This paper is published in *Demography and health issues (p51-62),* edited by Christos H Skiadas and Charilaos Skiadas, *The Springer series on demographic methods and population analysis 46.* edited by Kenneth C. Land, *Springer.*

# Using Child, Adult, and Old-age Mortality to Establish a Developing Countries Mortality Database (DCMD)

Nan Li[1], Hong Mi[2], and Patrick Gerland[1]

1. Population Division, Department of Economic and Social Affairs, United Nations, 2 UN Plaza, 233 East 44th Street, Room DC2-1938, 1934, New York, NY 10017, USA (Email: li32@un.org, garland@un.org). The views expressed in this paper are those of the author and do not necessarily reflect those of the United Nations.

2. Corresponding author，School of Public Affairs, Zhejiang University, Room 265#, Mengminwei Building, Zijin'gang Campus, Hangzhou, Zhejiang, P. R. China. (E-mail: spsswork@163.com). The work on this paper was supported by the Nature Science Foundation of China (NSFC) project (NO.71490732), NSFC project (NO.71490733) , Zhejiang Social Science Planning Project Key Program(NO.17NDJC029Z), and Nature Science Foundation of Zhejiang project(NO.LZ13G030001)

**Keywords:** Life table, Database, Developing countries

**Abstract.** Life-table databases have been established for developed countries and effectively used for various purposes. For developing countries of which the deaths counted 78% that of the world in 2010-2015, however, reliable life tables can hardly be found. Indirect estimates of life tables using empirical data on child and adult mortality are available for developing countries. But more than half of all deaths already occurred at age 60 and higher in developing countries in 2010-2015, which leads to the irony that worldwide the number of deaths at old-ages is the biggest, and also the least reliable. This reality indicates that improving the estimates of old-age mortality for individual developing countries is not enough, and that establishing a life-table database for all developing countries, which utilizes the improved estimations of old-age mortality, is necessary. To fulfill this task, we introduce two methods: (1) the Census Method that uses populations enumerated from census to estimate old-age mortality, and (2) the three-input model life table that utilizes child, adult, and old-age mortality to calculate life tables. Compared to using only child and adult mortality, applying the two methods to the data from the Human Mortality Database after 1950, the errors of fitting old-age mortality are reduced for more than 70% of all the countries. For the three non-European-origin populations in the Human Mortality Database the errors are reduced by 17% for Chile, 48% for Japan, and 17% for Taiwan, which is more relevant for developing countries. These results indicate that the methodology is adequate and empirical data are available to establish a mortality database for developing countries.

## 1 Introduction

Empirical data used in estimating life tables are collected from three types of source: (1) death registration that counts deaths by sex and age in a certain period, usually a calendar year; (2) census that enumerates the numbers of population by age and sex at a certain time point, and sometimes also death by age and sex during a period before the census time; and (3) sample survey that, in principle, could collect data on both death and population but cover only a small portion of the population in a country. Censuses are conducted in almost all the countries of the world. Besides providing middle-year populations to compute death rates for countries with reliable death registration, some developing countries rely also on census to obtain life tables directly. Since census interviewers must visit every household in a country to enumerate the number of residents at a certain time point, they could also ask just one more question about whether there was a death, or were deaths, in the household in past year; and if yes what is the gender and age of the death, or the genders and ages of the deaths (United Nations Statistics Division (UNSD), 2008). Furthermore, using population data of two successive censuses, some mortality indicators of the period between the two censuses could be estimated, especially for old ages at which the effect of migration is negligible

(Li and Gerland, 2013). For many countries, census data on population by age and sex can be found from the United Nations Demographic Yearbook (e.g., UNSD, 2013a). Occasionally, surveys using large sample size could also provide life tables.

