## Chapter - 2

# General demographic inputs for the estimation of disease burden. Mortality in Andhra Pradesh, India.

*Prasanta Mahapatra*

This chapter describes the general demographic inputs required for estimation of National burden of disease (NBD) and also explains the process by which I have arrived at these inputs for Andhra Pradesh. The intermediate steps required to generate demographic inputs for NBD will naturally vary according to specific circumstances of the country. For example, in a country or state with full coverage of vital registration system, there will not be any need to estimate the degree of under - registration. Since the coverage of vital registration system in India is not complete, the intermediate steps described here will be similar for many developing countries, with a similar vital statistics reporting system. The GBD estimates (World Bank, 1993; Murray and Lopez, 1996) are for the year 1990. I have used 1991 as the reference year to generate the demographic inputs, mainly because there was a population census in 1991 and the year is quite close to the GBD reference year of 1990.

The chapter is organised as follows. An overview of the demographic inputs required for estimation of disease burden is presented. Population data sources and estimates are then described. The next three sections deal with local mortality experience. Indirect demographic estimations and reference to stable population models are often resorted to in situations of partial or poor data availability. The first section on local mortality experience titled "identification of stable population models" examines stable characteristics of the state's population. It describes an approach to identify closely resembling stable population models for the state. This is followed by a description of mortality data sources in Andhra Pradesh and an analysis of the completeness of registration of deaths. The next section describes local life table functions. Usual life table functions are presented for a conventional appreciation of the current mortality experience in the state. Life table functions and mortality estimates required as inputs for computation of DALYs are also identified. Measurement of disease burden using a health gap measure like the DALY, define a standard life - table with respect to which the gap is to be measured. Murray (1994, 1996) chose a model from the extensions to Coale and Demeny model life tables (Coale and Guo, 1989). These models give life - table functions in five year age groups.

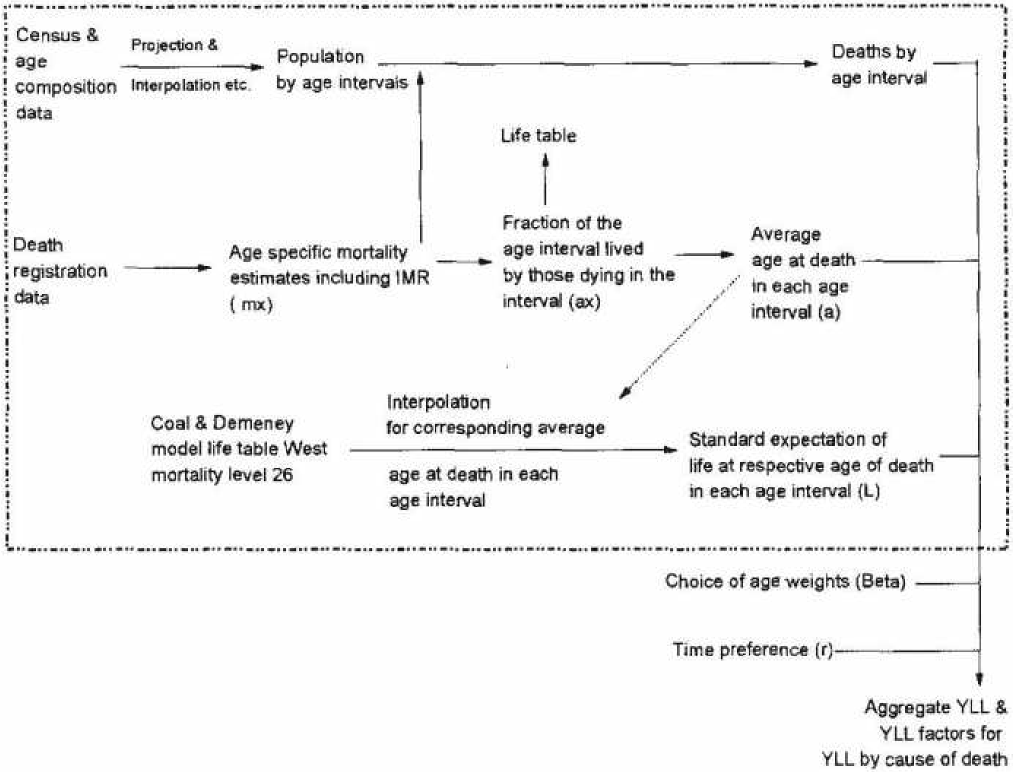
Computation of DALYs often require interpolated figures of standard life expectancy. NBD teams may adopt different techniques of interpolation, giving rise to slight variations in the estimates. To avoid this source of variation, the section titled "standard life expectancies by single year" compares different interpolation techniques. The chapter concludes with a list of demographic estimates required for computation of DALYs and corresponding appendices containing the estimates for Andhra Pradesh.

## Overview of demographic inputs for estimation of disease burden:

For the mortality (YLL) component of DALY the local data required must comprise the number of deaths by age, sex and average age at death within respective age groups. For the disability (YLD) component of DALY, no demographic information is directly required. The age at death can be approximated by identifying the midpoint of respective age intervals. YLDs can be computed from number of incident cases, and other descriptive epidemiology parameter estimates like age at death, duration, etc. Thus the only indispensable set of demographic estimates required for computation of DALYs appear to be the deaths. It is customary to ponder as to why we need more elaborate demographic estimates? For example, a frequently asked question is, what is the role of local life table in burden of disease estimation. It will be useful to take stock of the demographic estimates that may be required at different stages of disease burden. We are essentially talking about the role of a demographer in a NBD team.

Let us start with the single most important local data on mortality experience, namely, the number of deaths. This serves two purposes. The number of deaths is an input for computation of YLL. In addition, the general demographic estimate of deaths is an envelope for reconciliation of cause - specific mortality. Deaths due to all causes cannot exceed deaths due to all causes in respective age groups. Since the mortality estimates form a very important input for computation of YLL, all efforts should be made to improve accuracy of these estimates. Hence an assessment of the completeness of registration of deaths is necessary. Life - table functions are computed to have a conventional appreciation of the local mortality levels. Implausible life table functions might raise a suspicion about accuracy of the mortality data. For large age intervals, the midpoint is a poor estimate of the age at death in that interval. The a(x) function from the life table gives a better estimate of the average age at death in different age groups. Note that the life expectancy at the average age at death (which is the years of life lost, in the absence of age preference and discounting) is taken from the standard life table chosen to define the health gap measure. Thus the age at death value is arrived at from the local life table and the unrealised life expectancy at the age at death is arrived at from the standard life table.

Figure-2.1 The general demographic estimation process for computation of DALYs.




