin oo lwin	943,857	267,124	761	8,167	126,808		46.1	49.6
	Kyaukse	716,888	210,352	550	5,410	87,711		51.6	47.5
	Myingyan	1,033,495	303,267	618	7,745	106,803		55.0	55.8
	Myaung U	232,189	69,909	152	2,206	24,994		59.2	67.9
	Yame thin	491,578	145,237	546	5,442	60,776		75.3	68.9
	Meikhtila	847,980	244,204	614	7,263	91,790		59.6	60.9
Mon		1,949,821	511,442	1,096	10,555	216,214	41.8		
	Maulamyie	1,168,625	311,342	527	4,954	120,263		36.9	35.3
	Thaton	781,196	200,100	569	5,601	95,951		47.9	50.0

(continue)

(continued)

State/ Region	District	Population	Female reproductive age	Dead children from women 20-24	Dead children from women 20-34	Children ever born from women 20-34	IMR Province/ Region	IMR according to a Bras- type method	IMR according the sub-area method
Rakhine		2,034,148	559,208	2,622	23,863	317,880	61.1		
	Sittway	520,903	149,087	488	4,777	80,928		49.9	48.0
	Mayauk U	657,999	182,889	1,003	9,757	111,559		66.7	71.2
	Maung Taw	89,604	24,471	170	1,214	17,327		63.5	57.0
	Kyauk phyu	422,568	111,344	576	4,898	61,349		66.9	65.0
	Thandwe	343,074	91,417	385	3,217	46,717		56.0	56.0
Yangon		6,949,440	2,124,560	3,942	35,245	724,924	44.9		
	North	2,456,529	754,050	1,848	15,847	291,620		49.7	50.2
	East	2,258,040	695,360	833	7,116	200,538		35.8	32.8
	South	1,379,853	393,344	1,072	10,775	178,604		49.1	55.8
	West	855,018	281,806	189	1,507	54,162		34.9	25.7
Shan		5,500,933	1,465,146	8,471	67,636	948,188	55.5		
	Taunggyi	1,619,507	454,556	2,824	21,494	272,391		68.1	61.4
	Lailin	537,402	145,625	872	7,210	92,150		56.8	60.9
	Linkhe	130,875	34,497	332	1,982	20,859		83.4	74.0
	Lashio	583,342	158,044	673	6,425	94,512		46.2	52.9
	Muse	429,439	110,867	381	3,592	70,738		34.9	39.5
	Kyaukme	721,369	196,873	1,113	10,205	114,808		63.3	69.2
	Kunlon	55,355	11,697	59	480	10,701		33.0	34.9
	Lauling	143,425	33,871	91	597	26,770		21.6	17.4
	Hopan	225,742	54,203	147	1,236	43,710		21.5	22.0
	Makman	222,506	55,209	534	3,804	49,773		64.8	59.5
	Kengtaung	335,559	84,443	462	3,222	58,105		44.6	43.2
	Minesat	225,271	53,873	628	4,769	49,628		67.2	74.8
	Tachilate	167,493	46,103	186	1,415	25,891		42.1	42.6
	Minephyat	103,648	25,285	169	1,205	18,152		46.0	51.7
Ayeyawady		6,053,594	1,615,645	10,365	89,186	881,790	86.2		
	Pathein	1,587,709	425,258	2,149	18,690	222,095		73.1	71.7
	Phayapon	1,007,450	265,507	2,192	19,035	168,800		94.4	96.1
	Maubin	954,655	256,113	1,476	12,048	133,916		80.1	76.7
	Maugmya	767,340	203,241	1,214	10,880	110,697		87.6	83.7
	Labutta	617,284	163,082	1,929	16,183	111,685		108.8	123.5
	Hinthada	1,119,156	302,444	1,405	12,350	134,597		79.5	78.2
Nay Pyi Taw		1,072,833	307,391	1,130	10,955	152,051	55.4		
	Ottara	499,894	142,928	519	5,145	73,465		52.1	53.8
	Dekkhima	572,939	164,463	611	5,810	78,586		58.5	56.8

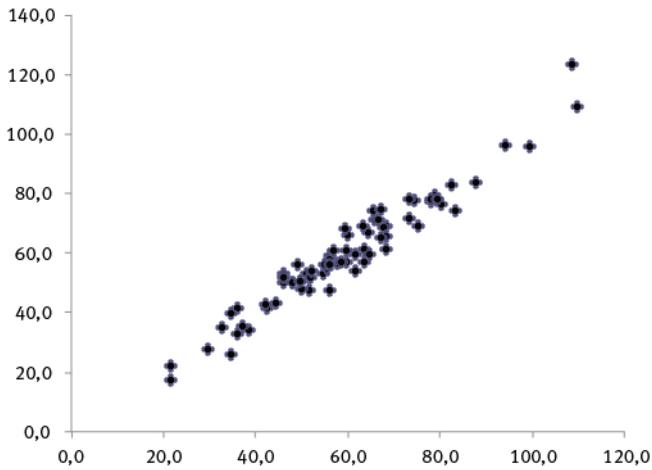
Source: Own estimates, based on 2014 Myanmar Census data.

Discussion

First, there are two main issues that are important to examine at this point: the differences in the IMR obtained by the two estimation approaches, and the differences between the district and the state/region according to each estimation method.

Table 2 shows the difference (absolute and percentage) between the IMR estimated with a Brass-type method and the sub-area method. Some of the differences are very small (Thandwe, Dewei) and others are relatively larger (Thandwe West, Labutta, Bawlakhe, Linkhe), but in no case is the difference very large. The correlation coefficient between the Brass-type and the sub-area method estimates is 0.97. Figure 1 shows the scatter graph corresponding to these two estimates; it indicates a clear linear trend. Evidently, the two estimates cannot be equal, since the database is different, but the values are close enough, indicating that both approaches are providing results for the same indicators, which are of a similar order of magnitude for the estimated variable (${}_1q_0$).

FIGURE 1
Scattergraph for IMR estimated by a Brass-type method and by the sub-area method



Source: Own estimates, values presented in the last 2 columns of Table 1.

TABLE 2
Differences between IMR estimated by a Brass-type method and by the sub-area method

District	IMR according to a Brass-type method	Difference between Brass-type and sub-area IMR	Per cent difference (1)	District	IMR according to a Brass-type method	Difference between Brass-type and sub-area IMR	Per cent difference (1)
Myitkyina	54.8	2.0	3.7	Myingyan	55.0	0.8	1.4
Mohntin	48.1	1.8	3.7	Myaung U	59.2	8.7	14.7
Bhamo	51.5	0.3	0.6	Yame thin	75.3	6.3	8.4
Putao	68.0	1.3	1.9	Meikhtila	59.6	1.3	2.1
Loikaw	61.8	2.3	3.8	Maulamyie	36.9	1.6	4.3
Bawlakhe	56.3	9.7	17.3	Thaton	47.9	2.2	4.6
Hpa-An	55.3	0.3	0.6	Sittway	49.9	1.9	3.7
Pharpon	56.0	2.7	4.9	Mayauk U	66.7	4.4	6.6
Myawady	35.9	5.2	14.6	Maung Taw	63.5	6.5	10.2
Kawkareik	59.7	2.4	4.0	Kyauk phyu	66.9	2.0	3.0
Haka	29.8	2.3	7.7	Thandwe	56.0	0.0	0.0
Falam	61.4	7.1	11.6	North	49.7	0.5	1.1
Mindat	109.9	1.2	1.1	East	35.8	3.0	8.3
Sagaing	42.4	0.8	1.8	South	49.1	6.7	13.6
Shwebo	49.6	1.1	2.3	West	34.9	9.2	26.3
Moniwa	56.1	8.6	15.4	Taunggyi	68.1	6.7	9.8
Katha	65.9	4.8	7.3	Lailin	56.8	4.1	7.2
Kalay	63.7	2.5	4.0	Linkhe	83.4	9.4	11.3
Tamu	51.0	1.9	3.8	Lashio	46.2	6.8	14.7
Mawlaik	79.3	0.9	1.1	Muse	34.9	4.6	13.2
Hkamti	74.3	3.5	4.8	Kyaukme	63.3	5.9	9.4
Yinmarpin	64.3	2.4	3.7	Kunlon	33.0	1.9	5.9
Dewei	51.9	0.1	0.1	Lauling	21.6	4.2	19.6
Myeik	77.9	0.7	0.9	Hopan	21.5	0.5	2.5
Kawthaung	73.3	4.9	6.7	Makman	64.8	5.3	8.2
Bago	56.3	1.5	2.7	Kengtaung	44.6	1.4	3.2
Taungoo	67.9	0.7	1.0	Minesat	67.2	7.6	11.3
Pyay	57.8	1.1	1.9	Tachilate	42.1	0.5	1.1
Thayawady	68.3	2.7	4.0	Minephyat	46.0	5.7	12.3
Magway	78.9	0.6	0.8	Pathein	73.1	1.4	1.9
Minbu	82.7	0.4	0.5	Phayapon	94.4	1.6	1.7
Thayet	77.8	0.3	0.4	Maubin	80.1	3.5	4.4
Pakokku	99.2	3.4	3.4	Maugmya	87.6	3.8	4.4
Gangaw	65.3	8.5	13.0	Labutta	108.8	14.7	13.5
Mandalay	38.7	4.7	12.1	Hinthada	79.5	1.4	1.7
Pyin oo lwin	46.1	3.5	7.6	Ottara	52.1	1.7	3.3
Kyaukse	51.6	4.1	8.0	Dekkhima	58.5	1.7	2.8

Source: Own estimates based on 2014 Myanmar Census data.

