

INTRODUCTION

The *Demographic Yearbook* is an international compendium of national demographic statistics, provided by national statistical authorities to the Statistics Division of the United Nations Department of Economic and Social Affairs. The *Yearbook* is part of the set of coordinated and interrelated publications issued by the United Nations and its specialized agencies,¹ designed to supply basic statistical data for such users as demographers, economists, public-health workers and sociologists. Through the co-operation of national statistical services, official demographic statistics are compiled in the *Yearbook*, as available, for more than 230 countries or areas throughout the world.

The *Demographic Yearbook 2004* is the fifty-sixth in a series published by the United Nations since 1948. It contains general tables including a world summary of selected demographic statistics, statistics on the size, distribution and trends in national populations, natality, foetal mortality, infant and maternal mortality, general mortality, nuptiality and divorce. Data are shown by urban/rural residence, as available. In addition, the volume provides Technical Notes, a synoptic table, a subject-matter index, a historical index and a listing of the issues of the *Yearbook* published to date.

The Technical Notes on the Statistical Tables are provided to assist the reader in using the tables. Table A, the synoptic table, provides a glance of the completeness of data coverage of the current *Yearbook*. The subject-matter index is a detailed guide to the current *Yearbook*. The cumulative historical index is a guide on content and coverage of all fifty-six issues, and indicates for each of the topics that have been published, the issues in which they are presented and the years covered. A list of the *Demographic Yearbook* issues, with their corresponding sales number and the special topics featured in each issue are shown on pages iv and v.

Until the 48th issue (1996), each issue consisted of two parts, the general tables and special topic tables, published in the same volume with the regular topics.² Beginning with the 49th issue (1997), the special topic tables were being disseminated on CD-ROMs as supplements to the regular issues. Two CD-ROMs have so far been issued: the *Demographic Yearbook Historical Supplement*, which presents a wide panorama of basic demographic statistics for the period 1948 to 1997, and the *Demographic Yearbook: Natality Statistics*, which contains a series of detailed tables dedicated to natality and covering the period from 1980 to 1998. Three volumes of a new Demographic Yearbook Special Census Topics have now been prepared and are presented at: <http://unstats.un.org/unsd/demographic/products/dyb/default.htm>

Population statistics are not available for all countries or areas, for a variety of reasons. In an effort to provide estimates of mid-year population and of selected vital statistics for all countries and areas, two annexes have been introduced since the 53rd issue of the *Demographic Yearbook*. Annex 1 presents United Nations population estimates for the period 1995-2004 and the second presents the medium variant estimates of crude birth and death rates, infant mortality and total fertility rates, as well as expectation of life at birth over the period 2000-2005. These data are produced by the United Nations Population Division and are published in the *World Population Prospects - The 2004 Revision*.³

Demographic statistics shown in this issue of the *Yearbook* are available online at the *Demographic Yearbook* website <http://unstats.un.org/unsd/demographic/products/dyb/default.htm>. Information about the Statistics Division's data collection and dissemination programme is also available on the same website. Additional information can be made available by contacting the Statistics Division of the United Nations Secretariat, at demostat@un.org.

TECHNICAL NOTES ON THE STATISTICAL TABLES

1. GENERAL REMARKS

1.1 Arrangement of Technical Notes

These Technical Notes are designed to provide the reader with relevant information for using the statistical tables. Information pertaining to the *Yearbook* in general is presented in the sections dealing with geographical aspects, population and vital statistics. In addition, preceding each table are notes describing the variables, remarks on the reliability and limitation of the data, countries and areas covered, and information on the presentation of earlier data. When appropriate, details on computation of rates, ratios or percentages are presented.

1.2 Arrangement of tables

This issue contains general tables only. Since the numbering of the tables does not correspond exactly to those in previous issues, the reader is advised to use the historical index that appears at the end of this book to find the reference to data in earlier issues.

1.3 Source of data

The statistics presented in the *Demographic Yearbook* are national data provided by official statistical authorities unless otherwise indicated. The primary source of data for the *Yearbook* is a set of questionnaires sent annually by the United Nations Statistics Division to over 230 national statistical services and other appropriate government offices. Data reported on these questionnaires are supplemented, to the extent possible, with data taken from official national publications, official websites and through correspondence with national statistical services. In the interest of comparability, rates, ratios and percentages have been calculated by the Statistics Division of the United Nations, except for the life table functions, the total fertility rate, and also crude birth rate and crude death rate for some countries or areas as appropriately noted. The methods used by the Statistics Division to calculate these rates and ratios are described in the Technical Notes for each table. The population figures used for these computations are those pertaining to the corresponding years published in this or previous issues of the *Yearbook*.

In cases when data in this issue of the *Demographic Yearbook* differ from those published in earlier issues or related publications, statistics in this issue may be assumed to reflect revisions to these data received by September 2006.

2. GEOGRAPHICAL ASPECTS

2.1 Coverage

Data are shown for all individual countries or areas that provided the information. Table 3 is the most comprehensive in geographical coverage, presenting data on population and surface area for all countries or areas with a population of at least 50 persons. Not all of these countries or areas appear in subsequent tables. In many cases the data required for a particular table are not available. In general, the more detailed the data required for a table, the fewer the number of countries or areas that can provide them.

In addition, rates and ratios are presented only for countries or areas reporting at least a minimum number of relevant events. The minimums are stated in the Technical Notes to individual tables.

Except for summary data shown for the world and by major areas and regions in tables 1 and 2 and data shown for capital city and cities with a population of 100 000 or more in table 8, all data are presented at the national level. The number of countries shown in each table is provided in table A, the synoptic table.

2.2 Territorial composition

To the extent possible, all data, including time series data, relate to the territory within 2004 boundaries, when the data were requested from the countries or areas. Exceptions are footnoted in individual tables. Additionally, in table 3, recent changes and other relevant clarifications are specified.

Data relating to the People's Republic of China generally do not include those for Taiwan Province except in tables 1 and 2.