Although population data is not a direct input for computation of DALYs, it is required for intermediate computations and presentation of results. If mortality statistics are based on sample registration or are available in the form of rates, then population figures are required to estimate deaths. Population figures are required for estimation of epidemiological parameters. For example, the DISMOD (Murray and Lopez, 1996, p204-209 ) program for checking consistency of epidemiological estimates requires population at least by five year age groups and preferably in single year intervals. Figure-2.1 explains the steps and interrelationships of various demographic estimates for computation of DALYs. A demographer would pursue three threads of data collection and estimation, namely (a) population, (b) mortality, and (c) handling of the standard life - table to give interpolated values of standard life expectancy. Interpolation and projection techniques may be used to estimate population figures according to the required age groups. Mortality data need to be validated, by checking for under - registration. The adjusted age- specific death rates are then applied to the population figures to estimate the number of deaths. Local life - table functions are computed from estimates of age specific death rate. Life table functions can be used to arrive at average age at death in respective age groups. This is an input for YLL computation. Standard life- tables are usually defined for five year age groups. The NBD team demographer has to either use some interpolation technique to prepare a table of $e_x$ values by single year or obtain such a table from other researchers.

Finally, a brief on choice of different age groups. Murray and Lopez (1994, 1996) use different age groups for the YLL and YLD component estimates. YLL estimates are made in seven age groups. YLD component is estimated in five age groups. These are collapsed to conform to the least common denominator of the two age groups, namely the five age groups for which YLD estimates are made. The rationale is that, given the general paucity of descriptive epidemiological data, it will be futile to try and estimate epidemiological parameters for a large number of age intervals. Hence the number of age groups for the YLD component is kept at five to reflect infancy and childhood, adolescence, adulthood, older adults and the elderly. For the YLL component, more accurate estimates of mortality experience are usually available for a wider range of age groups. Hence the YLL estimates are made in seven age groups. The adult age group of 15-44 is split up into two to reflect differences in mortality experience of young adults and middle aged persons. The elderly age group of 60+ is split up into 60-69 and 70+ years. Premature mortality contributes roughly about two thirds of total DALY loss. Using seven age groups would improve the accuracy in estimation of years of life lost (YLL) due to premature mortality. The disability adjusted life year (DALY) estimates are presented in five age groups (0-4, 5-14, 15-44, 45-59, and 60+). For consistency check of incidence prevalence estimates in DISMOD, at least 5 year age groups with a further split -up of the first group into 0 and 1-4 years is used. Accuracy of DISMOD estimates would improve if population data are available by single year age intervals. This means that general demographic estimates are required to be presented in many age groups. My experience is that, people sometimes confuse between the age groups. I define the following age group labels to help achieve clarity in communication among NBD team members:

**Five age groups**, i.e., 0-4, 5-14, 15-44, 45-59 and 60+. YLD and final DALY estimates are presented for these age groups.

**Seven age groups**, i.e., 0-4, 5-14, 15-29, 30-44, 45-59, 60-69 and 70+. The years lost due to premature mortality (YLL) estimates are presented in this manner.

**Five year age groups** i.e., 0, 1-4, 5-9, 10-15, .... Though this is called five year age groups, data for first five years is usually given for 0-1 and 1-4 age groups. Demographic data on population and mortality are usually collected in five year age groups. Data in these age groups can be used by DISMOD, although single year data is preferable.

**Single year age groups**: When available, improves accuracy of DISMOD computations.

## Population data for burden of disease study:

Population census is the natural source of population data. But a population census usually takes place at long and fixed intervals of about 10 years. If the reference year of the NBD study coincides with the census year, as is the case here, then the census figures could be used straightaway. However, there is usually a gap in publication of detailed age specific counts. For example tabulation of population by five - year age groups from the 1991 Indian census was not available until 1998. Primary census abstracts giving the total population count and its break up into two age groups were available by 1992 (Registrar General, 1992). In such a situation the task is to arrive at age - specific distribution of the population. For this an independent source of age distribution and a technique for interpolation may be required. In the Indian context, the Sample Registration Scheme (SRS) provides an estimate of age composition on an annual basis. Sampling design, registration and validation methodology of the SRS have been published (Registrar General 1993a). Several options for interpolation of population data are available. The osculatory interpolation technique using Sprague's coefficients is commonly used for interpolation of population and deaths. Sprague's coefficients and osculatory interpolation procedure is described in Shryock and Siegel (1976 p555). For Andhra Pradesh, detailed tables from 1991 census are now available giving the distribution of population by single year and five year age intervals. Hence I have used the population data from the census directly. The state's population tabulated by the five age groups, seven age groups and five year age groups are given in Appendix 2.1.

## Identification of stable population models:

In situations where mortality statistics are incomplete or their accuracy is in doubt, indirect estimation techniques are used. Many indirect estimation techniques make use of the relationship of demographic parameters in stable populations. Tables of demographic functions for different models of stable population with varying levels of mortality and population growth are available. For example, the Coale and Demeny model life-tables (Coale and Demeny, 1983). It's a good idea to start demographic estimations for national burden of disease studies, by locating a stable population model to which the local population has closest resemblance. Once a stable population model is identified, its demographic parameters become the first approximation of mortality estimates for computation of burden of disease.

Table - 2.1 Mortality changes in Andhra Pradesh

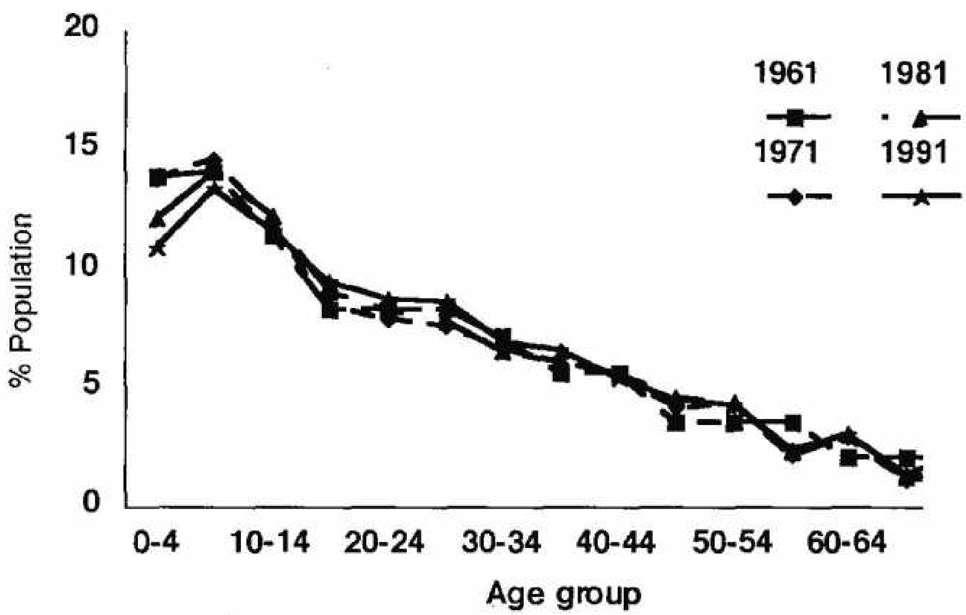
| Period | Crude death rate | Method | Source |
|---|---|---|---|
| 1961-71 | 18.7-19.9 | Intercensal survival | Bhatt, |
| 1970-73 | 17.5 | SRS adjusted for 10% | Preston, and |
| 1974-76 | 16.6 | under registration. | Dyson (1984) |
| 1977-79 | 14.9 | | |
| 1991 | 10.07 | My estimate based on SRS, adjusted for under registration, described later in this chapter. | |