Typical sample surveys often collect information only from a small portion of the population. Subsequently, they cannot produce life tables. This is because death rates at some ages, for example 10-20 years, could be very low, and hence require a large population to be estimated reliably. Nonetheless, sample surveys could provide reliable indicators of mortality for certain age groups when death is not a rare event or when the age group is wide enough. The most commonly sampled mortality indicator is child mortality, which is the probability of dying between birth and age 5, and is often denoted as ${}_5q_0$. The United Nations Children's Fund (UNICEF) as part of the United Nations Inter-agency Group for Child Mortality Estimation (IGME) has been regularly collecting, analyzing, and publishing child mortality for most of the countries back to the 1970s or earlier (see United Nations Children's Fund, 2013; http://www.childmortality.org). Based on the same principles used to estimate child mortality using birth histories, surveys such as the Demographic and Health Surveys (DHS, http://www.measuredhs.com/) have been collecting sibling histories since the 1990s to measure adult mortality, allowing to derive the probability of dying between age 15 and 50 or 60 years, namely ${}_{35}q_{15}$ or ${}_{45}q_{15}$, respectively, for an increasing number of developing countries (Timæus, 2013). Combining data of surveys and other sources, Wang and colleagues (2012) at the Institute for Heath Metrics and Evaluation (IHME) estimated adult mortality for 187 countries from 1970 to 2010.

Life-table databases have been established for developed countries (e.g., Human Mortality Database (HMD), 2016) and effectively used for various purposes. For developing countries of which the deaths counted 78% that of the world in 2010-2015 (United Nations Population Division (UNPD), 2015), however, reliable life tables can hardly be found. Indirect estimates of life tables have been provided by the UNPD (2015) and IHME (Wang and colleagues, 2012) for developing countries, using empirical data on child mortality (${}_5q_0$) and adult mortality (${}_{45}q_{15}$). But more than half of all deaths already occurred at age 60 and higher in developing countries in 2010-2015. Thus, estimating old-age mortality (${}_{15}q_{60}$), and using it together with the ${}_5q_0$ and ${}_{45}q_{15}$ estimated by the UNICEF and IHME mentioned above, to establish a life-table database for developing countries is a relevant and urgent task. To fulfil this task, this paper introduces two methods: (1) the Census Method that uses populations enumerated in census to estimate ${}_{15}q_{60}$, and (2) the three-input model life table that utilizes ${}_5q_0$, ${}_{45}q_{15}$, and ${}_{15}q_{60}$ to calculate life tables. Compared to using only child and adult mortality, applying the two methods to the data of HMD after 1950, the errors of fitting old-age mortality are reduced for more than 70% of all the countries. To be more specific to developing countries, the errors are reduced by 17% for Chile, 48% for Japan, and 17% for Taiwan, for two sexes combined, which are the three non-European-origin populations in Human Mortality Database. These results indicate that, in order to establish a life-table database for developing countries, the methodology is adequate and the empirical data are available.

## 2 Methods

The methods include the Census Method and the three-input model life table (three-input MLT).

### 2.1 The Census Method

The Census Method utilizes populations enumerated from census to estimate ${}_{15}q_{60}$, and includes two models. The first is the Census Method with variable-r model (Bennett.andHoriuchi,1981;Li and Gerland, 2013), which is more suitable when the period between the two successive censuses is not close to 10 years; and the second is the Census Method with survival model, which should work better when the period is close to 10 years.

#### *2.1.1 The Census Method with variable-r model*

The variable-r model (Bennett.andHoriuchi,1981) assumes zero migration and evenly distributed enumeration errors over age. Let *p(x, t)* be the observed number of population in age group *[x,x+5)* enumerated from a census conducted at time *t*, where *x*=60, 65, 70. The growth rates at age *x* are computed as

$$r(x) = Log[\frac{p(x,t_2)}{p(x,t_1)}]/(t_2 - t_1), \quad x = 60,65,70, \qquad (1)$$

where $t_1$ and $t_2$ represent the date of the first and second census, respectively. And the accumulated growth rates are

$$\begin{aligned} s(60) &= 2.5r(60), \\ s(65) &= 5r(60) + 2.5r(65), \\ s(70) &= 5[r(60) + r(65)] + 2.5r(70). \end{aligned} \qquad (2)$$