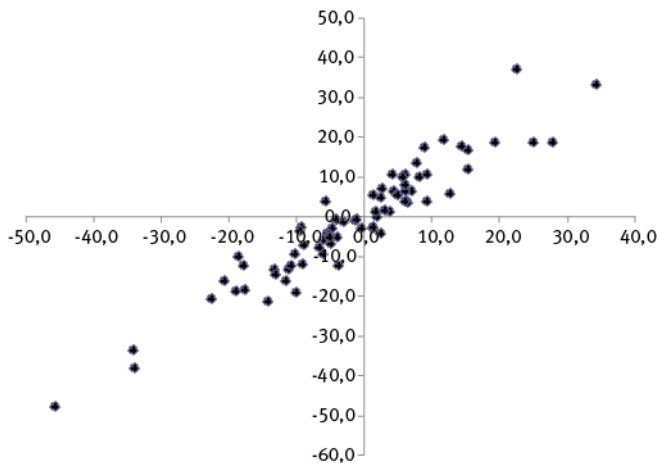
(1) Among 74 districts, the Brass' type and "sub-area" method estimates values differ by more than 10% in 17 cases (23%).

The second issue worthy of examination is the difference between the states/regions and their districts according to the Brass-type and the sub-area methods. The results of this exercise are in Table 3. Differences between IMR in the state/region and the districts are expected considering the size of the former and the diversity of development, environment and demography in the latter. The important point here is that the levels of mortality calculated with the two methods show very close results.

According to Table 3, in general, these differences are in fact very small. When there are large differences between the IMR in the district and in the state/region, as estimated with the Brass-type method, the difference is also large and in the same direction when the sub-area method is used for the calculation. Small differences between region/states and districts obtained with one method also correspond to small differences obtained with the other method.

The scatter graph presented in Figure 2 shows that the relationship between the differences obtained with the two methods is linear. The respective correlation coefficient is 0.95. This analysis also indicates that both methods measure indicators of the same variable, and reveal similar orders of magnitude. Obviously, the methods cannot record exactly the same differences, but these are similar enough to consider that the alternative method for estimation in small areas gives a result equivalent to the indirect Brass-type method.

FIGURE 2
Scattergraph of the absolute difference between state/region and district estimates according to a Brass-type and the sub-area methods



Source: Own estimates, values presented in Table 3.

TABLE 3
Absolute difference between state/region and district estimates according to a Brass-type and the sub-area methods

District	Absolute difference between state/region and district		District	Absolute difference between state/region and district	
	IMR Brass-type	IMR sub-area		IMR Brass-type	IMR sub-area
Myitkyina	2.0	0.0	Myingyan	4.8	5.5
Mohntin	-4.6	-2.8	Myaung U	8.9	17.6
Bhamo	-1.3	-1.0	Yame thin	25.0	18.6
Putao	15.3	16.6	Meik htilla	9.3	10.6
Loikaw	-0.4	-2.8	Maulamyie	-4.9	-6.5
Bawlakhe	-5.9	3.8	Thaton	6.0	8.2
Hpa-An	1.7	1.3	Sittway	-11.2	-13.1
Pharpon	2.4	5.1	Mayauk U	5.7	10.1
Myawady	-17.7	-12.5	Maung Taw	2.4	-4.1
Kawkareik	6.1	3.8	Kyauk phyu	5.9	3.9
Haka	-45.8	-48.1	Thandwe	-5.0	-5.1
Falam	-14.1	-21.3	North	4.8	5.3
Mindat	34.4	33.2	East	-9.2	-12.1
Sagaing	-17.6	-18.3	South	4.1	10.8
Shwebo	-10.3	-9.2	West	-10.1	-19.2
Moniwa	-3.8	-12.5	Taunggyi	12.6	5.9
Katha	5.9	10.7	Lailin	1.3	5.4
Kalay	3.7	1.2	Linkhe	27.9	18.4
Tamu	-9.0	-7.1	Lashio	-9.4	-2.6
Mawlaik	19.3	18.4	Muse	-20.6	-16.0
Hkamti	14.3	17.8	Kyaukme	7.7	13.7
Yinmarpin	4.3	6.7	Kunlon	-22.6	-20.6
Dewei	-18.9	-18.9	Lauling	-33.9	-38.2
Myeik	7.2	6.4	Hopan	-34.1	-33.5
Kawthaung	2.5	7.4	Makman	9.3	4.0
Bago	-5.6	-4.0	Kengtaung	-10.9	-12.4
Taungoo	5.9	6.6	Minesat	11.7	19.3
Pyay	-4.1	-5.2	Tachilate	-13.4	-13.0
Thayawady	6.3	3.6	Minephyat	-9.5	-3.9
Magway	-5.1	-4.5	Pathein	-13.1	-14.5
Minbu	-1.2	-0.8	Phayapon	8.3	9.9
Thayet	-6.1	-5.8	Maubin	-6.0	-9.5
Pakokku	15.3	11.9	Maugmya	1.4	-2.4
Gangaw	-18.6	-10.2	Labutta	22.6	37.3
Mandalay	-11.6	-16.3	Hinthada	-6.6	-8.0
Pyin oo lwin	-4.2	-0.7	Ottara	-3.3	-1.5
Kyaukse	1.3	-2.8	Dekkhima	3.1	1.4

Source: Own estimates based on 2014 Myanmar Census data.

As mentioned above, the method proposed here shares all the assumptions and limitations of the Brass-type methods. However, there are some additional issues that need to be considered.

The proportion of children who died among children born to women between 20-34 years old and the corresponding proportion for children born to women between 20-24 years old are both indicators of the level of mortality. However, these proportions differ in two aspects. First, the period of exposure to mortality is longer among children born to women between 20-34 years old than among children born to women between 20-24 years old. Second, in most countries, a process of mortality decline is likely to have occurred and, as a result, the level of mortality that has affected the children born to women aged 20-24 would be lower than the level that has affected the children born to women aged 20-34, since the former were born during a more recent period in a context of declining mortality. However, in practical terms, this does not necessarily represent a limitation for this estimation method. The small area estimation would be good, if the assumptions of the method are verified: that is, if the age pattern of fertility and age pattern of mortality are similar in the district and in the state/region (or if they do not differ radically), and that mortality trends have followed similar paths in the parent area and in the sub-areas. Under these conditions, estimates obtained by applying the scale factor $\lambda_{j,k}$, would be accurate, because $\lambda_{j,k}$ incorporates similar trends for parent area and sub-area. The factor $\lambda_{j,k}$ adjusts for differentials in the mortality level, incorporating at the same time the effects of mortality trends in both the parent area and in the sub-areas.

Obviously, under a mortality declining trend, the level of mortality implicit in the proportion of children dead to women aged 20-34 in the district j , $D_{20-34}^{j,k}$, as well as in state/region k , D_{20-34}^k , represent a higher mortality level and are located further in the past than the proportions of children dead to women 20-24 in similar areas: $D_{20-24}^{j,k}$ and D_{20-24}^k . When the scale factor $\lambda_{j,k}$ is used to derive the mortality indicator for a district from a similar indicator (either ${}_1q_0$ or ${}_5q_0$) in the state/region, the declining mortality trends are incorporated into the procedure. This is so because mortality trends are assumed to be similar in both geographic units and the scale factor $\lambda_{j,k}$ is calculated for similar age groups. Since the underlying age fertility and mortality patterns are similar, the estimates once adjusted by the scale factor (${}_1q_0^k \cdot \lambda_{j,k} = {}_1q_0^{j,k}$), would provide the correct level for the sub area. At the same time, the time location of this estimate would be also similar to parent area, because the areas have similar age fertility and mortality patterns, which are implicit in the calculation of the ${}_1q_0^k$, or ${}_5q_0^k$ through the traditional Brass' procedure applied to the state/region level. Hence, if the assumptions underlying the Brass estimation procedure are correct, it is legitimate to assign a similar time location to the mortality rate for the district, ${}_1q_0^{j,k}$, to the one estimated for the state/region, ${}_1q_0^k$.

Regarding whether the mortality decline might have been faster in one geographic unit than in the other, a way to assess this assumption would be to compare the ratio $w_{j,k} = D_{20-34}^{j,k} / D_{20-24}^{j,k}$ to the ratio $w_k = D_{20-34}^k / D_{20-24}^k$. A higher ratio in the district ($w_{j,k}$)

compared to the region (w_k) would indicate that mortality risks were relatively higher in the past in this district than in the state/region (faster mortality decline in the district than in the region). In this case, since the sub-area adjustment factor incorporates a larger component of past mortality than an estimation based only on age group 20-24, the result from the sub-area method would overestimate the recent infant mortality level in this district. If the ratio in the district ($w_{j,k}$) is lower than the ratio in the region (w_k), the method would underestimate the recent infant mortality level in this district. Table 4 shows the respective information.

According to Table 4, the ratio between these proportions observed in the districts is similar to that in the state/region to which they belong. There are some deviations, as expected, but these are relatively small, with no cases registering a substantial order of differences. Hence, the evidence from these results suggests that the assumption of similar mortality declining trends generally holds.

In addition to the assumption discussed above, systematic biases are also possible, in particular, those resulting from children born from very young women (15-19 years old). The level of mortality of children from women too young is usually higher than the average corresponding to all children ever born. If the proportion of children ever born from women aged 15-19 was substantially larger in a given district than in the state/region, then, under similar environmental risks IMR would be expected to be higher in this district due to distribution factors introducing a selectivity bias, through a larger proportion of children in a category affected by relative higher risks. A relative higher average mortality would have been derived, because under similar risks in similar categories, a larger proportion of children in a group that has relative higher risks would cause a higher number of deceased children, even when overall risks are similar in the corresponding categories in the two areas. Infant mortality associated to mother age group 20-24 is not necessarily higher than overall infant mortality. However, since the proportion of children dead to women aged 20-24 is a *cohort* measure, this incorporates the experience of children born when these women were 15-19. In this way, when the proportion of children born to very young women is high, their higher mortality reflects also in the proportion of children dead to women 20-24. The possible incidence of this potential selectivity factor affecting the comparison of mortality estimates based on mothers aged 20-34 to mothers aged 20-24, is explored in Table 5, assessing such possible distortion.