The basic premise of stable population theory is that under constant fertility and mortality experienced for a long period by a closed population, its age composition remains stable (Lotka, 1907, 1922; Dublin and Lotka, 1925). The requirement of constant mortality and fertility can be relaxed, if changes in mortality and fertility are gradual (Shryock and Siegel, 1976 Ch-24). Thus two characteristics of the local population determine the degree of reliance we attach to indirect estimations that assume the local population to be stable. These are: (a) mortality trends, and (b) fertility trends. In case of Andhra Pradesh, the state, as in the rest of India, is also experiencing declining death rates. Table-2.1 shows estimates of crude death rate for different periods and Table - 2.2 shows the inter-censal growth rates for rural and urban areas of the state.

Table-2.2 Exponential intercensal growth factors for urban (U) and rural (R) population of AP

| Period | R Males | R Females | U Males | U Females | All Males | All Females |
|---|---|---|---|---|---|---|
| 1961-71 | 0.017 | 0.016 | 0.029 | 0.029 | 0.019 | 0.019 |
| 1971-81 | 0.016 | 0.016 | 0.04 | 0.04 | 0.021 | 0.021 |
| 1981-91 | 0.017 | 0.017 | 0.035 | 0.037 | 0.022 | 0.022 |

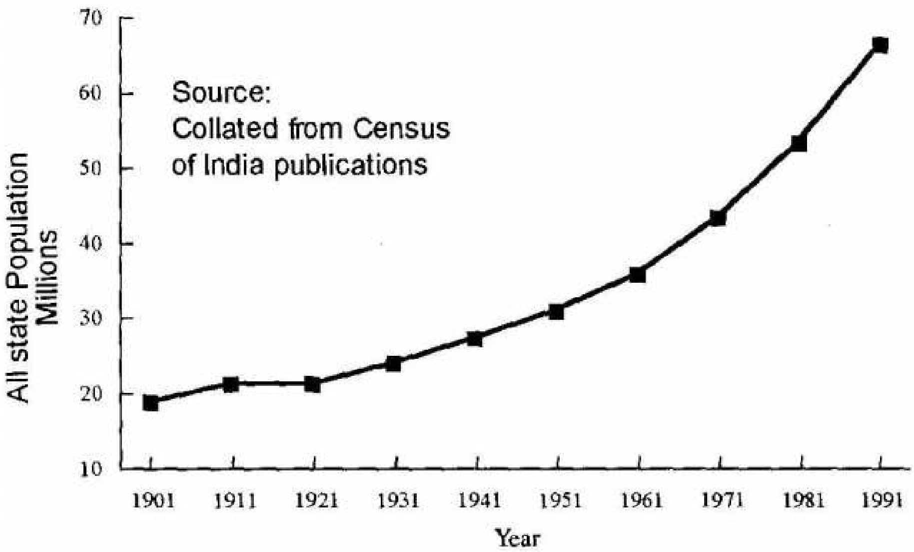
Figure 2.2: Andhra Pradesh - age composition of population over time 1961 to 1991, Census




[1] Appendix 2.5 shows data used for this graph and also rural / urban age composition in censuses from 1961

The state has been experiencing a gradual decline in mortality and a gradual decrease in fertility. Figure 2.2 shows age composition of the state's population at the last four decenial censuses. All age distribution plots follow the same pattern and are close to each other. Although population has been continuously rising (Figure 2.3), the age composition of the state's population has remained fairly stable over the last five decades. Net migration from the state to other states and other countries is negligible. However the rural urban migration within the state is substantial. Fertility in urban areas is known to be slightly less than that of rural areas. Thus the higher growth of urban population must have been due to rural urban migration.

Figure 2.3 Andhra Pradesh Population trend in the 20th century.




If reasonably accurate data on age composition of the local population is available, we can identify a closely related stable population model by comparing the age composition of local population with age composition of all stable populations. We can narrow down the comparison to a smaller set of stable population models, if we have some idea about the local population's mortality level and probable growth rates. For Andhra Pradesh, age composition of different sets of the study population was compared with the age composition of Coale and Demeny stable population models from both West and South patterns of mortality with mortality levels ranging from 14 to 20 and growth rates ranging from 15 to 30. The index of dissimilarity (ID) (Shryock and Siegel, 1976 p131) was used to compare agreement between age patterns. Index of dissimilarity is defined as $ID = \frac{1}{2} \sum |r_{2a} - r_{1a}|$ where $r_{1a}$ = percentage of first population in age group *a* and $r_{2a}$ = percentage of second population at in age group *a*. The summation is conducted over all age groups.

Table-2.3 Index of dissimilarity (ID) of age composition between AP SRS 1991 and selected Coale and Demeny stable populations.

| Population | Model, Level and Rate | ID | Population | Model, Level and Rate | ID |
|---|---|---|---|---|---|
| Rural males | WL14R15 | 3.95 | Rural | WL14R15 | 3.94 |
| | WL15R15 | 4.19 | females | WL15R15 | 4.16 |
| | WL19R20 | 4.25 | | WL17R20 | 4.31 |
| | WL18R20 | 4.28 | | WL18R20 | 4.32 |
| | WL17R20 | 4.35 | | WL19R20 | 4.35 |
| Urban males | WL14R20 | 6.15 | Urban | WL14R20 | 6.21 |
| | WL15R20 | 6.16 | | WL15R20 | 6.31 |
| | WL16R20 | 6.25 | | WL16R20 | 6.45 |
| | WL17R20 | 6.36 | | WL17R20 | 6.61 |
| | WL14R15 | 6.38 | | WL19R25 | 6.76 |
| All state males | WL14R15 | 3.9 | All state females | WL14R15 | 4.19 |
| | WL17R20 | 4.29 | | WL17R20 | 4.43 |
| | WL15R15 | 4.29 | | WL16R20 | 4.46 |
| | WL18R20 | 4.34 | | WL18R20 | 4.52 |
| | WL16R20 | 4.36 | | WL15R15 | 4.52 |

[1]Legend for model acronym: W=West, L=Level, R = Growth rate. Thus WL14R15 would mean model West, mortality level 14 and growth rate 0.015.