Further, the middle-point population in age group *[x,x+5)*, *N(x)*, are estimated as

$$N(x) = \sqrt{p(x,t_1)\,p(x,t_2)}, \quad x = 60,65,70\,. \qquad (3)$$

Furthermore, the person-years lived in 5-year age group *[x,x+5)*, $L_x$ , in the underlying stationary population, are obtained as (Bennett And Horiuchi, 1981)

$$L_x = N(x)\exp[s(x)], x = 60,65,70. \qquad (4)$$

At old ages such as 60 and over, migrants are negligible comparing to deaths. Thus, the zero-migration assumption is naturally satisfied. In developing countries, however, the errors in enumerating population often occur unevenly across age. A typical example is age heaping. When such errors are severe, the $L_x$ resulted from (4), would show implausible patterns of increasing with age, which cannot occur in a stationary population. When such implausible situations occur, adjusting $L_x$ is necessary. Li and Gerland (2013) proposed such an adjustment as is shown in the appendix A, which provides the adjusted $\hat{L}_x$ . After adjusting the age-reporting errors, the number of survivors at age x, $l_x$ , can be estimated using nonlinear optimization and a Gompertz model (Li and Gerland, 2013), or it can be estimated locally linearly as below:

$$\begin{aligned} l_{65} &= \frac{\hat{L}_{60} + \hat{L}_{65}}{2.5} \frac{\hat{L}_{65}}{(\hat{L}_{60} + 2\hat{L}_{65} + \hat{L}_{70})}, \\ l_{70} &= \frac{\hat{L}_{65} + \hat{L}_{70}}{2.5} \frac{\hat{L}_{65}}{(\hat{L}_{60} + 2\hat{L}_{65} + \hat{L}_{70})}, \\ l_{60} &= \frac{\hat{L}_{60}}{2.5} - l_{65}, \\ l_{75} &= \frac{\hat{L}_{70}}{2.5} - l_{70}. \end{aligned} \qquad (5)$$

In (5), the $\frac{\hat{L}_{60}+\hat{L}_{65}}{2.5}$ and $\frac{\hat{L}_{65}+\hat{L}_{70}}{2.5}$ are the first-step estimates of $l_{65}$ and $l_{70}$, which are linear interpolations between $\hat{L}_{60}$, $\hat{L}_{65}$ and $\hat{L}_{70}$. The $\frac{\hat{L}_{65}}{(\hat{L}_{60}+2\hat{L}_{65}+\hat{L}_{70})}$ is an adjustment that makes $2.5\cdot(l_{60}+l_{65})=\hat{L}_{65}$ . The last two lines in (5) are linear formulas of calculating $\hat{L}_{60}$ and $\hat{L}_{70}$.

Finally, after estimating $l_x$, ${}_{15}q_{60}$ is obtained as

$$ {}_{15}q_{60}=1-\frac{l_{75}}{l_{60}} \tag{6} $$

*2.1.2 The Census Method with survival model*

When the period between the two successive censuses is close to 10 years, the populations between the period of exactly 10 years can be reliably estimated assuming over-time constant growth rates and using (1). Consequently, the 10-year survival ratio of the stationary population is estimated as

$$ S=\frac{L_{70}}{L_{60}}=\frac{p(70-74,t_2)}{p(60-64,t_1)}. \tag{7} $$

Assuming that the over-age survival ratio is constant, the 1-year and 15-year survival ratios are therefore $S^{\frac{1}{10}}$ and $S^{\frac{15}{10}}$, respectively. Subsequently, the 15-year probability of death between age 60 and 75 can be estimated as

$$ q=1-S^{\frac{15}{10}}. \tag{8} $$

The assumption of constant over-age survival ratio can be adjusted using the United Nations general model life table (UNPD, 1982), which leads to a more accurate estimate of old-age mortality as

$$ {}_{15}q_{60}=\begin{cases} q\cdot(1.021-0.0002\cdot q+0.0002\cdot q^2), & R^2=0.999, \quad \mathit{female}, \\ q\cdot(1.0153-0.0003\cdot q+0.0002\cdot q^2), & R^2=0.999, \quad \mathit{male}. \end{cases} \tag{9} $$

**2.2 The three-input model life table**

The three-input model life table is an augmentation of the flexible two-dimensional model life table (two-input MLT, Wilmoth et al, 2012), which is expressed as

$$ \log(m_x)=a_x+b_x\cdot\log({}_5q_0)+c_x\cdot[\log({}_5q_0)]^2+v_x\cdot k, \tag{10} $$

where $m_x$ stands for the five-year age-specific death rates with *x=0,1,5,10,…*; coefficient vectors $a_x$, $b_x$, $c_x$, and $v_x$ are obtained from fitting mortality data of the Human Mortality Database; and parameter *k* is flexible, which can be solved to fit an additional $_{45}q_{15}$. Obviously, the two-input MLT can be used to produce a life table when $_5q_0$ and $_{45}q_{15}$ are used as two inputs.