TABLE 4
Proportion of dead children among ever born to women 20-34 year old and women 20-24 year old and ratio between the proportions

State/Region	District	Proportion 20-34	Proportion 20-24	Ratio	State/Region	District	Proportion 20-34	Proportion 20-24	Ratio
Union		0.0755	0.0606	1.2	Mandalay		0.0653	0.0492	1.3
Kachin		0.0683	0.0511	1.3		Mandalay	0.0442	0.0371	1.2
	Myitkyina	0.0682	0.0530	1.3		Pyin oo lwin	0.0644	0.0445	1.4
	Mohntin	0.0646	0.0463	1.4		Kyaukse	0.0617	0.0500	1.2
	Bhamo	0.0670	0.0502	1.3		Myingyan	0.0725	0.0530	1.4
	Putao	0.0897	0.0667	1.3		Myaung U	0.0883	0.0576	1.5
Kayah		0.0698	0.0584	1.2		Yame thin	0.0895	0.0739	1.2
	Loikaw	0.0691	0.0599	1.2		Meik htla	0.0791	0.0578	1.4
	Bawlakehe	0.0741	0.0507	1.5	Mon		0.0488	0.0403	1.2
Kayin		0.0692	0.0523	1.3		Maulamyie	0.0412	0.0354	1.2
	Hpa-An	0.0709	0.0537	1.3		Thaton	0.0584	0.0461	1.3
	Pharpon	0.0757	0.0644	1.2	Rakhine		0.0751	0.0597	1.3
	Myawady	0.0531	0.0352	1.5		Sittway	0.0590	0.0477	1.2
	Kawkareik	0.0740	0.0586	1.3		Mayauk U	0.0875	0.0653	1.3
Chin		0.0894	0.0744	1.2		Maung Taw	0.0701	0.0626	1.1
	Haka	0.0325	0.0288	1.1		Kyauk phyu	0.0798	0.0665	1.2
	Falam	0.0642	0.0598	1.1		Thandwe	0.0689	0.0557	1.2
	Mindat	0.1286	0.1100	1.2	Yangon		0.0486	0.0432	1.1
Sagaing		0.0761	0.0586	1.3		North	0.0543	0.0480	1.1
	Sagaing	0.0528	0.0406	1.3		East	0.0355	0.0342	1.0
	Shwebo	0.0644	0.0480	1.3		South	0.0603	0.0475	1.3
	Moniwa	0.0603	0.0544	1.1		West	0.0278	0.0330	0.8
	Katha	0.0897	0.0649	1.4	Shan		0.0713	0.0549	1.3
	Kalay	0.0776	0.0628	1.2		Taunggyi	0.0789	0.0675	1.2
	Tamu	0.0671	0.0499	1.3		Lailin	0.0782	0.0569	1.4
	Mawlaik	0.0995	0.0790	1.3		Linkhe	0.0950	0.0855	1.1
	Hkamti	0.0987	0.0730	1.4		Lashio	0.0680	0.0452	1.5
	Yinmarpin	0.0845	0.0630	1.3		Muse	0.0508	0.0338	1.5
Tanintharyi		0.0817	0.0696	1.2		Kyaukme	0.0889	0.0627	1.4
	Dewei	0.0599	0.0504	1.2		Kunlon	0.0449	0.0322	1.4
	Myeik	0.0892	0.0766	1.2		Lauling	0.0223	0.0207	1.1
	Kawthaung	0.0903	0.0731	1.2		Hopan	0.0283	0.0205	1.4
Bago		0.0755	0.0604	1.2		Makman	0.0764	0.0642	1.2
	Bago	0.0706	0.0545	1.3		Kengtaung	0.0555	0.0443	1.3
	Taungoo	0.0836	0.0664	1.3		Minesat	0.0961	0.0689	1.4
	Pyay	0.0692	0.0565	1.2		Tachilate	0.0547	0.0422	1.3
	Thayawady	0.0799	0.0672	1.2		Minephyat	0.0664	0.0466	1.4
Magway		0.1052	0.0834	1.3	Ayeyawady		0.1011	0.0865	1.2
	Magway	0.0996	0.0778	1.3		Pathein	0.0842	0.0725	1.2
	Minbu	0.1042	0.0821	1.3		Phayapon	0.1128	0.0954	1.2
	Thayet	0.0979	0.0771	1.3		Maubin	0.0900	0.0800	1.1
	Pakokku	0.1201	0.0995	1.2		Maugmya	0.0983	0.0874	1.1
	Gangaw	0.0925	0.0653	1.4		Labutta	0.1449	0.1120	1.3
						Hinthada	0.0918	0.0795	1.2
					Nay Pyi Taw		0.0720	0.0542	1.3
						Ottara	0.0700	0.0509	1.4
						Dekkhima	0.0739	0.0573	1.3

Source: Own estimates based on 2014 Myanmar Census data.

TABLE 5
Proportion of children ever born from women 15-24 years

State/ Region	District	Proportion CEB 15-24	State/ Region	District	Proportion CEB 15-24	State/ Region	District	Proportion CEB 15-24
Union		0.0465	Bago		0.0418	Yangon		0.0417
Kachin		0.0496		Bago	0.0422		North	0.0460
	Myitkyina	0.0504		Taungoo	0.0423		East	0.0388
	Mohntin	0.0462		Pyay	0.0377		South	0.0425
	Bhamo	0.0544		Thayawady	0.0437		West	0.0304
	Putao	0.0482	Magway		0.0378	Shan		0.0666
Kayah		0.0490		Magway	0.0348		Taunggyi	0.0594
	Loikaw	0.0472		Minbu	0.0400		Lailin	0.0671
	Bawlakhe	0.0602		Thayet	0.0399		Linkhe	0.0711
Kayin		0.0445		Pakokku	0.0370		Lashio	0.0630
	Hpa-An	0.0384		Gangaw	0.0447		Muse	0.0634
	Pharpon	0.0546	Mandalay		0.0391		Kyaukme	0.0579
	Myawady	0.0605		Mandalay	0.0440		Kunlon	0.0738
	Kawkareik	0.0466		Pyin oo lwin	0.0466		Lauling	0.0713
Chin		0.0553		Kyaukse	0.0414		Hopan	0.0764
	Haka	0.0656		Myingyan	0.0305		Makman	0.0837
	Falam	0.0567		Myaung U	0.0306		Kengtaung	0.0795
	Mindat	0.0502		Yame thin	0.0383		Minesat	0.0897
Sagaing		0.0423		Meik htla	0.0334		Tachilate	0.0735
	Sagaing	0.0333	Mon		0.0376		Minephyat	0.0894
	Shwebo	0.0383		Maulamyie	0.0368	Ayeyawady		0.0494
	Moniwa	0.0346		Thaton	0.0387		Pathein	0.0486
	Katha	0.0507	Rakhine		0.0491		Phayapon	0.0517
	Kalay	0.0493		Sittway	0.0425		Maubin	0.0484
	Tamu	0.0551		Mayauk U	0.0490		Maugmya	0.0437
	Mawlaik	0.0471		Maung Taw	0.0631		Labutta	0.0624
	Hkamti	0.0490		Kyauk phyu	0.0498		Hinthada	0.0446
	Yinmarpin	0.0366		Thandwe	0.0558	Nay Pyi Taw		0.0494
Tanintharyi		0.0476					Ottara	0.0512
	Dewei	0.0380					Dekkhima	0.0479
	Myeik	0.0488						
	Kawthaung	0.0610						

Source: Own estimates based on 2014 Myanmar Census data.

Table 5 shows the proportion of children ever born to women aged 15 to 24. Only a small proportion of women declared to have had children between ages 15 to 24 in Myanmar: a mere 4.7%. In all state/regions and districts, the percentages are below 10%; the highest around 8% (Makman, Minesat, and Minephyat). In addition, differences in the proportions between the state/region and its districts are quite small. Therefore, in the districts, a distortion in the level of infant mortality caused by a large proportion of children ever born to young women (15-24 years old) does not occur. The differences in IMR (or U-5MR) among districts or between the state/region and its districts would not be related to eventual distortions of this kind.

As a conclusion, it is possible to state that the method proposed here to estimate IMR in small areas provides reliable results or, at least, as reliable as a Brass-type method. The basic indicator of IMR in the smaller and larger sub-divisions (proportion of dead children from women aged 20-34) appears to be valid and reliable. The sub area estimation method provides suitable estimates of the mortality levels in the early stages of life, which are comparable to a traditional Brass-type estimate. The validation tests applied to these results, which involved internal consistency checks and comparisons with the traditional Brass estimation technique, are considered appropriate and sufficient. Further validations would require comparisons with vital statistic data, which, as a rule, are not available, and in any case they would be equally (or more severely) affected by random variations in the small numbers. In addition, countries with good vital statistics do not include retrospective questions on mortality in their censuses, so no countries would allow for validations by comparing with vital statistics. Exceptionally, Brazil included those questions in the census. Yet, Brazil has undergone radical fertility changes, having registered below replacement fertility for some years, as well as significant declines in adolescent fertility more recently. This introduces biases in the estimates due to distribution effects, affecting the comparisons of data from different age *cohorts*. Under the conditions prevailing in Myanmar, as well as in most other countries where these methods could be applied, the validation procedures and consistency checks, as applied here, are considered appropriate and relevant.