2.3 Nomenclature

Because of space limitations, the country or area names listed in the tables are generally the commonly employed short titles in use in the United Nations as of September 2006⁴, the full titles being used only when a short form is not available. The latest version of the *Standard Country or Area Codes for Statistics Use* can be accessed at <http://unstats.un.org/unsd/methods/m49/m49.htm>.

2.3.1 Order of presentation

Countries or areas are listed in English alphabetical order within the following continents: Africa, North America, South America, Asia, Europe and Oceania.

The designations and presentation of the material in this publication were adopted solely for the purpose of providing a convenient geographical basis for the accompanying statistical series. The same qualification applies to all notes and explanations concerning the geographical units for which data are presented.

2.4 Surface area data

Surface area data, shown in tables 1 and 3, represent the total surface area, comprising land area and inland waters (assumed to consist of major rivers and lakes) and excluding only Polar Regions and uninhabited islands. The surface area given is the most recent estimate available. They are presented in square kilometres, a conversion factor of 2.589988 having been applied to surface areas originally reported in square miles.

2.4.1 Comparability over time

Comparability over time in surface area estimates for any given country or area may be affected by changes in the surface area estimation procedures, increases in actual land surface by reclamation, boundary changes, changes in the concept of "land surface area" used or a change in the unit of measurement used. In most cases it was possible to ascertain the reason for a revision; otherwise, the latest figures have generally been accepted as correct and substituted for those previously on file.

2.4.2 International comparability

Lack of international comparability between surface area estimates arises primarily from differences in definition. In particular, there is considerable variation in the treatment of coastal bays, inlets and gulfs, rivers and lakes. International comparability is also impaired by the variation in methods employed to estimate surface area. These range from surveys based on modern scientific methods to conjectures based on diverse types of information. Some estimates are recent while others may not be. Since neither the exact method of determining the surface area nor the precise definition of its composition and time reference is known for all countries or areas, the estimates in table 3 should not be considered strictly comparable from one country or area to another.

3. POPULATION

Population statistics, that is, those pertaining to the size, geographical distribution and demographic characteristics of the population, are presented in a number of tables of the *Demographic Yearbook*.

Data for countries or areas include population census figures, estimates based on results of sample surveys (in the absence of a census), postcensal or intercensal estimates and those derived from continuous population registers. In the present issue of the *Yearbook*, the latest available census figure of the total population of each country or area and mid-year estimates for 2000 and 2004 are presented in table 3. Mid-year estimates of total population for 10 years (1995-2004) are shown in table 5 and mid-year estimates of urban and total population by sex for 10 years (1995-2004) are shown in table 6. The latest available data on population by age, sex and urban/rural residence are given in table 7. The latest available figures on the population of capital cities and of cities or urban agglomerations of 100 000 or more inhabitants are presented in table 8.

Summary estimates of the mid-year population of the world, major areas and regions for selected years and of its age and sex distribution in 2004 are set forth in tables 1 and 2, respectively.

The statistics on total population, population by age, sex and urban/rural distribution are used in the calculation of rates in the *Yearbook*. Vital rates by age and sex were calculated using data that appear in table 7 in this issue or the corresponding tables of previous issues of the *Demographic Yearbook*.

3.1 Sources of variation of data

The comparability of data is affected by several factors, including (1) the definition of total population; (2) the definition used to classify the population into its urban/rural components; (3) the accuracy of age reporting; (4) the extent of over-enumeration or under-enumeration in the most recent census or other source of benchmark population statistics; and (5) the quality of population estimates. These five factors will be discussed in some detail in sections 3.1.1 to 3.2.2 below. Other relevant problems are discussed in the technical notes to the individual tables. Readers interested in more detail, relating in particular to the basic concepts of population size, distribution and characteristics as elaborated by the United Nations, should consult the *Principles and Recommendations for Population and Housing Censuses, Revision 1*.⁵

3.1.1 Total population

The most important impediment to comparability of total populations is the difference between the concept of a de facto and de jure population. A de facto population should include all persons physically present in the country or area at the reference date. The de jure population, by contrast, should include all usual residents of the given country or area, whether or not they were physically present in the area at the reference date. By definition, therefore, a de facto total and a de jure total are not entirely comparable.

Comparability of even two de facto or de jure totals is often affected by the fact that strict conformity to either of these concepts is rare. For example, some so-called de facto counts do not include foreign military, naval and diplomatic personnel present in the country or area on official duty, and their accompanying family members and servants; some do not include foreign visitors in transit through the country or area or transients on ships in the harbour. On the other hand, they may include such persons as merchant seamen and fishermen who are temporarily out of the country or area working at their trade.

The de jure population figure presents even greater variations in comparability, in part because it depends in the first place on the concept of "usual residence", which varies from one country or area to another and is difficult to apply consistently in a census or survey enumeration. For example, non-national civilians temporarily in a country or area as short-term workers may officially be considered residents after a stay of a specified period of time or they may be considered as non-residents throughout the duration of their stay; at the same time, these individuals may be officially considered as residents or non-residents of the country or area from which they came, depending on the duration and/or purpose of their absence. Furthermore, regardless of the official treatment, individual respondents may apply their own interpretation of residence in responding to the inquiry. In addition, there may be considerable differences in the accuracy with which countries or areas are informed about the number of their residents temporarily out of the country or area.

As far as possible, the population statistics presented in the tables of the *Yearbook* refer to the de facto population. Those reported to have been based on the de jure concept are identified as such. Figures not otherwise qualified may, therefore, be assumed to have been reported by countries or areas as being based on a de facto definition of the population. In an effort to overcome, to the extent possible, the effect of the lack of strict conformity to either the de facto or the de jure concept given above, significant exceptions with respect to inclusions and exclusions of specific population groups, are footnoted when they are known.

It should be remembered, however, that the necessary detailed information has not been available in many cases. It cannot, therefore, be assumed that figures not thus qualified reflect strict de facto or de jure definitions.

A possible source of variation within the statistics of a single country or area may arise from the fact that some countries or areas collect information on both the de facto and the de jure population in, for example, a census, but prepare detailed tabulations for only the de jure population. Hence, even though the total population shown in table 3 is de facto, the figures shown in the tables presenting various characteristics of the population, for example, urban/rural distribution, age and sex distribution, may be de jure.