How to utilize the estimated old-age mortality ($_{15}\hat{q}_{60}$)? A simple answer (Li, 2014) can be found by following the logic of the Logit transformation: $\log[_x\hat{q}_0/(1-_x\hat{q}_0)] = \alpha + \beta \log[_xq_0/(1-_xq_0)]$, in which the standard $_xq_0$ is naturally that of the two-input MLT, and level $\alpha$ and pattern $\beta$ can be chosen to fit some function of observed probability of death ($_x\hat{q}_0$). When there is only $_{15}\hat{q}_{60}$, a customary is to set $\beta = 1$ and solve $\alpha$ to fit $_{15}\hat{q}_{60}$ (see Preston, Heuveline and Guillot, 2001; p.200). The rationale for using the Logit transformation is that $\log[_xq_0/(1-_xq_0)]$ would be close to linear at all the ages. It is worth noting that, at old ages, $\log(\hat{m}_x)$ would be close to linear according to the Gompertz law. Thus, at old ages, the linear relationship of the Logit transformation can be simplified as:

$$\log(\hat{m}_x) = \alpha + \log(m_x). \quad (11)$$

Because

$$_{15}\hat{q}_{60} \approx 1 - \exp[-5 \cdot (\hat{m}_{60} + \hat{m}_{65} + \hat{m}_{70})], \quad (12)$$

$\alpha$ is solved by inserting (11) to (12):

$$\alpha \approx \log[\frac{\log(1-_{15}\hat{q}_{60})}{\log(1-_{15}q_{60})}] \quad (13)$$

where $_{15}q_{60}$ is the old-age mortality of the two-input MLT. Subsequently, (10) is augmented to the three-input MLT:

$$\log(m_x) = \hat{a}_x + b_x \cdot \log(_5q_0) + c_x \cdot [\log(_5q_0)]^2 + v_x \cdot k, \quad (14)$$

$$\hat{a}_x = \begin{cases} a_x, & x < 60, \\ a_x + \log[\frac{\log(1-_{15}\hat{q}_{60})}{\log(1-_{15}q_{60})}], & x \geq 60, \end{cases} \quad (15)$$

which will exactly fit the three inputs: child, adult, and old-age mortality.

## 3 Validations

We use the data of HMD to test whether or not the three-input MLT (with $_5q_0$, $_{45}q_{15}$, and $_{15}q_{60}$) can improve the performance of the two-input MLT with only $_5q_0$ and $_{45}q_{15}$. We choose the periods after 1950 to avoid the irregular effect of World War II, and all the countries or areas except Israel, for which the Census Method could not work because of territory change. In HMD, all ‘census’ dates are adjusted to January first. Consequently, periods 1950-1959, 1960-1969, …, and 2000-2009, and the Census Method with survival model, are chosen to carry out the validations. In real census, there are undercounts. Nonetheless, these undercounts tend to cancel each other in causing the errors of estimating mortality level, as is indicated in appendix B.

We first choose the observed ${}_5q_0$ and ${}_{45}q_{15}$ of a certain population in a certain period as the inputs of two-input MLT, which will produce a life table that includes an estimated ${}_{15}\tilde{q}_{60}$. This ${}_{15}\tilde{q}_{60}$ will differ from the observed old-age mortality, ${}_{15}q_{60}$. We then use the 'census' populations at the two ends of each period to estimate the values of old-age mortality, and use an exponential model to smooth them. The results are denoted as ${}_{15}\hat{q}_{60}$.