To sum up, under the assumption of similar age mortality patterns in the parent area and sub-areas, the sub-area estimation procedure can be utilized to derive estimates of infant, child and under-five mortality rates, by using the scale factor $\lambda_{j,k}$, adjusting the mortality level in the sub-area to the respective rate estimated in the parent area: ${}_nq_x^k \cdot \lambda_{j,k} = {}_nq_x^{j,k}$. The selected mortality indicator, ${}_1q_0$, ${}_5q_0$, or ${}_4q_1$ would be derived from the corresponding indicator estimated for the parent area – based on a large number of observations – by adjusting the rate for the scale factor $\lambda_{j,k}$.

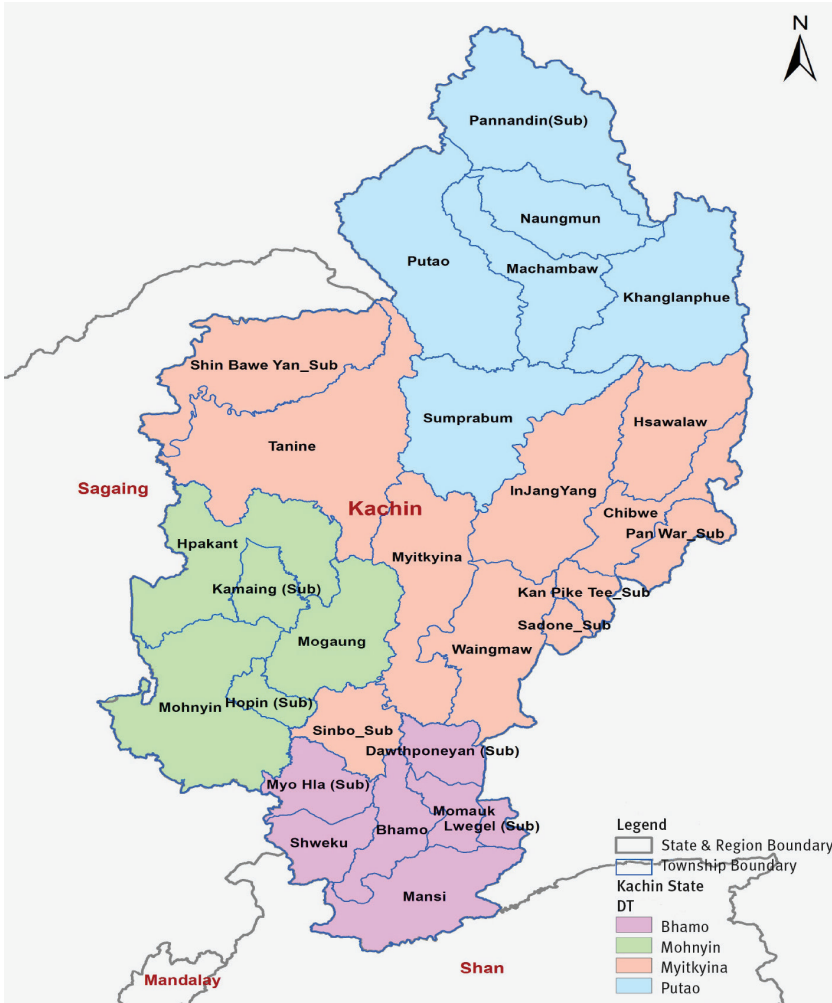
In the following section, the method is applied to sub-areas much smaller than the district: townships. Similar and additional validation tests are applied to the method's results.

Estimates of infant mortality at the township level

The administrative sub-division that follows the district in Myanmar is the *township*. There are 412 townships in the country. Some districts are sub-divided into just 2 townships while others contain more than 10.

The Kachin State was selected for this exercise. It is the most northern state of Myanmar. Shan State to the South, Sagaing Region and India to the West, and China to the North and East border the Kachin State. The state capital is Myitkina. Map 2 shows the division of the state by districts (4) and townships (29).

MAP 2
Kachin state, districts and townships



Several reasons would justify selecting Kachin State, but three stand out. First, there are many townships and most of their populations are small, which is particularly suitable for testing this method. Second, it is a remote and conflict-ridden area, with unstable and volatile situations, where problems of enumeration occur.⁵ This situation allows for testing the method under more challenging data quality conditions. Third, due to the above, the

⁵ In this state, the Kachin Independence Organisation (KIO), prevented the enumeration in some villages (in spite of a previous agreement between the government and KIO leaders). About 97 villages were not enumerated, being Mansi, Lwegel and Momauk the most affected townships. In these areas, total population collected during mapping was compared with the counted population in the townships where enumeration took place. An adjustment factor was calculated as the ratio between the counted population and the population estimated through mapping. This factor was used to estimate the population not covered during enumeration. It was estimated that 46,600 people were not enumerated in Kachin State. This total population was added to the overall population. Other indicators measured in the census were not estimated for these areas. This limitation has to be considered in the analyses.

state is not homogeneous. Some areas are under the control of the KIO separatists, some are affected by the conflict, while others have received a large number of internally displaced persons (in both government and KIO controlled areas). Townships in the Chinese border have more economic and social interaction with the Chinese province of Yunnan than with the central or local Myanmar government. There are also substantial socio-economic variations as well as variations in access to basic services (MINISTRY OF IMMIGRATION AND POPULATION, 2015).

Table 6 shows IMR and related basic information. Some townships have very small populations; therefore, very few births and deaths have occurred. Children dead among the ever born to women between 20-24 years old were included in Table 6 because these data are the basis to calculate IMR with a Brass-type method. Some small numbers suggest that Brass' method may not be suitable: in 7 townships the number of children dead among the ever born to women aged 20-24 is extremely small, less than 10 cases, and in 17 townships, it is less than 50 cases. For this reason, a Brass-type method was not used to estimate IMR. The sub-area method proposed here, based on the number of children deaths among those ever born to women between 20-39 years old seems more suitable. It should be noted that in this instance, to further increase the number of cases, four five-year age intervals have been used: 20-24, 25-29, 30-34 and 35-39. In the previous calculation of infant mortality at the district level, only three age intervals were used. It is also important to mention that factor $\lambda_{j,k}$ was calculated on the basis of the relationship between the township and the state (Kachin) and not with regard to the districts. Table 6 shows the results from this exercise.

According to Table 6, there are important variations in IMR among townships. These differences are expected since the state is not homogeneous, as mentioned above. There are two cases, however, that appear to be out of line, with very low IMR: Lwe'ge' (13.6) and Khaunglanphoo (22.6). Since it was not advisable to estimate infant mortality with a Brass-type method due to the small number of cases, it is not possible to compare results from the sub-area method and Brass' results. However, it is clear that in those two townships IMRs are too low, given the levels prevailing in the state as well as in most districts. A possible cause of this is the very small number of cases, combined with erratic changes in security and living conditions caused by the conflictive situation affecting parts of the state. These two townships have the smallest number of dead children: Lwe'ge' has no registered deaths among children born to women aged 20-24, Khaunglanphoo has only one death reported. Not even the deaths among children born to women aged 20-39 represent a number large enough to provide a reliable IMR estimate. It should also be noted that Lwe'ge' is among those townships where the population could not be wholly enumerated because of the separatist conflict, all this pointing to some serious data problems.

TABLE 6
Estimates of infant mortality rates (IMR) in Kachin state, districts and townships

Districts and townships	Population	Women in reproductive ages	Dead children from women 20-24	Dead children from women 20-39	Children ever born from women 20-24	Children ever born from women 20-39	IMR	Proportion children dead (group 20-39)
KACHIN	1,642,841	364,566	1455	27,165	28,473	342,416	52.8	0.079
Myitkyina							52.7	
Myitkyina	306,949	78,409	201	3,677	4,519	59,648	41.0	0.062
Waingmaw	106,366	24,637	125	2,518	2,319	27,063	61.9	0.093
Ingyanyan	1,732	357	2	85	23	536	105.5	0.159
Tanaing	48,566	9,658	55	966	997	10,623	60.5	0.091
Chiphwe	11,303	2,389	23	333	369	3,450	64.2	0.097
Hsotlaw	6,518	1,370	17	243	224	2,109	76.7	0.115
Hsinbo	10,655	2,417	26	410	240	2,859	95.4	0.143
Hsadone	10,496	1,458	20	200	239	2,739	48.6	0.073
Kanpaikti	8,682	2,139	18	173	297	2,365	48.7	0.073
Shinbwayyan	11,453	2,785	18	333	339	3,222	68.8	0.103
Panwa	8,736	1,920	22	212	381	2,753	51.3	0.077
Mohnyin							49.9	
Mohnyin	160,598	40,397	123	2,733	2,687	36,230	50.2	0.075
Mogaung	132,608	33,981	109	2,182	2,622	30,622	47.4	0.071
Phakant	312,278	39,915	149	3,305	3,218	39,054	56.3	0.085
Hopin	48,694	13,148	39	580	610	10,043	38.4	0.058
Kamine	19,430	4,004	20	512	362	4,821	70.7	0.106
Bhamo							51.8	
Bhamo	135,877	32,501	101	2,055	2,377	28,154	48.6	0.073
Shwegu	90,691	21,729	116	1,601	1,854	21,352	49.9	0.075
Momauk	41,562	9,948	38	856	850	10,042	56.7	0.085
Mansi	52,945	13,464	65	1,117	1,241	13,935	53.3	0.080
Myohla	4,093	975	6	150	112	1,188	84.0	0.126
Lwe'ge'	10,039	2,205	0	35	115	1,711	13.6	0.020
Dotphoneyan	11,313	2,631	21	359	363	3,062	78.0	0.117
Putao							69.4	
Putao	61,075	14,078	117	1,947	1,599	17,328	74.8	0.112
Sumprabum	2,546	543	2	71	48	714	66.2	0.099
Machanbaw	8,858	1,817	11	233	229	2,495	62.2	0.093
Khaunglanphoo	11,655	3,023	1	78	89	2,297	22.6	0.034
Naungmoon	5,365	1,194	5	130	102	1,420	60.9	0.092
Pannandin	1,758	354	5	71	48	581	81.3	0.122

Source: Own estimates based on 2014 Myanmar Census data.