3.1.2 Urban/rural classification

International comparability of urban/rural distributions is seriously impaired by the wide variation among national definitions of the concept of "urban". The definitions used by individual countries or areas and their implications are shown at the end of technical notes for table 6.

3.1.3 Age distribution

The classification of population by age is a core element of most analyses, estimation and projection of population statistics. Unfortunately, age data are subject to a number of sources of error and non-comparability. Accordingly, the reliability of age data should be of concern to users of these statistics.

3.1.3.1 Collection and compilation of age data

Age is the estimated or calculated interval of time between the date of birth and the date of the census or survey, expressed in completed solar years.³ There are two methods of collecting information on age. The first is to obtain the date of birth for each member of the population in a census or survey and then to calculate the completed age of the individual by subtracting the date of birth from the date of enumeration.⁶ The second method is to record the individual's completed age at the time of the census or survey, that is to say, age at last birthday.

The recommended method is to calculate age at last birthday by subtracting the exact date of birth from the date of the census. Some practices, however, do not use this method but instead calculate the difference between the year of birth and the year of the census. Classifications of this type are footnoted whenever possible. They can be identified to a certain extent by a smaller than expected population under one year of age. However, an irregular number of births from one year to the next or age selective omission of infants may also obscure the expected population under one year of age.

3.1.3.2 Errors in age data

Errors in age data may be due to a variety of causes, including ignorance of the correct age; reporting years of age in terms of a calendar concept other than completed solar years since birth,⁷ carelessness in reporting and recording age; a general tendency to state age in figures ending in certain digits (such as zero, two, five and eight); a tendency to exaggerate length of life at advanced ages; a subconscious aversion to certain numbers; and wilful misrepresentations.

These reasons for errors in reported age data are common to most investigations of age and to most countries or areas, and they may significantly impair comparability of the data.

As a result of the above-mentioned difficulties, the age-sex distribution of population in many countries or areas shows irregularities which may be summarized as follows: (1) a deficiency in the number of infants and young children; (2) a concentration at ages ending with zero and five (that is, 5, 10, 15, 20...); (3) heaping at even ages (for example, 10, 12, 14...) relative to odd ages (for example, 11, 13, 15...); (4) unexpectedly large differences between the frequency of males and females at certain ages; and (5) unaccountably large differences between the frequencies in adjacent age groups. Comparing of identical age-sex cohorts from successive censuses, as well as studying the age-sex composition of each census, may reveal these and other inconsistencies, some of which in varying degree are characteristic of even the most modern censuses.

3.1.3.3 Evaluation of accuracy

To measure the accuracy of data by age, based on the evidence of irregularities in 5-year groups, an index was devised for presentation in the *Demographic Yearbook 1949-1950*.⁸ Although this index was sensitive to various sources of inaccuracy in the data, it could also be affected considerably by real fluctuations in past demographic processes. It could not, therefore, be applied indiscriminately to all types of statistics, unless certain adjustments were made and caution used in the interpretation of results.

The publication of population statistics by single years of age in the *Demographic Yearbook 1955* made it possible to apply a simple, yet highly sensitive, index known as Whipple's Index, or the Index of Concentration,⁹ the interpretation of which is relatively free from consideration of factors not connected with the accuracy of age reporting. More refined methods for the measurement of accuracy of distributions by

single year of age have been devised, but this particular index was selected for presentation in the *Demographic Yearbook* for its simplicity and the wide use it has already found in other sources.

Whipple's Index is obtained by summing the age returns between 23 and 62 years inclusive and finding what percentage is borne by the sum of the returns of years ending with 5 and 0 to one-fifth of the total sum.

The results would vary between a minimum of 100, representing no concentration at all, and a maximum of 500, if no returns were recorded with any digits other than the two mentioned.¹⁰

The index is applicable to all age distributions for which single years are given at least to the age of 62, with the following exceptions: (1) where the data presented are the result of graduation, no irregularity is scored by Whipple's Index, even though the graduated data may still be affected by inaccuracies of a different type; and (2) where statistics on age have been derived by reference to the year of birth, and tendencies to round off the birth year would result in an excessive number of ages ending in odd numbers, the frequency of age reporting with terminal digits 5 and 0 is not an adequate measure of their accuracy.

Most recently, the index has been computed for all the single-year age distributions from censuses held between 1985 and 2003, with the exception of those excluded on the criteria set forth above. Such data are published in the special issue of the *Demographic Yearbook* special topic on population censuses, Volume 1, which is available online at <http://unstats.un.org/unsd/demographic/products/dyb/dybcens.htm>.

Although Whipple's Index measures only the effects of preferences for ages ending in 5 and 0, it can be assumed that such digit preference is usually connected with other sources of inaccuracy in age statements and the index can be accepted as a fair measure of the general reliability of the age distribution.

3.2 Methods used to indicate quality of published statistics

To the extent possible, efforts have been made to give the reader an indication of reliability of the statistics published in the *Demographic Yearbook*. This has been approached in several ways. Any information regarding a possible under-enumeration or over-enumeration, coming from a postcensal survey, for example, has been noted in the footnotes to table 3. Any deviation from full national coverage, as explained in section 2.1 under Geographical Aspects, has also been noted. In addition, national statistical offices have been asked to evaluate the estimates of total population they submit to the Statistics Division of the United Nations.

3.2.1 Treatment of time series of population estimates

When a series of mid-year population estimates are presented, the same indication of quality is shown for the entire series as was determined for the latest estimate. The quality is indicated by the type face employed.

No attempt has been made to split the series even though it is evident that in cases where the data are now considered reliable, in earlier years, many may have been considerably less reliable than the current classification implies. Thus it will be evident that this method overstates the probable reliability of the time series in many cases. It may also understate the reliability of estimates for years immediately preceding or following a census enumeration.

3.2.2 Treatment of estimated distributions by age and other demographic characteristics

Estimates of the age-sex distribution of population may be constructed by two major methods: (1) by applying the specific components of population change to each age-sex group of the population as enumerated at the time of the census, and (2) by distributing the total estimated for a postcensal year proportionately according to the age-sex structure at the time of the census. Estimates constructed by the latter method are not published in the *Demographic Yearbook*.