Table-2.3 shows the five lowest indexes of dissimilarity for each population subset. All of them belong to the West family. Dissimilarity of age pattern is the least for model West level 14, which indicates a life expectancy at birth of 55 years. This is consistent with the life expectancy estimates reported later in this chapter (Table-2.6). These estimates are: 53 years for rural males, 55 years for rural females, and 56 years for urban persons. The indexes for urban population are consistently higher than that of the rural population. This is mainly due to the rural-urban migration. The size of outmigration in relation to the rural population is smaller compared to the size of the corresponding immigration to urban areas in relation to the urban population. It is of interest to note that the mortality in AP appears to have changed from its similarity with South (as was reported by Bhatt and others, 1984) to the West pattern of Coale & Demeny models.

## Mortality data in Andhra Pradesh and completeness of registration of deaths:

Reference has already been made to SRS, the primary purpose of which is to provide representative data for estimation of mortality. For this study I have taken the SRS estimate of age- specific mortality rates and adjusted them for under registration. Incomplete registration of deaths can affect burden of disease estimates in many ways. Non -differential under- registration across all ages and both sexes would not usually pose a problem for a purely mortality - based analysis of disease burden. Since burden is now sought to be measured combining mortality and morbidity, under- reporting of deaths would bias the disease burden estimate towards morbidity-oriented conditions. In addition, differential under registration by age and or sex would bias both mortality - based and synthetic analyses of disease burden.

### Completeness of adult death registration by SRS:

Completeness of adult death registration by SRS with respect to the SRS population is estimated by Preston and Coale method (United Nations Manual - X, 1983). The overall inter-censal growth rate for males (2.18%) and females (2.15%) was used. To arrive at the coefficients for estimation of age factor for the open interval, it was assumed that mortality in AP conforms to Coale and Demeny West pattern[1]. Completeness of death registration (C) is the proportion registered of actual deaths; which is same as ratio of registered deaths to actual deaths. This proportion is estimated by the ratio of population implied by the age distribution of registered deaths to actual population count in corresponding age group. Thus

$$\hat{C} = \frac{\text{Population estimated from age distribution of deaths}}{\text{Corresponding census count i.e. actual population}}$$

Two ratios are used by the Preston and Coale method to facilitate interpretation of the nature of data and identification of the completeness estimate. These are:

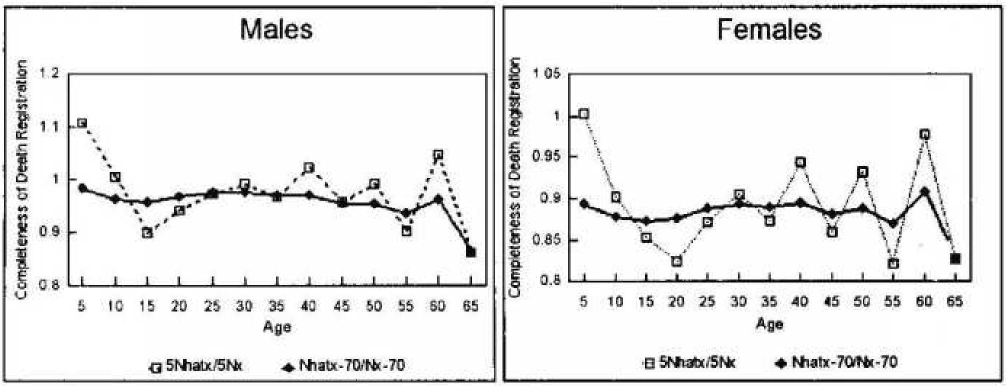
$$\frac{{}_5\hat{N}_x}{{}_5N_x} = \frac{\text{Estimated population in age group x to x+5}}{\text{Actual population in age group x to x+5}} \quad \text{... ... 5Nxhat/5Nx in Figure - 2.4}$$

[1]The set of coefficients for calculation of z (alpha) are separate for each mortality pattern but same for all levels within each pattern. Hence saying West without specifying the level of mortality is enough.

For both ratios the estimated population is obtained by first cumulating the population implied by the age distribution of deaths, the chosen growth rate and mortality pattern. Hence the effect of misreporting of age at death gets smoothed out for the estimates of population derived from it. Thus the first ratio of five year age groups (5Nhatx/5Nx) is sensitive to age misreporting from the census which reflects in the denominator. In case of the second ratio of population upto age 70 (Nhatx-70/Nx-70) age misreporting from the census also get smoothed out. So the ratio of population upto age 70 gives us a more consistent estimate of the extent of completeness of death registration.

Figure-2.4 shows the graphs of the two completeness ratios referred above for both sexes in rural and urban Andhra Pradesh. The ratio of five year age groups shows many ups and downs, more pronounced in females. There appears to be some age heaping around 15-24 which is more pronounced among females. Under-counting of children appears to be prevalent for ages upto 10.

Figure -2.4 Ratio of completeness of death registration by SRS, by age.




The ratio of population upto age 70 shows very good consistency for ages upto 55 years. The median of completeness fractions at each age group is read off to arrive at the adjustment factors (Table-2.4). In case of males the SRS appears to be registering about 96% of deaths. Registration of female deaths is lower at 89%. Bhat et al (1984) had estimated the completeness of SRS death registration around 1981 to be 90%. The present estimate is consistent with their estimate. Registration of male deaths appears to have improved in the mean while and registration of female deaths have reduced marginally.

Table-2.4 Estimated completeness of registration and adjustment factors

| Population sub division | Completeness of registration | | Adjustment factor |
|---|---|---|---|
| | Chosen growth rate | Completeness | |
| Males | 0.0218 | 0.96 | 1.04 |
| Females | 0.0215 | 0.89 | 1.13 |

## Completeness of death registration at younger ages:

For childhood mortality, the SRS estimates were compared with the findings of the National Family and Health Survey (NFHS) 1992, which is the Indian implementation of the Demographic Health Survey (DHS). According to the NFHS-92, IMR () for the period 1988-92 in AP was 70.4 / 1000 live births (Kanitkar et al, 1994). The 1990-92 three year average from SRS gives us slightly higher estimates i.e., 71 / 1000 for males and 73/100 for females. The two estimates are indeed very close to each other. The SRS estimate relates to the last three years (1990-92) of the five year period to which the NFHS estimate relate. Considering that the IMR has been generally declining, though slowly, the SRS estimates should have been slightly less than the NFHS. The NFHS is based on a 3 year recall which might have led to the slight underestimate. Hence I decided to keep the SRS child mortality estimates as such.

