USAID
FROM THE AMERICAN PEOPLE

GUIDE TO DHS STATISTICS

Demographic and Health Surveys Methodology

This document is part of the Demographic and Health Survey's *DHS Toolkit* of methodology for the MEASURE DHS Phase III project, implemented from 2008-2013.

This publication was produced for review by the United States Agency for International Development (USAID). It was prepared by MEASURE DHS/ICF International.

DHS TOOLKIT

[THIS PAGE IS INTENTIONALLY BLANK]

GUIDE TO DHS STATISTICS

Shea Oscar Rutstein, Ph.D.
Guillermo Rojas, M.C.S., M.A.

Demographic and Health Surveys
ORC Macro
Calverton, Maryland

September 2006

Table of Contents

Table of Contents i
Foreword iv
I. Introduction 1
Purpose of the Guide 1
Basic Description of the Demographic and Health Surveys 1
II. Overview of Key Aspects of the Organization of DHS Data 5
Structure of DHS Data Files 5
DHS Data 5
Household 5
Household 5
Household 5
Woman 5
Woman 5
Woman 5
Recode Files and Recode Variable Naming Conventions 8
Missing Values and Other Special Codes 10
Sampling Weights 12
Century Month Code (CMC) 15
Median Calculations 16
All Women Factors 18
Matching Data Files 21
Relationships 22
Normalization of Household Members and Children 23
III. Guide to Individual Statistics 26
Fertility 26
Current Fertility Rates 26
Children Ever Born and Living 34
Birth Intervals 36
Age at First Birth 38
Teenage Pregnancy and Motherhood 40
Contraception 41
Knowledge of Contraceptive Methods 41
Ever Use of Contraceptive Methods 41
Current Use of Contraceptive Methods 44
Number of Children At First Use of Contraception 47
Use of Social Marketing Brand of Pill 49
Knowledge of the Fertile Period 50
Age at Sterilization 51
Source of Contraception 53
Informed Choice 54
First-Year Contraceptive Discontinuation Rates 55
Reasons for Discontinuing Contraception 58
Future Use of Contraception 60
Reason for Not Intending to Use Contraception 61

Preferred Method of Contraception for Future Use 62
Exposure to Family Planning Messages 63
Contact of Nonusers with Family Planning Providers 64
Discussion of Family Planning With Husband 65
Attitudes Toward Family Planning 66
Nuptiality 67
Current Marital Status 67
Number of Co-Wives 68
Age at First Marriage 69
Median Age at First Marriage 70
Sexual Intercourse 71
Age at First Sexual Intercourse 71
Postpartum Exposure 73
Postpartum Amenorrhea, Abstinence, and Insusceptibility 73
Menopause 77
Fertility Preferences 78
Need for Family Planning—Currently Married Women 78
Need for Family Planning—Not Currently Married Women and All Women 82
Wanted Fertility Rates 87
Infant and Child Mortality 90
Early Childhood Mortality Rates 90
Mortality Rate Calculation 94
Perinatal Mortality Rate 96
High-Risk Fertility Behavior Risk Ratios 98
Maternal and Child Health 100
Antenatal Care 100
Number of Antenatal Care Visits and Timing of First Visit 101
Antenatal Care 103
Antenatal Care 104
Delivery Care 105
Place of Delivery 105
Delivery Care 106
Delivery Care 107
Postnatal Care 108
Vaccination Rates 109
Child Morbidity Prevalence and Treatment: Acute Respiratory Illness and Fever 112
Child Morbidity Prevalence and Treatment: Diarrheal Disease 114
Child Morbidity Prevalence and Treatment: Diarrheal Disease 116
Childhood Morbidity Prevention: Malaria 117
Childhood Morbidity Prevention: Diarrhea 118
Women’s Access to Health Care 119
Women’s Use of Smoking Tobacco 120
Infant Feeding 121
Initial Breastfeeding 121
Infant Feeding 123
Breastfeeding and Complementary Feeding 123
Infant Feeding 127
Foods 127