Estimated age-sex distributions are categorized as "reliable" or otherwise, according to the method of construction established for the latest estimate of total mid-year population. Hence, the quality designation of the total figure, as indicated by the code, is considered to apply also to the whole distribution by age and sex, and the data are set in *italic* or roman type, as appropriate, on this basis alone. Further evaluation of detailed age structure data has not been undertaken to date.

4. VITAL STATISTICS

For purposes of the *Demographic Yearbook*, vital statistics have been defined as statistics of live birth, death, foetal death, marriage and divorce.

This volume of the *Yearbook* presents general tables on natality, nuptiality and divorce as well as tables on mortality referring to: foetal mortality, infant and maternal mortality and general mortality.

4.1 Sources of variation of data

Most of the vital statistics data published in this *Yearbook* come from national civil registration systems. The completeness and the accuracy of the data that these systems produce vary from one country or area to another.

The provision for a national civil registration system is not universal, and in some cases, the registration system covers only certain vital events. For example, in some countries or areas only births and deaths are registered. There are also differences in the effectiveness with which national laws pertaining to civil registration operate in the various countries or areas. The manner in which the law is implemented and the degree to which the public complies with the legislation determine the reliability of the vital statistics obtained from the civil registers.

It should be noted that some statistics on marriage and divorce are obtained from sources other than civil registers. For example, in some countries or areas, the only source for data on marriages is church registers. Divorce statistics, on the other hand, are obtained from court records and/or civil registers according to national practice. The actual compilation of these statistics may be the responsibility of the civil registrar, the national statistical office or other government offices.

Other factors affecting international comparability of vital statistics are much the same as those that must be considered in evaluating the variations in other population statistics. Differences in statistical definitions of vital events, differences in geographical and ethnic coverage of the data and diverse tabulation procedures may also influence comparability.

In addition to vital statistics from civil registers, some vital statistics published in the *Yearbook* are official estimates. These estimates are frequently from sample surveys. As such, their comparability may be affected by the national completeness of reporting in household surveys, non-sampling and sampling errors and other sources of bias.

Readers interested in more detailed information on standards for vital statistics should consult the *Principles and Recommendations for a Vital Statistics System Revision 2*;¹¹ *Handbook of vital statistics systems and Methods Volume 1, Legal, Organizational and Technical Aspects*;¹³ *Handbook of Vital Statistics Systems and Methods, Volume 2, Review of national practices*;¹² *Handbook on Civil Registration and Vital Statistics Systems: Management, Operation and Maintenance*;¹³ *Handbook on Civil Registration and Vital Statistics Systems: Preparation of a Legal Framework*;¹⁴ *Handbook on Civil Registration and Vital Statistics Systems: Developing Information, Education and Communication*;¹⁵ *Handbook on Civil Registration and Vital Statistics Systems: Policies and Protocols for the Release and Archiving of Individual Records*;¹⁶ and *Handbook on Civil Registration and Vital Statistics Systems: Computerization*.¹⁷ The *Handbook of Household Surveys*¹⁸ provides information in collection and evaluation of data on fertility, mortality and other vital events collected in household surveys. These publications are also available on the website at <http://unstats.un.org/unsd/demographic/sources/civilreg/default.htm>.

4.1.1 Statistical definition of events

An important source of variation lies in the statistical definition of each vital event. The *Demographic Yearbook* attempts to collect data on vital events, using the standard definitions put forth in paragraph 57 of *Principles and Recommendations for a Vital Statistics System Revision 2*.¹⁴ These are as follows:

LIVE BIRTH is the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy, which after such separation breathes or shows any other evidence of life such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached; each product of such a birth is considered live-born regardless of gestational age.

DEATH is the permanent disappearance of all evidence of life at any time after live birth has taken place (postnatal cessation of vital functions without capability of resuscitation). This definition therefore excludes foetal deaths.

FOETAL DEATH is death prior to the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy; the death is indicated by the fact that after such separation the foetus does not breathe or show any other evidence of life, such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles. Late foetal deaths are those of twenty-eight or more completed weeks of gestation. These are synonymous with the events reported under the pre-1950 term stillbirth¹⁹.

MARRIAGE is an act, ceremony or process by which the legal relationship of husband and wife is constituted. The legality of the union may be established by civil, religious or other means as recognized by the laws of each country or area.

DIVORCE is a final legal dissolution of a marriage, that is, that separation of husband and wife which confers on the parties the right to remarriage under civil, religious and/or other provisions, according to the laws of each country.

In addition to these recommended definitions, the *Demographic Yearbook* collects and presents data on abortions, defined as:

ABORTION is defined, with reference to the woman, as any interruption of pregnancy before 28 weeks of gestation with a dead foetus. There are two major categories of abortion: spontaneous and induced. Induced abortions are those initiated by deliberate action undertaken with the intention of terminating pregnancy; all other abortions are considered spontaneous.

4.1.2 Problems relating to standard definitions

A basic problem affecting international comparability of vital statistics is deviations from the standard definitions of vital events. An example of this can be seen in the cases of live births and foetal deaths.²⁰ In some countries or areas, an infant must survive for at least 24 hours, to be inscribed in the live-birth register. Infants who die before the expiration of the 24-hour period are classified as late foetal deaths and, barring special tabulation procedures, they would not be counted either as live births or as deaths. Similarly, in several other countries or areas, those infants who are born alive but die before registration of their birth, are also considered late foetal deaths.

Unless special tabulation procedures are adopted in such cases, the live-birth and death statistics will both be deficient by the number of these infants, while the incidence of late foetal deaths will be increased by the same amount. Hence the infant mortality rate is underestimated. Although both components (infant deaths and live births) are deficient by the same absolute amount, the deficiency is proportionately greater in relation to the infant deaths, causing greater errors in the infant mortality rate than in the birth rate.