A significant number of childhood deaths are concentrated around the first year of life. Further breakdown of deaths in the first five year age interval is required for analytical purposes. Particularly distribution of these deaths between age interval 0 (i.e. upto 1 year) and 1-4 years is used to focus analytical attention on infant mortality. A simple distribution of deaths in 0-5 years equally to each year would not be correct as most of the deaths are concentrated in the first year. Hence the number of deaths in the first year was calculated on the basis of reported birth rate and IMR. The total number of infant deaths was then distributed to each sex by applying the male : female ratio of 105:100. The resultant number of deaths to infants of each sex is used as such for age interval 0 year. These deaths were subtracted from the deaths in age interval 0-4 years to arrive at the number of deaths in the age interval 1-4 years.

## Calculating $M_x$ values (age specific death rates):

SRS death rates for age groups 15-19 years and above were adjusted upwards by applying the factors obtained in Table-2.4 above. The SRS death rates for age group 0-4, 5-9 and 10-14 were used as such. In case of urban males a minor inconsistency in age pattern of the $M_x$ values for age group 40-44 and 45-49 was noticed. The general observation in all populations has been that adult mortality rates rise monotonically with age. This is the case for all subsets of population in Andhra Pradesh except for data relating to two age groups of urban males. In case of the later $M_x$ value for age group 40-44 ( ${}_5M_{40}$ =0.00832) was found to be slightly higher than that for the next higher age group ( ${}_5M_{45}$ =0.00686). The general pattern of increasing mortality rates with rising age is maintained for this group except for this isolated discrepancy. This could be due to a preference for age 40. To restore the monotonic pattern of $M_x$ values I have manually switched them for these two age groups of urban males. This does not affect other life table functions like life expectancy etc. Its only effect is that the mortality rate for age group 45-49 is higher than that for age 40-44, which is plausible. The $M_x$ values for infants and 1-4 year children described earlier and the adjusted age specific mortality rates for adults were combined to obtain a complete set of values for construction of abridged life table. The resultant estimates of $M_x$ values for each region are plotted in Figure 2.5 and are shown in Table-2.5.

Figure 2.5: Age pattern of $M_x$ values for different sub populations in AP, 1991

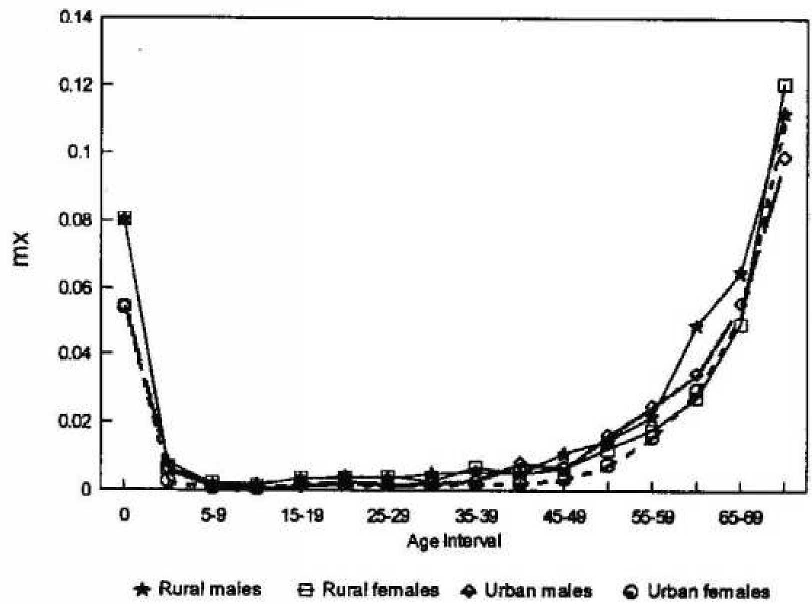


Table-2.5 Estimated $M_x$ values in 1991 for different regions of Andhra Pradesh

| Age interval | Rural | | Urban | | All State | |
|---|---|---|---|---|---|---|
| | Male | Female | Male | Female | Male | Female |
| 0 | 0.08053 | 0.0804 | 0.05426 | 0.05423 | 0.07499 | 0.07711 |
| 1-4 | 0.008 | 0.00629 | 0.00632 | 0.00275 | 0.00784 | 0.00511 |
| 5-9 | 0.0018 | 0.002 | 0.0011 | 0.0006 | 0.0017 | 0.0017 |
| 10-14 | 0.0018 | 0.0014 | 0.0008 | 0.0003 | 0.0016 | 0.0011 |
| 15-19 | 0.00291 | 0.00339 | 0.00104 | 0.00203 | 0.0025 | 0.00305 |
| 20-24 | 0.00416 | 0.00362 | 0.0026 | 0.00158 | 0.00385 | 0.00316 |
| 25-29 | 0.00343 | 0.00407 | 0.00177 | 0.00181 | 0.00302 | 0.0035 |
| 30-34 | 0.00478 | 0.00249 | 0.00229 | 0.00158 | 0.00416 | 0.00226 |
| 35-39 | 0.00541 | 0.00655 | 0.00302 | 0.00215 | 0.00489 | 0.00565 |
| 40-44 | 0.0051 | 0.00509 | 0.00686 | 0.00215 | 0.00572 | 0.00441 |
| 45-49 | 0.01102 | 0.00644 | 0.00832 | 0.00339 | 0.01019 | 0.00588 |
| 50-54 | 0.01466 | 0.01232 | 0.01633 | 0.00746 | 0.01498 | 0.0113 |
| 55-59 | 0.02205 | 0.01763 | 0.02506 | 0.01571 | 0.02257 | 0.01729 |
| 60-64 | 0.04909 | 0.02723 | 0.03463 | 0.02983 | 0.0467 | 0.02769 |
| 65-69 | 0.06479 | 0.04938 | 0.05585 | 0.05209 | 0.06344 | 0.04983 |
| 70+ | 0.11201 | 0.12057 | 0.09922 | 0.11085 | 0.11014 | 0.11899 |