Although this method was designed to solve the problem of a small number of cases, there are instances with an extremely small number of infant deaths, too small to provide a valid IMR, irrespective of the method used to estimate it. Similarly, there are situations beyond the possibilities of the method to provide a reliable estimate of IMR, such as when an underenumeration occurs, especially if it is associated to some kind of selectivity.

In those instances where the number of cases is too small, a possible solution is to combine the township in question with a neighboring one. In this particular exercise, this rule can be applied to all those townships with less than 100 deaths among children born to 20-39 year old women. Before implementing these calculations, it is convenient to repeat the consistency analyses previously done with districts.

As explained before, the sub-area method assumes that the ratio of the proportion of dead children among those born to women aged between 20-39 years old (or 20-34) to similar proportion from women in *cohort* 20-24 in the parent area is equal to the corresponding ratio in the small area. In countries experiencing a process of mortality decline, the level of mortality of children born to women aged 20-24 is usually lower than the level corresponding to children born to women 20-39 (or 20-34), because children born to younger women were born more recently, under lower mortality risks. If the declining trend is similar in the parent and in the small area, these ratios would have similar values. When applying the sub-area method, it is assumed that mortality trends are similar in the small area (townships) and in the parent area (state/region). The ratio between these two proportions is an indicator of the underlying trends, if they are similar in the two areas the assumption is satisfied.

The problem in this case is that the number of deceased children among children born to women aged 20-24 is extremely small in some townships. When the number of cases is too small, the evaluation is not very reliable, as random factors may interfere in the results. In addition, as indicated before, location specific factors have a greater impact on population changes in small areas (natural disasters, conflicts, closure of services, and the like). This analysis was conducted in spite of these limitations. The results are presented in Table 7.

TABLE 7
Proportion of dead children among those born alive to women 20-39 years old and 20-24 years old and ratio between these proportions

Townships	Proportion 20-39	Proportion 20-24	Ratio	Townships	Proportion 20-39	Proportion 20-24	Ratio
Kachin State	0.0793	0.0511	1.6	Kamine	0.1062	0.0552	1.9
Myitkyina	0.0616	0.0445	1.4	Bhamo	0.0730	0.0425	1.7
Waingmaw	0.0930	0.0539	1.7	Shwegu	0.0750	0.0626	1.2
Ingyanyan	0.1586	0.0870	1.8	Momauk	0.0852	0.0447	1.9
Tanaing	0.0909	0.0552	1.6	Mansi	0.0802	0.0524	1.5
Chiphwe	0.0965	0.0623	1.5	Myohla	0.1263	0.0536	2.4
Hsotlaw	0.1152	0.0759	1.5	Lwe'ge'	0.0205	0.0000	-
Hsinbo	0.1434	0.1083	1.3	Dotphoneyan	0.1172	0.0579	2.0
Hsadone	0.0730	0.0837	0.9	Putao	0.1124	0.0732	1.5
Kanpaikti	0.0732	0.0606	1.2	Sumprabum	0.0994	0.0417	2.4
Shinbwayyan	0.1034	0.0531	1.9	Machanbaw	0.0934	0.0480	1.9
Panwa	0.0770	0.0577	1.3	Khaunglanphoo	0.0340	0.0112	3.0
Mohnyin	0.0754	0.0458	1.6	Naungmoon	0.0915	0.0490	1.9
Mogaung	0.0713	0.0416	1.7	Pannandin	0.1222	0.1042	1.2
Phakant	0.0846	0.0463	1.8				
Hopin	0.0578	0.0639	0.9				

Source: Own estimates on the basis of 2014 Myanmar Census data.

Table 7 shows that the assumption is satisfied in most of the townships. The ratio in the townships is similar to the one observed in Kachin (1.6). There are, however, four cases where differences are large: Myohla (2.4), Dotphoneyan (2.0), Sumprabum (2.4), and Khaunglanphoo (3.0). It is also necessary to mention Lwe'ge', where the ratio cannot be calculated because there are no dead children among those born to women aged 20-24 years old. It should be noted, however, that there are two townships with a very small number of children dead among children born to women aged 20-24, still satisfying the assumption: Pannandin (1.2) and Ingyanyan (1.8). The analysis gives somehow mixed results in a few cases, mostly because the number of observations is exceptionally small. In general, results suggest that in most townships the assumption is satisfied.

It is convenient to replicate the second analysis conducted previously, when estimating the IMR for the districts. As stated earlier, if the proportion of children ever born to young women (15-19 and 20-24 years) in a given district is large, then infant mortality is likely to be high, mainly because there are a large number of children in a situation of higher mortality risks. This is because the level of mortality affecting children born to very young women is usually higher than the average mortality for all children. Hence, the level of mortality in a township would be influenced by the proportion of children born into a category of higher risk (very young mothers). Table 8 examines this possible source of bias.

According to Table 8, in Kachin, as in the case of districts, only a small percentage of women declared to have had children between ages 15 to 24: only 5.0%. In all townships, percentages are below 10%. Therefore, in the townships, a bias in the level of infant mortality caused by a large proportion of children ever born from young women (15-24 years) does not occur. The differences in IMR among townships or between the state/region and its townships cannot be explained by this factor.

It can be concluded that the sub-area method to estimate IMR also provides reliable results in most townships. However, in those townships with very small populations and, therefore, very few infant deaths, the method may provide unreliable results. As suggested above, in this case it is recommended to merge two townships; the issue is which townships are more appropriate to merge.

TABLE 8
Proportion of children ever born from women 15-24 years old

Kachin	Proportion	Kachin	Proportion	Kachin	Proportion
Total	0.0496				
Myitkyina	0.0434	Panwa	0.0913	Myohla	0.0650
Waingmaw	0.0510	Mohnyin	0.0426	Lwe'ge'	0.0426
Ingyanyan	0.0285	Mogaung	0.0483	Dotphoneyan	0.0737
Tanaing	0.0682	Phakant	0.0516	Putao	0.0534
Chiphwe	0.0682	Hopin	0.0340	Sumprabum	0.0390
Hsotlaw	0.0652	Kamine	0.0450	Machanbaw	0.0528
Hsinbo	0.0493	Bhamo	0.0513	Khaunglanphoo	0.0190
Hsadone	0.0495	Shwegu	0.0552	Naungmoon	0.0352
Kanpaukti	0.0824	Momauk	0.0514	Pannandin	0.0539
Shinbwayyan	0.0719	Mansi	0.0580		

Source: Own estimates based on 2014 Myanmar Census data.

To address this issue, the following approach was used: three census variables referring to townships were considered: percentage of households with electricity, percentage of households with at least one of selected household items,⁶ and percentage of women aged 15 and older who are literate. With these three variables, a simple summatory index was constructed.⁷ This index can be considered a rough indicator of townships' socio-economic development. The townships with small number of cases were merged with the adjacent townships with the closest index value. The larger townships with which the small townships were merged had to belong to the same district. The respective data are in Table 9 and the infant mortality rates for the merged townships are in Table 10. As previously noted, there were five townships that were considered too small for this estimation; they have less than 100 dead children among those born to women between 20-39 years old.

For example, Ingyanyan, with only 85 dead children from births to women 20-39 years, was merged with Chiphwe. There are five townships in Myitkyina District that share borders with Ingyanyan: Waingmaw, Chiphwe, Hsotlaw, Kanpaikti and Myitkyina (see Map 2). The index value for Ingyanyan is 47.7 and the adjacent township with the nearest value is Hsadone with 47.9. However, the border they share is very small; therefore, the township with the second nearest value was selected, Chiphwe, with an index value of 51.2 and a large shared border. The new unit with the combined populations of these two townships has an IMR of 69.8. The same procedure was followed in the other four small population townships. There was, however, an exception: Khaunglanphoo

In some cases, the resulting unit from the merger of two townships may reflect a somewhat distorted reality. This is probably the case of Khaunglanphoo (see Table 10). In fact, this township is a special case. Its IMR is unrealistically low; the number of children dead among children ever born to women aged 20-24 or even to women aged 20-39 is small, but the population of this township is comparatively large (see Table 6). As a result of its population size, any township with which it merges will result in a new unit with a possibly biased IMR. The rate estimated for Khaunglanphoo will prevail given its population size. If its IMR is underestimated, which is likely the case, the rate of the new unit will also be underestimated. This township was combined with Naungmoon, which is adjacent and has a similar socio-economic level according to the index (see Table 9). The resulting unit exhibits an IMR of 37.2. If the Khaunglanphoo rate were underestimated, the combination of the two townships would not render a unit with a reliable IMR. The merger would not solve the problem. An option could be assigning to Khaunglanphoo the same rate estimated for Naungmoon. This was the solution adopted here (see Tables 6 and 10). A different solution could be adopted by conducting an in-depth study of the characteristics of both townships. This, however, is beyond the purposes of this study.

⁶ The items are: radio, TV, landline phone, computer, and internet.