Moreover, the practice exaggerates the late foetal death ratios. Some countries or areas make provision for correcting this deficiency (at least in the total frequencies) at the tabulation stage. Data for which the correction has not been made are indicated by footnote whenever possible.

The definitions used for marriage and divorce also present problems for international comparability. Unlike birth and death, which are biological events, marriage and divorce are defined only in terms of law and custom and as such are less amenable to universally applicable statistical definitions. They have therefore been defined for statistical purposes in general terms referring to the laws of individual countries or areas. Laws pertaining to marriage and particularly to divorce, vary from one country or area to another. With respect to marriage, the most widespread requirement relates to the minimum age at which persons may marry but frequently other requirements are specified.

When known the minimum legal age of men and women at which marriage can occur with (or in some cases without) parental consent is presented in table 24-1. Laws and regulations relating to the dissolution of marriage by divorce range from total prohibition, through a wide range of grounds upon which divorces may be granted, to the granting of divorce in response to a simple statement of desire or intention by husbands.

4.1.3 Fragmentary geographical or ethnic coverage

Ideally, vital statistics for any given country or area should cover the entire geographical area and include all ethnic groups. Fragmentary coverage is, however, not uncommon. In some countries or areas, registration is compulsory for only a small part of the population, limited to certain ethnic groups, for example. In other places there is no national provision for compulsory registration, but only municipal or state ordinances that do not cover the entire geographical area. Still others have developed a registration area that comprises only a part of the country or area, the remainder being excluded because of inaccessibility or for economic and cultural considerations that make regular registration practically impossible.

4.1.4 Tabulation procedures

4.1.4.1 *By place of occurrence*

Vital statistics presented on the national level relate to the de facto, that is, the present-in-area population. Thus, unless otherwise noted, vital statistics for a given country or area cover all the events that occur within its present boundaries and among all segments of the population therein. They may be presumed to include events among nomadic tribes and indigenous peoples, and among nationals and foreigners. When known, deviations from the de facto concept are footnoted.

Urban/rural differentials in vital rates for some countries may vary considerably depending on whether the relevant vital events were tabulated on the basis of place of occurrence or place of usual residence. For example, if a substantial number of women residing in rural areas near major urban centres travel to hospitals or maternity homes located in a city to give birth, urban fertility and neo-natal and infant mortality rates will usually be higher (and the corresponding rural rates will usually be lower) if the events are tabulated on the basis of place of occurrence rather than on the basis of place of usual residence. A similar process will affect general mortality differentials if substantial numbers of persons residing in rural areas use urban health facilities when seriously ill.

4.1.4.2 *By date of occurrence versus by date of registration*

To the extent possible, the vital statistics presented in the *Demographic Yearbook* refer to events that occurred during the specified year, rather than to those that were registered during that period. However, a considerable number of countries or areas tabulate their vital statistics not by date of occurrence, but by date of registration. Because such statistics can be very misleading, the countries or areas known to tabulate vital statistics by date of registration are identified in the tables by a plus sign (+). Since information on the method of tabulating vital statistics is not available for all countries and areas, tabulation by date of registration may be more prevalent than the symbols on the vital statistics tables would indicate.

Because quality of data is inextricably related to the timeliness of registration, this must always be considered in conjunction with the quality code description in section 4.2.1 below. If registration of births is complete and timely (code C), the ill effects of tabulating by date of registration, are, for all practical purposes, nullified. Similarly, with respect to death statistics, the effect of tabulating events by date of registration may be minimized in many countries or areas in which the sanitary code requires that a death must be registered before a burial permit can be issued, and this regulation tends to make registration prompt. With respect to foetal death, registration is usually done right away or not at all. Therefore, if registration is prompt, the difference between statistics tabulated by date of occurrence and those tabulated by date of registration may be negligible. In many cases, the length of the statutory time period allowed for registering various vital events plays an important part in determining the effects of tabulation by date of registration on comparability of data.

With respect to marriage and divorce, the practice of tabulating data by date of registration does not generally pose serious problems. In many countries or areas marriage is a civil legal contract which, to establish its legality, must be celebrated before a civil officer. It follows that for these countries or areas registration would tend to be almost automatic at the time of, or immediately following, the marriage ceremony. Because the registration of a divorce in many countries or areas is the responsibility solely of the court or the authority which granted it, and since the registration record in such cases is part of the records of the court proceedings, it follows that divorces are likely to be registered soon after the decree is granted.

On the other hand, if registration is not prompt, vital statistics by date of registration will not produce internationally comparable data. Under the best circumstances, statistics by date of registration will include

primarily events that occurred in the immediately preceding year; in countries or areas with less developed systems, tabulations will include some events that occurred many years in the past. Examination of available information reveals that delays of many years are not uncommon for birth registration, though the majority is recorded between two to four years after birth.

As long as registration is not prompt, statistics by date of registration will not be internationally comparable either among themselves or with statistics by date of occurrence.

It should also be mentioned that lack of international comparability is not the only limitation introduced by date-of-registration tabulation. Even within the same country or area, comparability over time may be lost by the practice of counting registrations rather than occurrences. If the number of events registered from year to year fluctuates because of ad hoc incentives to stimulate registration, or to the sudden need, for example, for proof of (unregistered) birth or death to meet certain requirements, vital statistics tabulated by date of registration are not useful in measuring and analyzing demographic levels and trends. All they can give is an indication of the fluctuations in the need for a birth, death or marriage certificate and the work-load of the registrars. Therefore, statistics tabulated by date of registration may be of very limited use for either national or international studies.

4.2 Methods used to indicate quality of published vital statistics

The quality of vital statistics can be assessed in terms of a number of factors. Most fundamental is the completeness of the civil registration system on which the statistics are based. In some cases, the incompleteness of the data obtained from civil registration systems is revealed when these events are used to compute rates. However, this technique applies only where the data are markedly deficient, where they are tabulated by date of occurrence and where the population base is correctly estimated. Tabulation by date of registration will often produce rates which appear correct, simply because the numerator is artificially inflated by the inclusion of delayed registrations and, conversely, rates may be of credible magnitude because the population at risk has been underestimated. Moreover, it should be remembered that knowledge of what is credible in regard to levels of fertility, mortality and nuptiality is extremely scanty for many parts of the world, and borderline cases, which are the most difficult to appraise, are frequent.