## Local Life Tables:

Table-2.6 Selected life table functions for different sub populations in AP

| Region and sex | | ${}_1q_0$ | ${}_4q_1$ | ${}_5q_1$ | ${}_{45}q_{15}$ | $e_0$ | $e_{10}$ |
|---|---|---|---|---|---|---|---|
| Rural | Male | 0.076 | 0.0314 | 0.105 | 0.3083 | 55.94 | 52.9 |
| | Female | 0.076 | 0.0248 | 0.0989 | 0.2655 | 57.75 | 54.57 |
| Urban | Male | 0.052 | 0.0249 | 0.0756 | 0.2864 | 60.38 | 55.56 |
| | Female | 0.052 | 0.0109 | 0.0624 | 0.1729 | 63.81 | 55.56 |

The $M_x$ values in Table - 2.5 were used to generate life tables for different regions of AP. The Mortpak Lite Lifetab procedure was used (United Nations, 1988). The procedure chooses a set of $a_i$ values (fraction of the age interval lived by those dying in the interval) based on the $m_x$ inputs to it. These $a_i$ values are required to arrive at the average age at death in each age interval, which is a required input for estimation of years of life lost (YLL). Four sets of life tables corresponding to the rural and urban population of males and females respectively were constructed. Table - 2.6 shows selected life table functions for each set of population. All of the above life table functions can be read straightaway from the Mortpak Lifetab output, except for the ${}_{45}q_{15}$ values which is obtained from the ${}_{5}q_{x}$ values as follows:

$$ {}_{45}q_{15} = 1 - (1-{}_5q_{15})\,(1-{}_5q_{20})\,(1-{}_5q_{25})\,(1-{}_5q_{30})\,(1-{}_5q_{35})\,(1-{}_5q_{40})\,(1-{}_5q_{45})\,(1-{}_5q_{50})\,(1-{}_5q_{55}) $$

## Final estimate of deaths:

The final estimate of deaths were made by applying the $M_x$ values arrived at above to the estimated age specific population in 1991. Thus our estimate is that in 1991 there were about 556000 deaths in rural areas consisting of 297000 males deaths and 259000 female deaths. In urban areas there were 129000 deaths in 1991 consisting of 70000 males deaths and 59000 females. The overall crude death rate of the state in 1991 is estimated at 10.07 /1000 population. For rural areas, the crude death rate works out to 11.16 / 1000 population. For urban areas the crude death rate was 7.2 / 1000 population. Appendix 2.2. shows detailed estimate of deaths in various age groups.

## Age at Death *(a)*:

Age at death is a required input for calculation of YLL. Age at death (*a*) in each age interval is arrived at as follows.

$a$ = Initial age of the interval + $A_{x-(x+n)}$

Where $A_{x-(x+n)}$ = Average number of years lived in the age interval by those dying in the interval *x-(x+n)*. This is equivalent to $a_i \times n$, n being the number of years in the age interval. I have used the average years lived in the interval ($A_{x-(x+n)}$) generated by the Mortpak Lifetab procedure. These ($A_{x-(x+n)}$) values can be used straightaway for the 5 year age groups as they correspond to the Mortpak lifetab output age groups exactly. The $A_{x-(x+n)}$ age intervals in the Five and Seven age groups spanning more than five years were calculated from the Mortpak life table output as follows:

$$A_{0-4} = \frac{{}_{1}q_{0}A_{0} + (1-{}_{1}q_{0})\,{}_{4}q_{1}(1+A_{1-4})}{1-(1-{}_{1}q_{0})(1-{}_{4}q_{1})}$$

$$A_{5-14} = \frac{{}_{5}q_{5}A_{5-9} + (1-{}_{5}q_{5})\,{}_{5}q_{10}(5+A_{10-14})}{1-(1-{}_{5}q_{5})(1-{}_{5}q_{10})}$$

$$A_{15-29} = \frac{{}_{5}q_{15}A_{15-19} + (1-{}_{5}q_{15})\,{}_{5}q_{20}(5+A_{20-24}) + (1-{}_{5}q_{20})\,{}_{5}q_{25}(10+A_{25-29})}{1-(1-{}_{5}q_{15})(1-{}_{5}q_{20})(1-{}_{5}q_{25})}$$

$$A_{30-44} = \frac{{}_{5}q_{30}A_{30-34} + (1-{}_{5}q_{30})\,{}_{5}q_{35}(5+A_{35-39}) + (1-{}_{5}q_{35})\,{}_{5}q_{40}(10+A_{40-44})}{1-(1-{}_{5}q_{30})(1-{}_{5}q_{35})(1-{}_{5}q_{40})}$$

$$A_{15-44} = \frac{\begin{array}{c}{}_{5}q_{15}A_{15-19} + (1-{}_{5}q_{15})\,{}_{5}q_{20}(5+A_{20-24}) + (1-{}_{5}q_{20})\,{}_{5}q_{25}(10+A_{25-29}) \\ + (1-{}_{5}q_{25})\,{}_{5}q_{30}(5+A_{30-34}) + (1-{}_{5}q_{30})\,{}_{5}q_{35}(20+A_{35-39}) \\ + (1-{}_{5}q_{35})\,{}_{5}q_{40}(25+A_{40-44})\end{array}}{1-(1-{}_{5}q_{15})(1-{}_{5}q_{20})(1-{}_{5}q_{25})(1-{}_{5}q_{30})(1-{}_{5}q_{35})(1-{}_{5}q_{40})}$$

$$A_{45-59} = \frac{{}_{5}q_{45}A_{45-49} + (1-{}_{5}q_{45})\,{}_{5}q_{50}(5+A_{50-54}) + (1-{}_{5}q_{50})\,{}_{5}q_{55}(10+A_{55-59})}{1-(1-{}_{5}q_{45})(1-{}_{5}q_{50})(1-{}_{5}q_{55})}$$

$$A_{60-69} = \frac{{}_{5}q_{60}A_{60-64} + (1-{}_{5}q_{60})\,{}_{5}q_{65}(5+A_{65-69})}{1-(1-{}_{5}q_{60})(1-{}_{5}q_{65})}$$

For the open age intervals i.e. 60+ and 70+ the average duration lived in the interval is simply the inverse of the $m_x$ value as follows:

$$A_{60+} = \frac{1}{{}_{\omega}m_{60}} \text{ and } A_{70+} = \frac{1}{{}_{\omega}m_{70}}$$

Note that even if the $(A_{x-(x+5)})$ values (i.e. $A_x$ values for 5 year age interval) of all groups contained in a broader group is exactly half the size of each interval (i.e., 2.5 years), the $(A_{x-(x+n)})$ for the broader age group need not be exactly half its size. The $(A_{x-(x+n)})$ for the broader age group is determined by the mortality pattern (i.e, $q_x$ values) within the broader age group. If mortality is increases within the broader group the $(A_{x-(x+n)})$ values will be more than half the interval's size. It will be less than half its size if mortality decreases within the broader interval. For more details regarding the linkages between mortality pattern and $a_i$ values refer to Chiang (1984). $A_{x-(x+n)}$ values for different age groups have been computed for AP from the respective life table functions. These are shown in Appendix 2-3.