⁷ Index values are calculated simply as the sum of the three variables divided by 3.

TABLE 9
Indicators of socio-economic development and simple summatory index

Township	Percentage household with electricity	Percentage household with at least one selected amenity	Percentage of literate women (15 years and older)	Summatory Index
Myitkyina				
Myitkyina	52.3	86.0	92.3	76.9
Waingmaw	12.8	70.0	79.4	54.1
Ingyanyan	16.5	50.5	76.1	47.7
Tanaing	19.2	74.1	82.2	58.5
Chiphwe	38.1	58.8	56.8	51.2
Hsotlaw	14.2	47.3	67.7	43.1
Hsinbo	16.8	68.3	91.5	58.9
Hsadone	4.2	62.1	77.4	47.9
Kanpaikti	9.4	62.0	64.1	45.2
Shinbwayyan	7.8	62.4	75.6	48.6
Panwa	27.7	52.1	49.3	43.0
Mohnyin				
Mohnyin	25.7	82.4	93.7	67.3
Mogaung	43.4	79.7	91.5	71.5
Phakant	15.2	82.3	92.8	63.4
Hopin	43.6	85.8	95.0	74.8
Kamine	8.6	71.7	89.1	56.5
Bhamo				
Bhamo	41.0	84.6	93.0	72.9
Shwegu	21.9	74.3	96.4	64.2
Momauk	23.1	71.6	81.9	58.9
Mansi	25.1	77.4	86.9	63.1
Myohla	14.1	74.9	94.2	61.1
Lwe'ge'	89.2	89.5	75.0	84.6
Dotphoneyan	6.3	62.0	71.9	46.7
Putao				
Putao	9.1	41.1	83.9	44.7
Sumprabum	25.1	76.0	70.2	57.1
Machanbaw	17.3	43.3	83.8	48.1
Khaunglanphoo	8.5	60.1	79.7	49.4
Naungmoon	16.2	34.8	88.9	46.6
Pannandin	9.7	46.1	71.7	42.5

Source: Own estimates based on 2014 Myanmar Census data.

TABLE 10
Infant mortality rates in merged townships

Merged townships	Infant mortality rate
Chiphwe-Ingyanyan	69.8
Mansi-Lwe'ge'	49.0
Naungmoon-Khaunglanphoo	60.9
Machanbaw-Sumprabum	63.0
Putao-Pannandin	75.0

Source: Own estimates based on 2014 Myanmar Census data.

To sum up, it is possible to state that the sub-area method provides reliable estimates of IMR in small areas. However, the consistency of the results should always be examined, especially in administrative units where the number of cases is too small. *Ad hoc* solutions should be adopted such as those proposed here: merging two similar adjacent small areas or adopting the rate of an adjacent area with similar characteristics.

Considerations on sources of errors and probable magnitude of statistical estimation errors

As in any estimation procedure based on sample surveys, two types of errors may affect the results obtained: the first one pertains to inaccuracies in the basic data, which may result from errors in the responses given by the informant during the survey interview and/or omissions in the survey or census enumeration. The second type of error relates to random statistical variations associated with the number of observations on which the parameter's estimate is based. Procedures to assess/adjust demographic data from census and surveys are abundant in demographic literature and will not be dealt with here.

In addition to technical procedures to assess and adjust data, demographers also judge the reliability of demographic estimates based on the consistency between these estimates and the inherent discipline most often observed in demographic processes, as well as the consistency between the demographic parameters and other social indicators, such as level of education, access/availability of social services, sanitation infrastructure, area of residence, and so on. All these tools are essential to evaluate the reliability of the estimation results; most frequently, they constitute a routine part of the demographic analyses, and obviously, they are part of the options available to assess the results obtained from the “*sub-area mortality estimation method*” discussed here.

As for random statistical variations that may affect estimates from the “*sub-area mortality estimation method*”, those results are derived from the proportions (D_i) of dead children, classified by age of the mother. Hence, the standard deviation of the estimate is related to the standard deviation of sample estimates of proportions. D_i represents the proportion of children dead by age of the mother, where “ i ” indicate the mother's age group:

$$D_i = (\text{number of children dead}) \div (\text{total children born alive to women in age group } i).$$

If we represent the number of children dead among the children born alive to women in age group i with x , and N is the total children born alive to women in age group i , then $D_i = x/N$. Hence, the mean number of children dead to women aged “ i ” would be $x = ND_i$, and the variance for children dead to women aged “ i ”, symbolized by σ^2 , would be $\sigma^2 = ND_i(1-D_i)$; which characterizes a binomial distribution.

To simplify notation in the formulae that follows, the “true” proportion of children dead among those born alive to women in age group “ i ” in the total population will be symbolized by “ π ”, while the proportion of children dead as measured from the sample observations (which we have been calling D_i) will be represented by p . Therefore, p will

be the sample estimate for π . The binomial distribution quickly converges to a normal distribution when N , the number of observations multiplied by the parameter π is larger than 5. For instance, if the true proportion of children deaths among all children born alive to women of age “ i ” in the population is $\pi = 0.1$, and $N = 50$ (the total number of children born alive is 50 or more), the normal distribution would provide a good approximation for the probability distribution of p , the estimate of the proportion of children deaths. If the “true” proportion π is smaller, such as $\pi = 0.05$, then the number of children born alive to women in the age group i would need to be over 100 to obtain a reasonable approximation to a normal distribution.

Taking into account the order of magnitude of cases observed in census or surveys, which we usually deal with when obtaining these estimates, it is evident that using the normal distribution to construct confidence intervals for these proportions would be a very safe assumption in most situations. In the case of this exercise, the smallest number of children born alive were recorded in the township of Ingyanyan, where $N = 536$, since the proportion of children who died was 0.15858, therefore $N\pi = 85$; then, in Lwe'ge' where $N = 1711$, the proportion of children dead shows the lowest level among these townships: 0.02046, thus $N\pi = 35$, still a number large enough to ensure a good approximation to a normal distribution for “ p ” (or proportion D_i in the sample), which is the indicator utilized to estimate of the parameter π (the “true” proportion in the population).

The proportion of children dead among the total number of children ever born observed in the sample, represented here by p , is an unbiased estimate of the true proportion of children deaths π among the total number of children born alive to women in age group i , under the prevailing mortality risks of the population under study. The estimate p is a minimum variance estimate that, when the sample size is large enough, it will have a normal distribution with media π , and standard deviation $\sigma_p = \sqrt{[\pi(1-\pi)/N]}$.

Hence, the estimation error for π , is the difference between the observed proportion in the survey, and the “true” value in the population: $e = p - \pi$. Therefore, the probability that the error e would be equal or smaller than $z_{(\alpha/2)} \sigma_p$ is equal to $1-\alpha$, where α is the significance level, which determines the value $z_{(\alpha/2)}$ in the normal distribution. In other words, the probability that $(p - \pi)$ in absolute value would be equal or smaller than $z_{(\alpha/2)} \sigma_p$ is equal to $1-\alpha$. Using statistical notation, this would be expressed as: $\Pr\{-z_{(\alpha/2)} \sigma_p \leq (p - \pi) \leq z_{(\alpha/2)} \sigma_p\} = 1-\alpha$; where \Pr stands for “probability”. The value of σ_p is expressed as a function of the true proportion π in the population, but we will use the observed proportion of children dead p (calculated from the sample, that is the proportion D_i) to approximate the value π , and from that we obtain an estimated value for the standard deviation; $\sigma_p = \sqrt{[p(1-p)/N]}$.

Then: $e = (p-\pi) \leq z_{(\alpha/2)} \sigma_p$; that is: $e \leq z_{(\alpha/2)} \sigma_p (1)$; where $\sigma_p = \sqrt{[\pi(1-\pi) / N]}$.

From (1) the sample size, N , can be estimated: $N \geq \pi(1-\pi) [z_{(\alpha/2)} \div e]^2$.

If the level of error e is given as a fixed value, we can obtain N , the number of observations (number of children born alive in the sample) required to keep the estimated value p within a given margin of error from the true value π . In some cases, it may be more

convenient to express the error not in absolute terms, but as a percentage “ ∂ ” of the true proportion π (that is, $e = \partial\pi$); then the number of observations “ N ” required (total number of children born alive to women in age group i) would be:

$$N \geq \pi (1 - \pi) \{z_{(\alpha/2)}\}^2 / (\partial \cdot \pi)^2 = [(1 - \pi) / \pi] \{z_{(\alpha/2)} / \partial\}^2$$

Since the standard deviation for the estimate p , which is a proportion, is a function of this proportion, the minimum sample size required for a given significance level and a given margin of error varies both with the maximum size of error requested and with the value of the proportion: the smaller the proportion (rare events, which in this case would mean lower level of mortality), the larger the required sample size (number of children born alive in the respective age group of women). When the level of mortality is very low, a larger number of children born alive would be needed in order to observe a significant number of deaths among those children ever born, and keep the error within the set margins.

To illustrate the level of magnitude of these variations, Table 11 and Table 12 have been calculated; Table 11 presents values for the standard deviation of the estimator (proportion of children dead) and table 12 presents the required number of observations (sample size) of children ever born to women in the given age group, in order to maintain the level of error within certain percentage difference from the “true” value of the mortality estimator (proportion of children deceased).