4.2.1 Quality code for vital statistics from registers.

In the *Demographic Yearbook* annual "Questionnaire on Vital Statistics" national statistical offices are asked to provide their own estimates of the completeness of the births, deaths, late foetal deaths, marriages and divorces recorded in their civil registers.

On the basis of information from the questionnaires, from direct correspondence and from relevant official publications, it has been possible to classify current national statistics from civil registers of birth, death, infant death, late foetal death, marriage and divorce into three broad quality categories, as follows:

C: Data estimated to be virtually complete, that is, representing at least 90 per cent of the events occurring each year.

U: Data estimated to be incomplete, that is representing less than 90 per cent of the events occurring each year.

]: Data not derived from civil registration systems but considered reliable, such as estimates derived from projections, other estimation techniques or population and housing census.

...: Data for which no specific information is available regarding completeness.

These quality codes appear in the first column of the tables which show total frequencies and crude rates (or ratios) over a period of years for live births (table 9), late foetal deaths (table 12), infant deaths (table 15), deaths (table 18), marriages (table 23), and divorces (table 25).

The classification of countries and areas in terms of these quality codes may not be uniform. Nevertheless, it was felt that national statistical offices were in the best position to judge the quality of their data. It was considered that even the very broad categories that could be established on the basis of the available information would provide useful indicators of the quality of the vital statistics presented in this *Yearbook*.

In the past, the bases of the national estimates of completeness were usually not available. In connection with the *Demographic Yearbook 1977*, countries were asked, for the first time, to provide some indication of the basis of their completeness estimates. They were requested to indicate whether the completeness estimates reported for registered live births, deaths, and infant deaths were prepared on the basis of demographic analysis, dual record checks or some other specified method. Relatively few countries or areas have responded to this new question; therefore, no attempt has been made to revise the system of quality codes used in connection with the vital statistics data presented in the *Yearbook*. It is hoped that, in the future, more countries will be able to provide this information so that the system of quality codes used in connection with the vital statistics data presented in the *Yearbook* may be revised.

Among the countries or areas indicating that the registration of live births was estimated to be 90 per cent or more complete (and hence classified as C in table 9), the following countries or areas provided information on the basis of this completeness estimate:

(a) Demographic analysis -- Argentina, Australia, Canada, Chile, Croatia, Cuba, Czech Republic, Egypt, French Guiana, Guadeloupe, Guernsey, Iceland, Ireland, Israel, Republic of Korea, Kuwait, Latvia, Mauritius, Puerto Rico, Romania, San Marino, Singapore, Sweden, Switzerland and United States.

(b) Dual record check -- Bahamas, Barbados, Belgium, Bulgaria, Cook Islands, Cuba, Cyprus, Denmark, Estonia, Fiji, Finland, France, French Guiana, Greece, Guam, Guadeloupe, Guernsey, Hungary, Iceland, Isle of Man, Japan, Kyrgyzstan, Lithuania, Maldives, New Zealand, Peninsular Malaysia, Romania, Saint Kitts and Nevis, Saint Lucia, Singapore, Sri Lanka, Sweden, Switzerland, Tokelau, Uruguay and Venezuela.

(c) Other specified methods -- Belgium, Bermuda, Cayman Islands, China Hong Kong SAR, China Macao SAR, Germany, Greenland, Iceland, Japan, Luxembourg, Netherlands, Norway, Poland, Singapore and Slovenia.

Among the countries or areas indicating that the registration of deaths was estimated to be 90 per cent or more complete (and hence classified as C in table 18), the following countries or areas provided information on the basis of this completeness estimate:

(a) Demographic analysis -- Argentina, Australia, Canada, Chile, Cuba, Egypt, French Guiana, Guadeloupe, Guernsey, Iceland, Ireland, Israel, Kuwait, Latvia, Mauritius, Puerto Rico, Romania, San Marino, Singapore, Switzerland and United States.

(b) Dual record check -- Bahamas, Bulgaria, Cook Islands, Cuba, Denmark, Fiji, Finland, France, Greece, Greenland, Guam, Guernsey, Iceland, Isle of Man, Maldives, New Zealand, Romania, Saint Kitts and Nevis, Saint Lucia, Singapore, Sri Lanka, Sweden, Switzerland, Tokelau and Uruguay.

(c) Other specified methods -- Belgium, Bermuda, Cayman Islands, Germany, Hong Kong SAR, Iceland, Ireland, Japan, Luxembourg, Netherlands, Norway, Poland, Singapore and Slovenia.

Among the countries or areas indicating that the registration of infant deaths was estimated to be 90 per cent or more complete (and hence classified as C in table 15), the following countries or areas provided information on the basis of this completeness estimate:

(a) Demographic analysis -- Argentina, Australia, Canada, Chile, Cuba, Egypt, Iceland, Ireland, Israel, Kuwait, Latvia, Mauritius, Puerto Rico, Romania, San Marino, Singapore, Sri Lanka, Switzerland and United States.

(b) Dual record check -- Bahamas, Bulgaria, Cook Islands, Cuba, Denmark, Fiji, Finland, France, Greece, Greenland, Guam, Guernsey, Iceland, Isle of Man, Japan, Maldives, New Zealand, Romania, Saint Kitts and Nevis, Saint Lucia, Singapore, Sweden, Switzerland, Tokelau and Uruguay.

(c) Other specified methods -- Belgium, Bermuda, Cayman Islands, Germany, Hong Kong SAR, Iceland, Japan, Luxembourg, Netherlands, Norway, Poland, Singapore and Slovenia.

4.2.2 Treatment of vital statistics from registers

On the basis of the quality code described above, the vital statistics shown in all tables of the *Yearbook* are treated as either reliable or unreliable. Data coded C are considered reliable and appear in roman type. Data coded U or ... are considered unreliable and appear in *italics*. Although the quality code itself appears only in certain tables, the indication of reliability (that is, the use of *italics* to indicate unreliable data) is shown in all tables presenting vital statistics data.