The purpose of two-input MLT is to use the ${}_5q_0$ and ${}_{45}q_{15}$ to best describe the corresponding life table, including particularly the ${}_{15}q_{60}$, using the mortality patterns of the HMD populations. Thus, ${}_{15}\tilde{q}_{60}$ is the best estimated ${}_{15}q_{60}$ that the two-input MLT could provide. We believe that for developing countries ${}_{15}\tilde{q}_{60}$ should also be reasonable to some extent. Therefore, we use

$$ {}_{15}\bar{q}_{60} = [w \cdot {}_{15}\hat{q}_{60} + (1-w) \cdot {}_{15}\tilde{q}_{60}] \tag{16} $$

as the estimated old-age mortality of the three-input MLT, where the w stands for the weight that can be determined flexibly, and is taken as 0.5 in all the validations here. The values of ${}_{15}\bar{q}_{60}$ are input to the three-input model life tables, which will have the same ${}_5q_0$ and ${}_{45}q_{15}$ as that of two-input MLT. But the values of old-age mortality of these life tables are ${}_{15}\bar{q}_{60}$, which will differ from the observed ${}_{15}q_{60}$.

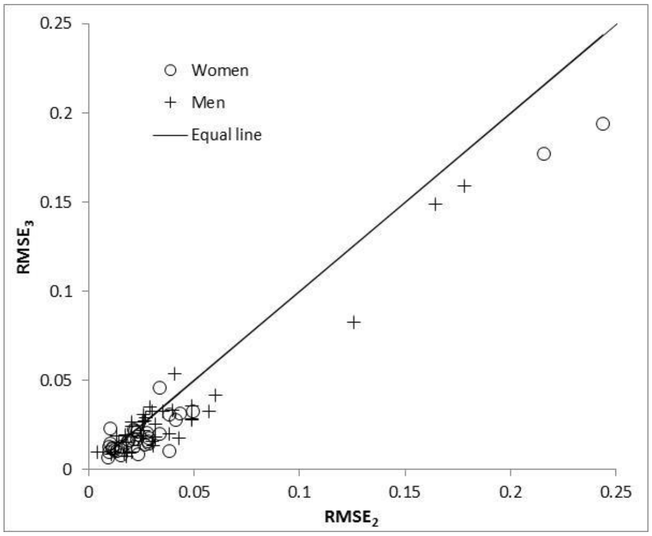
For a given population, we use the root-mean-squared error (RMSE) to measure errors. More specifically, we use $RMSE_2$ to indicate the difference between ${}_{15}\tilde{q}_{60}$ and ${}_{15}q_{60}$, and $RMSE_3$ to show the distance between ${}_{15}\bar{q}_{60}$ and ${}_{15}q_{60}$. Let the ith estimates be ${}_{15}\tilde{q}_{60}(i)$ and ${}_{15}q_{60}(i)$, and the total number of periods be n, there are

$$ \begin{aligned} RMSE_2 &= \sqrt{\sum_{i=1}^{n} [{}_{15}\tilde{q}_{60}(i) - {}_{15}q_{60}(i)]^2 / n}, \\ RMSE_3 &= \sqrt{\sum_{i=1}^{n} [{}_{15}\bar{q}_{60}(i) - {}_{15}q_{60}(i)]^2 / n}. \end{aligned} \tag{17} $$

If $RMSE_3 < RMSE_2$ for a given population, we conclude that the three-input MLT fits this population better than does the two-input MLT, and vice versa.

The validations use HMD data. If reliable life tables for developing countries were not rare, we would choose them to carry out the validation. For the 37 (excluding Israel) countries' 74 populations by sex in HMD, the results of validation are summarized in figure 1, in which the position of a population is marked by its $RMSE_2$ on the horizontal axis and $RMSE_3$ on the vertical axis. When the three-input MLT improves the performance of two-input MLT for a given population, the position of this population is below the equal line, and vice versa.