## Standard life expectancies by single year:

Calculation of the years lost due to premature mortality (YLL) component of the disability adjusted life year (DALY) measure requires standard expectation of life at respective ages at death. The formula for YLL is

$$YLL = -\left[\frac{Ce^{-\beta a}}{(\beta + r)}[e^{-(\beta+r)L}]\,(1 + (\beta + r)\,(L + a)) - (1 + (\beta + r)a)\right]$$

where L= Standard expectation of life at age *a*, and all other notations are the same as in Murray's original articles (1994, 1996). Since disability weight (*D*) for premature mortality is one, it has been dropped from the DALY formula to give the YLL formula above. In the above formula, age at death *a* is determined by the local mortality experience. The ages at death within five year intervals will thus vary from region to region. Even if some uniform age at death were to be used, these would be the midpoints of five year age intervals. The Coale and Guo (1989) standard life tables are published by five year age groups. To read off standard life expectancies at mid point of various age intervals, a complete life table by single year age group is required. Since the age at death number can have fractions of a year, a complete life-table by fractions of a year at least upto one decimal place would be ideal. In practice, however, we first compute a complete life table by single year age groups. Standard life expectancy at an age of death involving fraction of a year is then obtained by linear interpolation using the life expectancy at the integer part of the age at death and the next age interval. In this section I have described the construction of complete lifetables from the Coale and Guo abridged lifetables using different interpolation methods, and then proceed to compare their results. Using a least squares criterion of fit, I have chosen the piece-wise exponential method of interpolation.

Murray (1994) has chosen the Coale and Guo model west level 26 for females as the standard for females who have life expectancy at birth equal to 82.5 years. For males he has chosen a life expectancy of 80 years. Since none of the Coale and Guo model tables for males have exactly 80 years of life expectancy at birth he uses the female table at level 25 as the standard for males. So the interpolation exercise below is limited to these two tables. The following three methods are explored:

1. Piece-wise exponential survivorship functions within five year age intervals,
2. Six point Lagrangian interpolation for ages up to 0-74 and Gompertz curve for older ages,
3. Mortpak UNABR, which uses the eight point Helligman and Pollard formula.

The method for six point Lagrangian interpolation for ages 0-74 and Gompertz curve there after has been described by Elandt-Jhonson and Jhonson (1980, p111-114). The Mortpak UNABR procedure is described in the Mortpak manual (United Nations, 1988) and Helligman and Pollard (1980). I have used the Mortpak UNABR procedure to generate results by this method. This procedure accepts $q_x$ data only upto 80-85 year age interval. Discussions about the exponential survival distribution can be found in any text on survival analysis (for example Elandt-Jhonson and Jhonson, 1980). It is essentially assumed that the hazard rate (instantaneous probability of death) remains constant over time. This is unrealistic for the total human mortality experience. Since the abridged life- table already captures much of the age pattern of mortality, modelling a constant hazard rate within the five year age groups is usually a reasonable approximation. Further details of implementation of this method in this instant case is given below. In addition to the above three methods the interpolated life expectancy by single years used by Murray for the global burden of disease (GBD) estimates was obtained. The one for males was available. So in addition to the above three methods the figures used by Murray are also included in the comparison.

The exponential survivor function within an age interval say from $l_x$ to $l_{x+n}$ of the abridged life- table is given by $l_{x+n} = l_x e^{-t\lambda x}$ where $t$= time in years after x and $\lambda_x$ is the constant hazard rate within the interval (i.e. the "piece"). We estimate $\lambda_x$ using the $l_x$ and $l_{x+n}$ numbers from the abridged life table, using the above formula which can be rearranged as follows:

$$\lambda_x = -\frac{1}{n} \ln \left(\frac{l_{x+n}}{l_x}\right)$$

Once $\lambda_x$ is obtained, the $l_{x+t}$ values can be generated for each year within the interval by using the formula $l_{x+t} = l_x e^{-t\lambda x}$ for t = 1,2...99 and x = 1,5,10...95. In this manner the lx column is constructed for single year intervals using the $l_x$ values from the abridged table.

For the construction of a complete life table, the fraction of the last age interval lived by those dying within that interval $a_x$ is required in addition to the single year $l_x$ values as obtained above. The published Coale and Guo abridged tables do not provide $a_x$ values. However these can be derived from other columns of the table using the following relationship:

$$a_x = \frac{L_x - nl_{x+n}}{nd_x} = \frac{L_x - nl_{x+n}}{n(l_x - l_{x+n})}$$

The $a_x$ values thus obtained were used for each year within the five year age intervals.

The above procedure resulted in a life expectancy at birth of 79.89 against the original of 80 for level 25 and 82.41 against the original of 82.5 years for level 26. There is a discrepancy of 0.11 and 0.09 years respectively. The discrepancy is due to unavailability of $a_x$ values by a single year. To maintain exact correspondence with the standard life table figure, the difference was added back at all ages of the complete life table.

Degree of fit of the different interpolations were measured by computing the squared deviation of interpolated values from the original values at the overlapping age groups (i.e. the age groups like 0,1, 5, 10,...100 for which both original and interpolated numbers are available). The one with the least sum of squared deviation is judged to have provided the best fit. Comparative statement of life expectancies at different ages in the original Coale and Guo table and different interpolation results are given in Appendix 2-6 for both level 25 and 26. The MortPak UNABR over estimates life expectancies. For example UNABR interpolated life expectancy at birth is 80.76 against original of 80 years and 84.21 against the original of 82.5 years for level 25 and 26 respectively. Similar differences persist at other ages. On the other hand the six point Lagrangian method underestimates life expectancy at corresponding ages. This method gives a life expectancy at birth of 77.81 years for the original of 80 years for level 25 and 79.12 years compared to the original of 82.5 for level 26. Appendix 2-6 here shows details of the fitting criterion from different interpolations. The piece-wise exponential method provides the best fit. If the discrepancy of 0.11 and 0.09 years were not added back, the sum of squared deviations for this method would be 0.31 and 0.22 for level 25 and 26 respectively. This would still leave the piece-wise exponential method with the best fit to the original. Lack of fit of the six point Lagrangian and Mortpak method was anticipated. Both these methods assume Gompertzian mortality pattern at older ages. The Gompertz pattern assumes that mortality rate monotonically increases for ages 75 years and beyond. Coale and Guo recognised that this was not borne out by actual experience in very low mortality populations. Consequently they modified the projection method used in the Coale and Demeny regional model life tables. For the latest low mortality models, Coale and Guo incorporated a linearly decreasing hazard rate starting at 80 years. This is based on observed pattern from low mortality populations. In other words, the hazard rate at older ages increase up to 80 years of age and then decreases monotonically till the cohort exhausts itself completely by 110 years (Coale and Guo, 1989).