In general, the quality code for deaths shown in table 18 is used to determine whether data on deaths in other tables appear in roman or *italic* type. However, for some of the maternal deaths data shown in *italics* in table 17, the known quality code differs from that ascribed on the basis of the completeness of registration of the total number of deaths. In cases where the quality code in table 18 does not correspond with the quality level implied by the typeface used in table 17, relevant information regarding the completeness of maternal mortality is given in a footnote.

The same indication of reliability used in connection with tables showing the frequencies of vital events is also used in connection with tables showing the corresponding vital rates. For example, death rates computed using deaths from a register that is incomplete or of unknown completeness are considered unreliable and appear in *italics*. Strictly speaking, to evaluate vital rates more precisely, one would have to also take into account the accuracy of population data used in the denominator of these rates. The quality of population data is discussed in section 3.2 of the Technical Notes.

It should be noted that the indications of reliability used for infant mortality rates, maternal mortality rates and late foetal death ratios (all of which are calculated using the number of live births in the denominator) are determined on the basis of the quality codes for infant deaths, deaths and late foetal deaths respectively. To evaluate these rates and ratios more precisely, one would have to take into account the quality of the live-birth data used in the denominator of these rates and ratios. The quality codes for live births are shown in table 9 and described more fully in the text of the technical notes for that table.

4.2.3 Treatment of time series of vital statistics from registers

The quality of a time series of vital statistics is more difficult to determine than the quality of data for a single year. Since a time series of vital statistics is usually generated only by a system of continuous civil registration, it was assumed that the quality of the entire series was the same as that for the latest year's data obtained from the civil register. The entire series is treated as described in section 4.2.2 above. That is, if the quality code for the latest registered data is C, the frequencies and rates for earlier years are also considered reliable and appear in roman type. Conversely, if the latest registered data are coded as U or ... then data for earlier years are considered unreliable and appear in *italics*. It is recognized that this method is not entirely satisfactory because it is known that data from earlier years in many of the series were considerably less reliable than the current code implies.

4.2.4 Treatment of estimated vital statistics

In addition to data from vital registration systems, estimated frequencies and rates of the events, usually ad hoc official estimates that have been derived either from the results of a sample survey or by demographic analyses, also appear in the *Demographic Yearbook*. Estimated frequencies and rates have been included in the tables because it is assumed that they provide information that is more accurate than that from existing civil registration systems. By implication, they are assumed to be reliable and as such they are set in roman type. Estimated frequencies and rates continue to be treated in this manner even when they are interspersed in a time series with data from civil registers.

In tables showing the quality code, the code applies only to data from civil registers. If a series of data for a country or area contains both data from a civil register and estimated data, the code applies only to the registered data; if only estimated data are shown, the symbol (I) is shown.

4.3 Cause of death

World Health Organization (WHO) Member States are bound by the International Nomenclature Regulations to provide the Organization with cause of death data coded in accordance with the current revision of the International Statistical Classification of Diseases and Related Health Problems (ICD) as adopted from time to time by the World Health Assembly²¹. The data are collected by the WHO using the ICD. In order to promote international comparability of cause of death statistics, the World Health

Organization organizes and conducts an international Conference for the revision of the ICD on a regular basis in order to ensure that the Classification is kept current with the most recent clinical and statistical concepts. The data are now usually submitted to WHO at the full four-character level of detail provided by the ICD and are stored in the WHO Mortality Database at the level of detail as provided by the country. For earlier versions, however, the data are only available according to the ICD's list of 150 causes. Data from the WHO Mortality Database are available in electronic format at <http://www3.who.int/whosis/menu.cfm>.

Although revisions provide an up-to-date version of the ICD, such revisions create several problems related to the comparability of cause of death statistics. The first is the lack of comparability over time that inevitably accompanies the use of a new classification. The second problem affects comparability between countries and areas because they may adopt the new classification at different times. The more refined the classification becomes the greater is the need for expert clinical diagnosis of cause of death. In many countries or areas, few of the deaths occur in the presence of an attendant, who is medically trained, i.e., most deaths are certified by a lay attendant. Because the ICD contains many diagnoses that cannot be identified by a non-medical person, the ICD does not always promote international comparability particularly between countries and areas where the level of medical services differ widely.

The chapters of the tenth revision²², the latest revision of the ICD, consist of an alphanumeric coding scheme of one letter followed by three numbers at the four-character level. Chapter one contains infectious and parasitic diseases, chapter two refers to all neoplasms, chapter three to disorders of the immune mechanism including diseases of the blood and blood-forming organs; and chapter four to endocrine, nutritional and metabolic diseases. The remaining chapters group diseases according to the anatomical site affected, except for chapters that refer to mental disorders; complications of pregnancy, childbirth and the puerperium; congenital malformations; and conditions originating in the perinatal period. Finally, an entire chapter is devoted to symptoms, signs, and abnormal findings.

4.3.1 Maternal mortality

According to the tenth revision of the ICD, "Maternal death" is defined as the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and the site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes.

"Maternal deaths should be subdivided into direct and indirect obstetric deaths. Direct obstetric deaths are those resulting from obstetric complications of the pregnant state (pregnancy, labour and puerperium), from interventions, omissions, incorrect treatment, or from a chain of events resulting from any of the above. Indirect obstetric deaths are those resulting from previous existing disease or disease that developed during pregnancy and which was not due to direct obstetric causes, but which was aggravated by physiologic effects of pregnancy".

While the denominator maternal rate should be the number of pregnant women, it is impossible to determine the number of pregnant women. A further recommendation by the tenth revision conference is therefore that maternal mortality rates be expressed per 100,000 live births or per 100,000 total births (live births and foetal deaths).²³ The maternal mortality rate calculated here is expressed per 100,000 live births. Although live births do not represent an unbiased estimate of pregnant women, this figure is more reliable than other estimates in particular, live births are more accurately registered than live births plus foetal deaths.