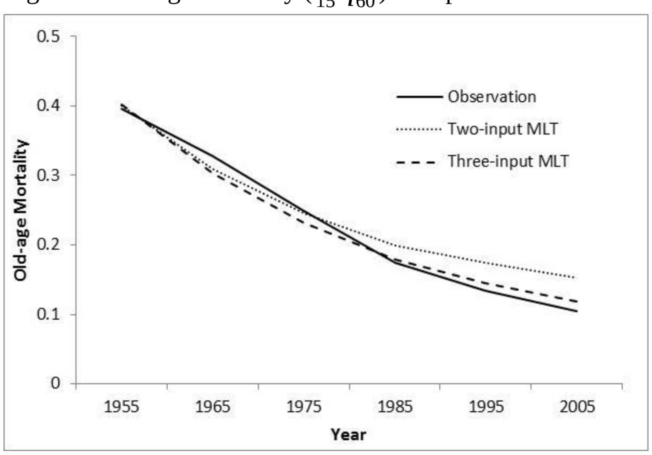
Figure 1. Root-mean-squared errors in predicting ${}_{15}q_{60}$ for the 74 populations by sex in HMD




We see that the three-input MLT improved the performance of the two-input MLT for most of the populations. To be more specific, the three-input MLT improved the performance of the two-input MLT for 55 of the 74 populations. We also see from figure 1 that the chance for the improvement to occur is bigger when the $RMSE_2$ is larger. Since the two-input MLT is based on the data of HMD of which the populations are almost exclusively of European origin, we expect that for non-European-origin populations the error of two-input MLT are more likely to be larger and therefore improvements are more likely to occur. This expectation turned to be true within the HMD populations. The errors are reduced by 17% for Chile, 48% for Japan, and 17% for Taiwan, which are the three non-European-origin populations in HMD. Furthermore, since developing countries are all non-European origin, we expect that the three-input should provide greater improvements than that in the validations.

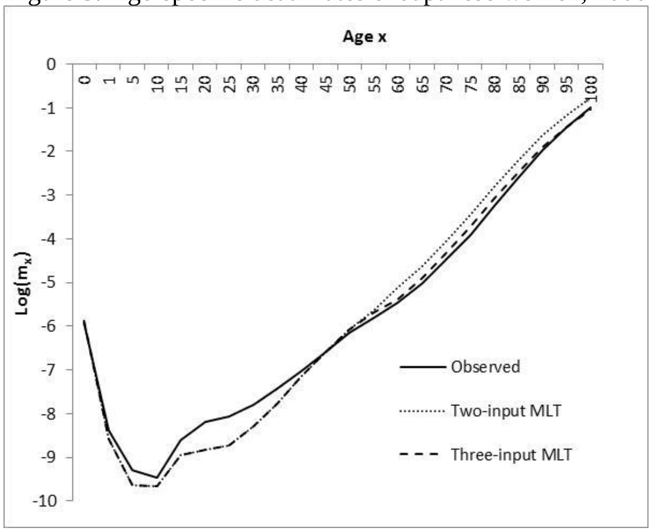
To see more details of the improvement, we choose Japanese women as an example, and show the fittings of old-age mortality in figure 2. We see that the three-input MLT performed slightly worse than did the two-input MLT for years before 1980, but remarkably better later. Overall, the three-input MLT reduced the errors of the two-input MLT by 49% (48% for both men and women).

Figure 2. Old-age mortality (${}_{15}q_{60}$) of Japanese women




Our final target is not only to better fit ${}_{15}q_{60}$, but to improve the estimates of life tables at old ages. To see how this target is reached, we choose Japanese women in 2000-2009 as an example, and show the result in figure 3. We see that the three-input MLT remarkably improved the estimates of age-specific death rate at old ages.

Figure 3. Age-specific death rates of Japanese women, 2000-2009




## 4Summary

In 2010-2015, for example, the deaths at age 60 and older already reached 60% of all deaths worldwide (UNPD, 2015). Compared to the numbers of deaths at child and adult ages, the number of deaths at old ages is the biggest and, ironically, also the least reliable. This is because, for most developing countries, the numbers of old-age deaths are not estimated on the basis of empirical data. They are extrapolations of mortality at younger ages. This reality indicates that improving the estimates of old-age mortality for individual developing countries is not enough, and that establishing a life-table database for all developing countries, which utilizes the improved estimations of old-age mortality, is necessary.