The interpolated figures currently being used (i.e. Murray's) by the Harvard Burden of Disease Unit (BDU) has fairly good fit. But the accuracy of DALY computations with respect to the adopted standard can be improved further by using the interpolated figures now arrived at by the piece-wise exponential method. Replication of the burden of disease estimation method can be achieved by adopting the interpolated standard life expectancy by single year now arrived at. For convenience of future work, the interpolated standard life expectancy numbers are furnished in Appendix 2-4. Standard expectation of life at a given age can be calculated from this table as follows. The average age at death in each interval for estimation of burden of disease usually has a fractional component (around 0.5 in view of the fact that these age intervals are in 5 years or its multiple). Hence the exact figure for is obtained from the detailed standard life table by linear interpolation according to the following formula:

$$E(a) = E(n) - [a'_n \times (E(n) - E(n+1))]$$

Where: $a$ = average age at death in the interval,

$n$ = Intiger part of $a$,

$a'_n$ = Fraction part of $a$,

$E(n)$ = Expectation of life at age $n$ in the detailed standard life table,

and

$E(n+1)$ = Expectation of life at age $n+1$ in the detailed standard life table.

## Summary of demographic estimates:

Demographic estimates are required for computation of DALY and for intermediate computations as well. Conventional mortality analysis provides the foundation for the YLL component of DALY. The number of deaths from general demographic estimates are an envelope for causes of death estimates. A National Burden of Disease study would usually start with inputs from general demographic estimates. Although a lot of demographic statistics are generated in India, their accuracy has to be examined. Data from different sources have to be collated to meet the requirements of NBD. Interpolation and projection techniques may have to be applied to fill in lacunae in data. The demographer plays an important role in generating these fundamental estimates needed for intermediate and final estimates of disease burden. The following list of demographic estimates prepared for this study and presented in Appendix -2 gives an overview of the nature of demographic estimates required for a NBD study.

| Appendix | Description of tables |
|---|---|
| 2-1 | Population estimates. |
| 2-2 | Age- specific estimate of deaths. |
| 2-3 | Age at death (a) and Expectation of life at age a [E(a)]. |
| 2-4 | Interpolated standard life expectancies by single year obtained from Coale and Guo model life tables using piece-wise exponential method. |

## References

Bhatt Marri, Preston Samuel and Dyson Tim; Vital rates in India 1961-1981, National Academy Press 1984.

Chiang Chin Long; The life table and its applications; Robert E. Kreiger publishing, Malabar, Florida USA, 1984 pp142-147.

Coal Ansley and Guang Guo; 1989 winter; Revised regional model life tables at very low levels of mortality; Population Index 55(44):613-643.

Coale Ansley J., Paul Demeny; Regional model life tables and stable populations, Academic Press New York, 1983.

Dublin Louis I. and Lotka Alfred J. On the True Rate of Natural Increase. Journal of the American Statistical Association. 1925 Sep; 20(150):305-339.

Elandt-Jhonson Regina C. and Jhonson Norman L.; 1980; Survival models and data analysis; Willey, New York, p111-114.

Helligman L. and Pollard JH; 1980 June; The age pattern of mortality; The Journal of the Institute of Actuaries, vol-107, par-1, no.434.

Kanitkar T, Ramesh B.M., Roy T.K., et al.; National Family Health Survey, India 1992-93 Introductory Reprot; International Institute for Population Sciences, Bombay, 1994, pp84-88.

Lotka Alfred J. Relation between birth rates and death rates, Science New Series 26(653): 21-22, July 05, 1907 reproduced in: Keyfitz Nathan. Introduction to the Mathematics of Population. Part I. The Life Table. Reading, MA: Addison-Wesley Publishing Co.

Lotka Alfred J. The stability of the normal age distribution, 1922 reproduced in: Keyfitz Nathan . Introduction to the Mathematics of Population. Part I. The Life Table. Reading, MA: Addison-Wesley Publishing Co.; pp. 101-107.

Murray Christopher J.L. Quantifying the burden of disease: the technical basis for disability-adjusted life years. in: Murray Christopher J.L. and Lopez Alan D., Editors. Global comparative assessments in the health sector: disease burden, expenditures and intervention packages. Collected reprints from the Bulletin of the World Health Organization. Geneva: WHO; 1994.

Murray Christopher J.L. Rethinking DALYs. in: Murray Christopher J.L. and Lopez Alan D., Editors. The global burden of disease. A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Boston: Harvard School of Public Health; 1996.

Murray Christopher J.L. and Lopez Alan D. Global and regional descriptive epidemiology of disability: incidence, prevalence, health expectancies and years lived with disability. in: Murray Christopher J.L. and Lopez Alan D., Editors. The global burden of disease. A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Boston: Harvard School of Public Health; 1996.

Registrar General & Census Commissioner, India, 1966; Census of India 1961 Volume II Andhra Pradesh Part II-C Cultural and Migration tables, p56,73.

Registrar General & Census Commissioner, India, 1980; Census of India 1971 Series 2-Andhra Pradesh Part II-C(ii) Social and cultural tables and fertility table; p137 table C-IV.

Registrar General & Census Commissioner, India, 1986; Census of India 1981 Social and cultural tables Part IV-A series-2 Andhra Pradesh, table C-6, p425-427.

Registrar General & Census Commissioner, India, 1992; Census of India 1991 Final Population Totals: brief analysis of primary census abstracts paper-2 of 1992 p78,106, 210-211.

Registrar General & Census Commissioner, India, 1993; Sample Registration System - Fertility & Mortality Indicators 1991, New Delhi.

Shryock Henry S., Jacob S. Siegel and associates, The methods and materials of demography; Academic Press New York 1976.

United Nations, Manual X Indirect techniques for demographic estimation p130-139, New York 1983.

United Nations; 1988; Mortpak-Lite. The United Nations Software Package for Mortality Measurement. Manual. United Nations, New York.

———o———