4.3.2 Perinatal mortality

The definition of perinatal death was recommended by the Study Group on Perinatal Mortality set up by the World Health Organization. The International Conference for the Eighth Revision of the International Classification of Diseases adopted the recommendation that the perinatal period be defined "as extending from the 28th week of gestation to the seventh day of life". Noting that several countries considered as late foetal deaths any foetal death of 20 weeks or longer gestation, the Conference agreed to accept a broader definition of perinatal death that extends from the 20th week of gestation to the 28th day of life. This alternative definition was believed to promote more complete registration of events between 28 weeks of gestation and the end of the first 6 days of life. In 1975, the Ninth Revision Conference recommended the collection of perinatal mortality statistics by use of a standard perinatal death certificate according to a definition which not only includes a minimum length of gestation but also minimum weight and length criteria.

In table 19 of the *Demographic Yearbook 1996* and previous issues of the Yearbook that included perinatal mortality statistics, the definition of perinatal deaths used is the sum of late foetal deaths (foetal deaths of 28 or more weeks of gestation) and infant deaths within the first week of life. In addition, in order to standardize the definition and eliminate differences due to national practice, the figures on perinatal death are calculated by the Statistics Division for inclusion in the *Demographic Yearbook*. Following the recommendations of the Tenth Revision Conference, the perinatal mortality rate is calculated per 1,000 live births in order to minimize the effect of limited foetal death registration on the magnitude of the denominator.

25

NOTES

- ¹ The data on maternal mortality are from the World Health Organization, and are available at <http://www3.who.int/whosis/menu.cfm>, as one cause of death.
- ² There are two exceptions – the 1978 and 1991 issues, which were disseminated in separate volumes from the respective regular issues.
- ³ *World Population Prospects: The 2004 Revision*, United Nations, New York, forthcoming. Until this is published, highlights and selected output are available by following links at www.unpopulation.org, and information about methodology and data are available in *World Population Prospects: The 2002 Revision*, Sales No. E.03.XIII.6, United Nations, New York, 2003.
- ⁴ ST/ESA/STAT/SER.M/49/Rev.4/WWW ; <http://unstats.un.org/unsd/methods/m49/m49.htm>; see also Standard Country or Area Codes for Statistical Use, Sales No. M.98.XVII.9, United Nations, New York, 1999.
- ⁵ Sales No. E.98.XVII.8, United Nations, New York, 1998.
- ⁶ Alternatively, if a population register is used, completed ages are calculated by subtracting the date of birth of individuals listed in the register from a reference date to which the age data pertain.
- ⁷ A source of non-comparability may result from differences in the method of reckoning age, is for example, the Western versus the Eastern or, as it is usually known, the English versus the Chinese system. By the latter, a child is considered one year old at birth and advances an additional year at each Chinese New Year. The effect of this system is most obvious at the beginning of the age span, where the frequencies in the under-one-year category are markedly understated. The effect on higher age groups is not so apparent. Distributions constructed on this basis are often adjusted before publication, but the possibility of such aberrations should not be excluded when census data by age are compared.
- ⁸ In this index, differences were scored from expected values of ratios between numbers of either sex in the same age group, and numbers of the same sex in adjoining age group. In compounding the score, allowance had to be made for certain factors such as the effects of past fluctuations in birth rates, of heavy war casualties, and of the smallness of the population itself. A detailed description of the index, with results from its application to the data presented in the 1949-1950 and 1951 issues of the *Demographic Yearbook*, is furnished in *Population Bulletin, No. 2* (United Nations publication, Sales No. 52.XIII.4), pp. 59-79. The scores obtained from statistics presented in *Demographic Yearbook 1952* are presented in that issue, and the index has also been briefly explained in that issue, as well as those of 1953 and 1954.
- ⁹ United States, Bureau of the Census, Thirteenth Census, Vol. I (Washington, D.C., U.S. Government Printing Office), pp. 291-292.
- ¹⁰ J. T. Marten, Census of India, 1921, vol. I, part I (Calcutta, 1924), pp. 126-127.
- ¹¹ Sales No. E. 01.XVII.10, United Nations, New York, 2001; http://unstats.un.org/unsd/demographic/vital_statistics/index.htm.
- ¹² Sales No. E.84.XVII.11, United Nations, New York, 1985; http://unstats.un.org/unsd/demographic/vital_statistics/index.htm.
- ¹³ *Handbook on Civil Registration and Vital Statistics Systems: Management, Operation and Maintenance*, Sales No. E.98.XVII.11, United Nations, New York, 1998. http://unstats.un.org/unsd/demographic/vital_statistics/index.htm.
- ¹⁴ Sales No. E. 98.XVII.7, United Nations, New York, 1998; http://unstats.un.org/unsd/demographic/vital_statistics/index.htm.
- ¹⁵ Sales No. E.98.XVII.4, United Nations, New York, 1998; http://unstats.un.org/unsd/demographic/vital_statistics/index.htm.

-
- ¹⁶ Sales No. E.98.XVII.6, United Nations, New York, 1998;
http://unstats.un.org/unsd/demographic/vital_statistics/index.htm.
- ¹⁷ Sales No. E.98.XVII.10, United Nations, New York, 1998;
http://unstats.un.org/unsd/demographic/vital_statistics/index.htm.
- ¹⁸ *Handbook of Household Surveys*, Sales No. E.83.XVII.13, United Nations, New York, 1984.
- ¹⁹ For more detailed discussion on this issue, refer to *Principles and Recommendations for a Vital Statistics System Revision 2*, Sales No. E. 01.XVII.10, United Nations, New York, 2001, para 57.
- ²⁰ For more information on historical and legal background on the use of differing definitions of live births and foetal deaths, comparisons of definitions used as of 1 January 1950, and evaluation of the effects of these differences on the calculation of various rates, see *Handbook of Vital Statistics Systems and Methods Volume 2, Review of National Practices*, Sales No. E.84.XVII.11, United Nations, New York, 1985, Chapter IV.
- ²¹ The World Health Assembly is the annual meeting of the Member States of the World Health Organization and its highest governing body.
- ²² *International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision, Volume 2, World Health Organization, Geneva, 1992.
- ²³ *International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision, Volume 2, World Health Organization, Geneva, 1992, pp. 129-136.