At old ages, migrants are rare comparing to deaths. Thus, census data on population by age and sex could be used to estimate old-age mortality; and such data are available for almost all the countries of the world. For example, among the 233 countries and areas (UNPD, 2015), 220 have conducted the 2010-round census between 2005 and 2014 (United Nations Statistics Division, 2013b). Moreover, some developing countries had surveys or censuses that collected information on old-age mortality, which can be used as supplementary data to more reliably estimate old-age mortality.

In recent years, new methodological developments have been made to use census population to estimate old-age mortality, and extend one-input model life tables to better utilize existing information. Furthermore, these methods are improved to work better for old ages in recent years. In this paper, we described and organized these methods as the three-input MLT; and we validated performance of the three-input MLT using the HMD data. We found that the three-input MLT could improve the performance of the previous methods for 55 of the 74 populations in HMD, and that the average improvement is 14%. To be more relevant to developing countries that are non-European-origin populations, confirm this suggestion, improvements are observed for all the non-European-origin populations in HMD, which are 17% for Chile, 48% for Japan, and 17% for Taiwan.

This paper indicated that establishing a life-table database for developing countries is necessary, that the methodology is adequate, and that the empirical data are available.

## References


Bennett, N.G.and S. Horiuchi.1981. Estimating the Completeness of Death Registration in a Closed Population.*Population Index* 47(2):207-221.

Haman Mortality Database, 2016. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de.

Li, N., 2014. Estimating Life Tables for Developing Countries. Technical paper of the United Nations Population Devision 2014/4. Available at http://www.un.org/en/development/desa/population/publications/pdf/technical/TP2014-4.pdf.

Li, N. and P. Gerland. 2013. Using census data to estimate old-age mortality for developing countries. Paper prepared for Session 17-05: Indirect methods of mortality and fertility estimation: new techniques for new realities, in *XXVII IUSSP International Population Conference*. Busan, Korea. http://iussp.org/en/event/17/programme/paper/1959

Preston, S. H., P. Heuveline and M. Guillot (2001). *Demography, Measuring and Modeling Population Process*. Oxford, United Kingdom: Blackwell Publishers Ltd.

Timæus I.M. 2013. Estimation of adult mortality from sibling histories. In Moultrie T.A., R.E. Dorrington, A.G. Hill, K. Hill, I.M. Timæus and B. Zaba (eds). *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/content/sibling-histories

United Nations Population Division. 2015. *World Population Prospects: The 2015 Revision*. United Nations, New York.

United NationsPopulation Division. 1982. *Model life tables for developing countries*. New York: United Nations.

United Nations Statistics Division. 2008.*Principles and Recommendations for Population and Housing Censuses, Revision 2,Statistical papers Series M,* Rev. 2 ed., New York, NY:United Nations. http://unstats.un.org/unsd/demographic/sources/census/census3.htm

United Nations Statistics Division. 2013a. *Demographic Yearbook 2012, Sixty-third issue*. New York: United Nations. https://unstats.un.org/unsd/demographic/products/dyb/2010_round_latest.htm

United Nations Statistics Division. 2013b. Availablehttp://unstats.un.org/unsd/demographic/sources/census/2010_PHC/censusclockmore.htm

Wang H., L. Dwyer-Lindgren, K.T. Lofgren, J.K. Rajaratnam, J.R. Marcus, A. Levin-Rector, C. Levitz, A.D. Lopez, C.J.L. Murray. 2012. Age-specific and sex-specific mortality in 187 countries, 1970–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*. 2012 Dec 13; 380: 2071–2094.

Wilmoth, J., S. Zureick, V. Canudas-Romo, V. Inoue, and C. Sawyer, 2012. A flexible two-dimensional mortality model for use in indirect estimation. *Population Studies*, (66). 1-28.