TABLE 11
Estimates of standard deviation by sample size (N) and by a range of values for the proportion "Di"

N value	Values of the parameter "Di"						
	0,02	0,05	0,1	0,15	0,2	0,25	0,3
30	0.02556	0.03979	0.05477	0.06519	0.07303	0.07906	0.08367
40	0.02214	0.03446	0.04743	0.05646	0.06325	0.06847	0.07246
50	0.01980	0.03082	0.04243	0.05050	0.05657	0.06124	0.06481
60	0.01807	0.02814	0.03873	0.04610	0.05164	0.05590	0.05916
70	0.01673	0.02605	0.03586	0.04268	0.04781	0.05175	0.05477
80	0.01565	0.02437	0.03354	0.03992	0.04472	0.04841	0.05123
90	0.01476	0.02297	0.03162	0.03764	0.04216	0.04564	0.04830
100	0.01400	0.02179	0.03000	0.03571	0.04000	0.04330	0.04583
110	0.01335	0.02078	0.02860	0.03405	0.03814	0.04129	0.04369
120	0.01278	0.01990	0.02739	0.03260	0.03651	0.03953	0.04183
130	0.01228	0.01912	0.02631	0.03132	0.03508	0.03798	0.04019
140	0.01183	0.01842	0.02535	0.03018	0.03381	0.03660	0.03873
150	0.01143	0.01780	0.02449	0.02915	0.03266	0.03536	0.03742

Sources: Own estimates.

TABLE 12

Number of observations (N) needed to keep the error within a percentage difference of the true proportion

Number of observations (N) needed to keep the error within one tenth of the true "Di" value			
Confidence level:	90%	95%	99%
"z" value:	1.645	1.960	2.576
"Di" values:	Required "N" values		
0.01	26790	38032	65694
0.02	13260	18824	32515
0.03	8749	12421	21456
0.05	5141	7299	12608
0.10	2435	3457	5972
0.15	1533	2177	3760
0.20	1082	1537	2654
0.25	812	1152	1991
0.30	631	896	1548
Number of observations (N) needed to keep the error within 1.5 tenth of the true "Di" value			
Confidence level:	90%	95%	99%
"z" value:	1.645	1.960	2.576
"Di" values:	Required "N" values		
0.01	11907	16903	29197
0.02	5893	8366	14451
0.03	3889	5521	9536
0.05	2285	3244	5604
0.10	1082	1537	2654
0.15	682	968	1671
0.20	481	683	1180
0.25	361	512	885
0.30	281	398	688
Number of observations (N) needed to keep the error within two tenth of the true "Di" value			
Confidence level:	90%	95%	99%
"z" value:	1.645	1.960	2.576
"Di" values:	Required "N" values		
0.01	6697	9508	16424
0.02	3315	4706	8129
0.03	2187	3105	5364
0.05	1285	1825	3152
0.10	609	864	1493
0.15	383	544	940
0.20	271	384	664
0.25	203	288	498
0.30	158	224	387
Number of observations (N) needed to keep the error within three tenth of the true "Di" value			
Confidence level:	90%	95%	99%
"z" value:	1.645	1.960	2.576
"Di" values:	Required "N" Values		
0.01	2977	4226	7299
0.02	1473	2092	3613
0.03	972	1380	2384
0.05	571	811	1401
0.10	271	384	664
0.15	170	242	418
0.20	120	171	295
0.25	90	128	221
0.30	70	100	172

Source: Own estimates.

The sample size is very sensitive to the level of mortality. The rarer the event, the higher the number of cases required in order to account for a significant number of events, at a given level of confidence. Similarly, if the margin of error were to be kept within a narrower interval around the true value of the estimated parameter, the number of observations would increase in order to reduce the standard deviation and narrow the margin of error. Therefore, to assess the level of confidence associated to the estimated mortality rates obtained, it is necessary to bear in mind both the level of mortality (proportion of children who have died) and the confidence interval or the margin of error within which the estimation of the parameter is expected to be maintained. As indicated earlier, in the calculations in Table 11 and Table 12 the value for the standard deviation σ_p , which is a function of the parameter π , was approximated by using the observed proportion p in place of π .

Table 13 presents the expected errors e , both in absolute value and as a percentage of the proportion of children deceased (D_i), given the number of children born alive (N) in the townships. The differences between the estimated value and the “true” value in the population are expected to be smaller than the calculated e value in 90, 95, and 99 per cent of cases, according to these levels of confidence, as showed in Table 13.

Given the numbers of children born alive observed (N), the level of errors are relatively small in general, particularly if we consider a level of confidence of 90 or 95 per cent. At 95 per cent, the error would be more than 10 per cent of the true proportion value in just 14 townships, and over 20 per cent of the true value in only 4 out of 29 townships. In relative terms, the largest error is in Lwe'ge', where the level of mortality is the lowest ($D_i = 0.02$) and the number of children born alive –compared to other townships – is not very high ($N=1711$). However, in absolute terms the error is expected to be smaller than 0.0067 in 95 out of 100 samples of this size; a difference plus/minus seven per thousand or smaller is not a big error in absolute terms, for these type of estimates. Relative differences in other townships are significantly smaller than in Lwe'ge', either because the number of children born alive was higher or the level of mortality was higher. Therefore, in general terms, we can conclude that the level of statistical accuracy expected from these estimates of infant and child mortality is very reasonable. Naturally, this conclusion is from a statistical point of view, which takes into account the random variations associated to the number of observations. Yet, the estimates may be unreliable on the basis of response errors or other factors which may generate biases or errors associated to the interview or the data gathering process.

As mentioned at the beginning of this section, the reliability of the estimates can be further assessed by verifying the degree of consistency these estimates have with other socio-economic and demographic variables, like education level in the population, sanitation infrastructure in the townships, access to social services, communication, electricity, safe drinking water, etc.

TABLE 13
Maximum margin of error "e" expected in the estimation of proportion Di (age group 20-39) for confidence level 90%, 95% and 99%, in absolute terms and as percentage of the observed "Di" value

Township	N	Di	N.Di	Probable margin of error (+/-) in the estimated proportion given the N and Di values in each Township, for 3 significance levels. Where: $e \leq z(\alpha/2) \sigma/\sqrt{n}$					
				Absolute value for			Value of error "e" as percentage of the proportion Di		
				$e = (p - \pi) \leq z(\alpha/2) \sigma/\sqrt{n}$					
				90%	95%	99%	90%	95%	99%
1.645	1.96	2.576	1.645	1.96	2.576	(% of Di)	(% of Di)	(% of Di)	
Myitkyina	59,648	0.06164	3,677	0.00162	0.00193	0.00254	2.6	3.1	4.1
Waingmaw	27,063	0.09304	2,518	0.00290	0.00346	0.00455	3.1	3.7	4.9
Ingyanyan	536	0.15858	85	0.02595	0.03092	0.04064	16.4	19.5	25.6
Tanaing	10,623	0.09093	966	0.00459	0.00547	0.00719	5.0	6.0	7.9
Chiphwe	3,450	0.09652	333	0.00827	0.00985	0.01295	8.6	10.2	13.4
Hsotlaw	2,109	0.11522	243	0.01144	0.01363	0.01791	9.9	11.8	15.5
Hsinbo	2,859	0.14341	410	0.01078	0.01285	0.01689	7.5	9.0	11.8
Hsadone	2,739	0.07302	200	0.00818	0.00974	0.01281	11.2	13.3	17.5
Kanpaikti	2,365	0.07315	173	0.00881	0.01049	0.01379	12.0	14.3	18.9
Shinbwayyan	3,222	0.10335	333	0.00882	0.01051	0.01382	8.5	10.2	13.4
Panwa	2,753	0.07701	212	0.00836	0.00996	0.01309	10.9	12.9	17.0
Mohnyin	36,230	0.07543	2,733	0.00228	0.00272	0.00357	3.0	3.6	4.7
Mogaung	30,622	0.07126	2,182	0.00242	0.00288	0.00379	3.4	4.0	5.3
Phakant	39,054	0.08463	3,305	0.00232	0.00276	0.00363	2.7	3.3	4.3
Hopin	10,043	0.05775	580	0.00383	0.00456	0.00600	6.6	7.9	10.4
Kamine	4,821	0.10620	512	0.00730	0.00870	0.01143	6.9	8.2	10.8
Bhamo	28,154	0.07299	2,055	0.00255	0.00304	0.00399	3.5	4.2	5.5
Shwegu	21,352	0.07498	1,601	0.00296	0.00353	0.00464	4.0	4.7	6.2
Momauk	10,042	0.08524	856	0.00458	0.00546	0.00718	5.4	6.4	8.4
Mansi	13,935	0.08016	1,117	0.00378	0.00451	0.00593	4.7	5.6	7.4
Myohla	1,188	0.12626	150	0.01585	0.01889	0.02482	12.6	15.0	19.7
Lwe'ge'	1,711	0.02046	35	0.00563	0.00671	0.00882	27.5	32.8	43.1
Dotphoneyan	3,062	0.11724	359	0.00956	0.01140	0.01498	8.2	9.7	12.8
Putao	17,328	0.11236	1,947	0.00395	0.00470	0.00618	3.5	4.2	5.5
Sumprabum	714	0.09944	71	0.01842	0.02195	0.02885	18.5	22.1	29.0
Machanbaw	2,495	0.09339	233	0.00958	0.01142	0.01501	10.3	12.2	16.1
Khaunglanphoo	2,297	0.03396	78	0.00622	0.00741	0.00973	18.3	21.8	28.7
Naungmoon	1,420	0.09155	130	0.01259	0.01500	0.01971	13.8	16.4	21.5
Pannandin	581	0.12220	71	0.02235	0.02663	0.03500	18.3	21.8	28.6

Source: Own estimates based on 2014 Myanmar Census data.

Conclusions

The method proposed here, the *sub-area method*, provides reliable and consistent infant mortality estimates even in very small areas. Its results were compared to those obtained with a Brass-type method in small sub-divisions, but large enough to use the classic method (districts). The results were similar. In a second exercise, the method was applied to very small units (townships) and, in general, results appear reliable, except in some cases where the number of dead children was too small. The method was illustrated here only to estimate infant mortality, but it can be also used to estimate child and under-five mortality. The principles and procedures are the same.