## Appendix A. Adjusting age reporting errors

It is hard to find a proper basis to adjust enumerating errors in a real population, which is affected by historical fertility, mortality and migration. But a stationary population is determined only by mortality. Thus, it is possible to find a proper basis to adjust age errors for stationary populations. According to the United Nations general model life table (United Nations Population Division, 1982), there is a common relationship between the survival ratios $S_{60} = \frac{L_{65}}{L_{60}}$ and $S_{65} = \frac{L_{70}}{L_{65}}$ among model life tables, which is

$$S_{65} = -0.29 + 1.27 \cdot S_{60}, \quad R^2 = 0.998\,. \qquad \text{(A.1)}$$

This relationship is called the model line. When the observed survival-ratio point, $(S_{60}, S_{65})$, is above the model line, or when the survival ratio is abnormally rising with age, the difference between the survival-ratio point and the model line is caused mainly by age heaping. Accordingly, assuming that the heaping ratio at age 60 equals to that at age 70, the adjustment is

$$\begin{aligned}
\hat{L}_{60} &= L_{60} - \frac{L_{60}}{L_{70}}\Delta, \\
\hat{L}_{65} &= L_{65} + \Delta, \qquad \text{(A.2)} \\
\hat{L}_{70} &= L_{70} - \Delta,
\end{aligned}$$

where

$$\begin{aligned}
\Delta &= \frac{-B + \sqrt{B^2 - 4AC}}{2A}, \\
A &= b - a\frac{L_{60}}{L_{70}} - \frac{L_{60}}{L_{70}}, \\
B &= a\left(L_{60} - \frac{L_{60}}{L_{70}} L_{65}\right) + 2bL_{65} + L_{60} + \frac{L_{60}}{L_{70}} L_{70}, \qquad \text{(A.3)} \\
C &= L_{65}(aL_{60} + bL_{65}) - L_{60}L_{70}, \\
a &= -0.29, \quad b = 1.27.
\end{aligned}$$

On the other hand, when the survival-ratio point is below the model line, the difference between the survival-ratio point and the model line is caused by nonspecific errors. Accordingly, the adjustment is to move the survival ratio point into the model line through minimal distance as

$$\begin{aligned}
\hat{S}_{60} &= \frac{-ab + S_{60} + bS_{65}}{1 + b^2}, \qquad \text{(A.4)} \\
\hat{S}_{65} &= a + b\hat{S}_{60}.
\end{aligned}$$

$$\begin{aligned}
\hat{L}_{60} &= w\frac{L_{60} + \hat{S}_{60}L_{65} + \hat{S}_{60}\hat{S}_{65}L_{70}}{1 + \hat{S}_{60}^2 + \hat{S}_{60}^2\hat{S}_{65}^2} + (1-w)L_{60}, \\
\hat{L}_{65} &= w\hat{S}_{60}\frac{L_{60} + \hat{S}_{60}L_{65} + \hat{S}_{60}\hat{S}_{65}L_{70}}{1 + \hat{S}_{60}^2 + \hat{S}_{60}^2\hat{S}_{65}^2} + (1-w)L_{65}, \qquad \text{(A.5)} \\
\hat{L}_{70} &= w\hat{S}_{65}\hat{S}_{60}\frac{L_{60} + \hat{S}_{60}L_{65} + \hat{S}_{60}\hat{S}_{65}L_{70}}{1 + \hat{S}_{60}^2 + \hat{S}_{60}^2\hat{S}_{65}^2} + (1-w)L_{70},
\end{aligned}$$

where $0 \le w \le 1$ is the weight, and is used as 0.5.

### *Appendix B. The errors of estimating survival ratio using census population*

Let the net undercounting rates be $u_1$ and $u_2$ for the first and second censuses, respectively. Neglecting intercensal migration, the estimated survival ratio (*Se*) is:

$$Se = \frac{p(70-74, t_2)\cdot(1-u_2)}{p(60-64, t_1)\cdot(1-u_1)} = S\,\frac{(1-u_2)}{(1-u_1)}\,. \qquad \text{(b.1)}$$

Subsequently, the relative error in estimating survival ratio is:

$$E(u_1, u_2) = \frac{Se - S}{S} = \frac{1-u_2}{1-u_1} - 1 = \frac{u_1 - u_2}{1-u_1}. \qquad \text{(b.2)}$$

It can be seen that the estimating error of survival ratio is determined only by census undercounts. In addition, census undercounts tend to cancel each other in causing the errors of estimating survival ratio, which would therefore be small in general.