As mentioned above, if the population of a unit is too small, the method may not work because of the influence of random factors, seasonal changes or location-specific factors, which are likely to produce distorted estimates in small areas. They may also affect the assumptions of the method. In this case, the solution proposed here was to merge two similar contiguous small areas or adopt the rate of an adjacent area with similar characteristics. The similarity was evaluated with a simple socio-economic index constructed with three census indicators. Other *ad hoc* solutions can be considered.

It is important to state that this new method proposed here shares all the assumption and limitations of the Brass-type methods. In addition, it does not deal with other problems that may affect small area statistics, such as undercounting of population or events (birth, deaths). Therefore, results should be treated with caution. There are, however, assumptions specific to this method. They refer mainly to the comparability of the distribution of dead children along age groups of women between the parent area and the small area. In the example, the assumptions were satisfied.

Finally, it is important to mention that the method was tested in small areas with poor quality data (townships) and results were consistent. Clearly, the result may not be reliable because of data problems, but the method was able to perform well under this difficult context. Assumptions were satisfied and results were consistent with expectations. The exceptions were those areas with very small number of cases, for which no method would deliver reliable results.

Declarations

Availability of data and material

All data generated or analyzed during this study are included in the following published census report:

THE REPUBLIC OF THE UNION OF MYANMAR. The 2014 Myanmar Population and Housing Census. Census report volume 4-B (mortality). Department of Population, Ministry of Labour, Immigration and Population (with technical assistance from UNFPA), Naypyitaw, 2016. Available in: <http://myanmar.unfpa.org/en/publications/thematic-report-mortality>.

References

- BRASS, W. Uses of census or survey data for the estimation of vital rates. *In: AFRICAN SEMINAR ON VITAL STATISTICS. Proceedings* [...]. Addis Ababa, 1964.
- FEENEY, G. Child survivorship estimation: methods and data analysis. *Asian and Pacific Population Forum*, v. 5, n. 2-3, p. 51-87, 1991.
- HILL, K. Indirect estimation of child mortality. *In: MOULTRIE, T. A.; DORRINGTON, R. E.; HILL, A. G.; HILL, K.; TIMÆUS, I. M.; ZABA, B. (ed.). Tools for demographic estimation*. Paris: International Union for the Scientific Study of Population, 2013. Available in: <<http://demographicestimation.iussp.org/content/indirect-estimation-child-mortality>>. Access in: 04 Dec. 2017.
- KINTNER, H. J. The life table. *In: SIEGEL, J.; SWANSON, D. (ed.). The methods and materials of demography*. Elsevier Academic Press, 2004. p. 301-340.
- MERCER, L. *et al.* Small area estimation of child mortality in the absence of vital registration. *The Annals of Applied Statistics*, v. 9, n. 4, p. 1889-1905, 2015.
- MINISTRY OF IMMIGRATION AND POPULATION. Department of Population. *The 2014 Myanmar Population and Housing Census*, volumes 3-A to 3-O (volumes for regions/states). Department of Population, Nay Pyi Taw, Myanmar, 2015. Available in: <http://www.dop.gov.mm/moip/index.php?route=product/category&path=54_52>. Access in: 04 Dec. 2017.
- PALLONI, A.; HELIGMAN, L. Re-estimation of structural parameters to obtain estimates of mortality in developing countries. *Population Bulletin of the United Nations*, n. 18, p. 10-33, 1985.
- POPOFF, C.; JUDSON, D. H. Some methods of estimation for statistically underdeveloped areas. *In: SIEGEL, J.; SWANSON, D. (ed.). The methods and materials of demography*. Elsevier Academic Press, 2004. p. 603-640.
- RAO, J. N. K. Some methods for small area estimation. *Rivista Internazionale di Scienze Sociali*, anno 116, n. 4, p. 387-406, 2008. Stable URL: <<http://www.jstor.org/stable/41625216>>. Access in: 04 Dec. 2017.
- ROWLAND, D. *Demographic methods and concepts*. New, York: Oxford University Press, 2003.
- SIAP – Statistical Institute for Asia and the Pacific. *Estimation of demographic parameters from census data*. London, UK: Institute for Asia and the Pacific, prepared by the Center of Population Studies, 1994.
- SWANSON, D. A. *Estimating the underlying infant mortality rates for small populations: a case study of counties in Estonia*. Center for Studies in Demography & Ecology University of Washington and Department of Sociology University of California Riverside, 2015.
- SWANSON, D. A.; KPOSOWA, A.; BAKER, J. Estimating the underlying infant mortality rates for small populations, including those reporting zero infant deaths: a case study of counties in California. *Population Review*, v. 58, n. 2, 2019.
- STONER, O.; ECONOMOU, T.; SILVA, G. D. M da. A hierarchical framework for correcting under-reporting in count data. *Journal of the American Statistical Association*, p. 1-17, 2019. 10.1080/01621459.2019.1573732.
- UNITED NATIONS. *Manual X*. Indirect techniques for demographic estimation. New York: United Nations, 1983.
- UNPD – United Nations Population Division. *MORTPAK for Windows* (Version 4.3). New York: United Nations Population Division, 2013.

About the authors

Ricardo Neupert is Ph.D. Sociology/Demography (Brown University, USA). Freelance consultant.

Rogelio Eduardo Fernandez Menjivar is M.Sc. Demography, Ph.D. candidate, Cedeplar – Universidade Federal de Minas Gerais, Brazil.

Rogelio Eduardo Fernandez Castilla is Ph.D. Medical Demography, London School of Hygiene and Tropical Medicine, University of London. Independent consultant.

Contact address

Ricardo Neupert

Alcalde Rutilio Rivas, 7031, La Reina
7850195 – Santiago, Chile

Rogelio Eduardo Fernandez Menjivar

Rua Jose Ribeiro Filho, 35, apto. 1101, Bloco 2, Bairro Ouro Preto
31330-500 – Belo Horizonte-MG, Brasil

Rogelio Eduardo Fernandez Castilla

Santa Fe, 532
4000 – San Miguel de Tucuman, Tucuman, Argentina

Resumo

Estimación indirecta da mortalidade infantil em pequenas áreas

Métodos indiretos do tipo Brass para estimar a mortalidade nas primeiras idades da vida têm sido usados por mais de quatro décadas, fornecendo estimativas muito robustas para países que não possuem sistemas confiáveis de estatísticas vitais. No entanto, quando as áreas de estimativa são pequenas, o número de crianças mortas pode ser muito baixo, especialmente entre os nascidos de mulheres jovens, que fornecem informações essenciais para estimar a mortalidade recente. Nestes casos, as estimativas podem ser afetadas por erros aleatórios e flutuações anuais inesperadas. Ao mesmo tempo, embora seja muito improvável que as tendências demográficas em uma área pequena sigam padrões muito diferentes daqueles prevalentes no ambiente mais amplo ao qual pertencem, é possível que certos eventos locais se tornem relevantes em pequenas áreas, causando alguns desvios de padrões que na área maior são válidos. O objetivo deste trabalho é propor uma adaptação da abordagem de estimativa indireta, que permita obter estimativas de mortalidade infantil e das crianças em pequenas áreas. Dessa forma, tal proposta pertence ao escopo dos métodos de estimativa indireta, compartilhando as limitações e vantagens que caracterizam essa metodologia de estimativa. O método é ilustrado com dados do Censo de População e Habitação de Myanmar, 2014. Os resultados indicam que o método proposto aqui fornece estimativas confiáveis e consistentes de mortalidade infantil, em comparação com os resultados do método original de Brass, mesmo em áreas muito pequenas.

Palavras-chave: Mortalidade infantil. Métodos indiretos. Censos. Pequenas áreas. Mianmar.

Resumen

Estimación indirecta de la mortalidad infantil en pequeñas áreas

Los métodos indirectos del tipo Brass para estimar la mortalidad en las edades tempranas se han utilizado durante más de cuatro décadas y han proporcionado estimaciones muy robustas para países que no cuentan con sistemas de estadísticas vitales fiables. Sin embargo, cuando las áreas de estimación son pequeñas, el número de niños muertos podría ser muy reducido, especialmente entre los nacidos de mujeres jóvenes, quienes proporcionan información esencial para estimar la mortalidad reciente. En estos casos, las estimaciones podrían verse afectadas por errores aleatorios y fluctuaciones anuales inesperadas. Al mismo tiempo, aunque es muy poco probable que las tendencias demográficas en un área pequeña sigan patrones muy diferentes de los que prevalecen en el entorno más amplio al cual pertenecen, es posible que ciertos eventos locales se tornen relevantes en áreas pequeñas, causando algunas desviaciones de los supuestos que para el área más grande pueden ser válidos. El objetivo de este trabajo es proponer una adaptación del enfoque de estimación indirecta que permita obtener estimaciones de mortalidad infantil y en la niñez para áreas pequeñas. Como tal, esta propuesta pertenece al ámbito de los métodos de estimación indirecta, por lo que comparte las limitaciones y ventajas que caracterizan a esta metodología de estimación. El método se ilustra con datos del censo de población y vivienda de Myanmar de 2014. Los resultados obtenidos indican que el método aquí propuesto brinda estimaciones de mortalidad infantil confiables y consistentes, en comparación con los resultados del método original de Brass, incluso en áreas muy pequeñas.

Palabras clave: Mortalidad infantil. Métodos indirectos. Censo. Áreas pequeñas. Myanmar.

Received for publication in 06/06/2019

Approved for publication in 13/10/2019