Infant Feeding 129
Infant Feeding 131
Micronutrients 133
Iodization of Household Salt 133
Micronutrient Intake Among Children 134
Micronutrient Intake Among Women 136
Children's and Women's Nutritional Status 138
Children's Anemia Status 138
Children's and Women's Nutritional Status 140
Women's Anemia Status 140
Children's and Women's Nutritional Status 143
Children's Nutritional Status 143
Children's and Women's Nutritional Status 147
Women's Nutritional Status 147
HIV/AIDS/STDs 150
HIV/AIDS Prevention Knowledge 150
Knowledge of STD Symptoms 151
Condom Use at Last Sex 152
Adult and Maternal Mortality 153
Adult Mortality (Direct Method) 156
Adult Mortality (Indirect Method) 158

Foreword

DHS surveys collect a wealth of information on widely different topics for a sample of the population of the countries that participate in the DHS program. The information is processed and presented in tables that describe the situation of the relevant country. Before producing these tables, many steps are necessary to ensure that the data properly reflect the situations they intend to describe. This involves weighting the sample results, considering such factors as the sample design and response rates, the construction of summary variables, and the calculation of mortality and fertility rates.

Many of the procedures involved are straightforward and are familiar to demographic analysts. However, other procedures need special attention and are built on the basis of experience accumulated over many years regarding the preferred way of calculating certain indicators, what to guard against, and what not to forget.

This ***Guide to DHS Statistics*** is meant to be a tool for all data users. It can also be an aid for those who are just starting out in data analysis, a tool for checking procedures for those whose skills are more advanced, and a reference document for all who deal with data generated from the database of a DHS survey.

I hope that this valuable tool will significantly increase the capacity of many researchers to do their own tabulations and analyses. I trust that it will help make data users aware of the precautions to take and the procedures that need to be followed to ensure that generated data are as correct as they can be. It should also ensure that the strengths and limitations of certain data and indicators are clear to all users.

I thank Shea Rutstein and Guillermo Rojas for an excellent job and hope that you will find this document very useful in your work with DHS data and beyond.

Martin Vaessen

Project Director

I. Introduction

Purpose of the Guide

The purpose of this guide is twofold:

1. To provide an overview on key aspects of the organization and use of DHS data sets
2. To serve as a reference document for researchers on the definitions and calculation of complex statistics used in DHS reports

The document has been prepared in response to feedback received from DHS data users on the need for a detailed guide to assist them in understanding DHS data sets and reproducing the statistics in DHS reports. The initial section of the guide deals with general topics including DHS data file structure, sampling weights, century month codes, and the adjustment factors employed in surveys in which only ever-married women are interviewed. The remaining sections deal with the definitions of individual statistics used in DHS reports.

Basic Description of the Demographic and Health Surveys

THE DHS PROGRAM

The DHS program was established by the United States Agency for International Development (USAID) in 1984. It was designed as a follow-up to the World Fertility Survey and the Contraceptive Prevalence Survey projects. The DHS project was first awarded in 1984 to Westinghouse Health Systems (which subsequently evolved into part of OCR Macro). The project has been implemented in overlapping five-year phases; DHS-I ran from 1984 to1990; DHS-II from 1988 to1993; and DHS-III from 1992 to1998. In 1997, DHS was folded into the new multi-project MEASURE program as MEASURE *DHS+*.

The objectives of the DHS program are to:

- Provide decision makers in participating countries with improved information and analyses useful for informed policy choices
- Improve coordination and partnerships in data collection at the international and country levels
- Develop in participating countries the skills and resources necessary to conduct high-quality demographic and health surveys
- Improve data collection and analysis tools and methodology
- Improve the dissemination and utilization of data.

Since 1984, more than 130 nationally representative household-based surveys have been completed under the DHS project in about 70 countries. Many of the countries have conducted multiple DHS surveys to establish trend data that enable them to gauge progress in their programs. Countries that participate in the DHS program are primarily countries that receive USAID assistance; however, several non-USAID supported countries have participated with funding from other donors such as UNICEF, UNFPA or the World Bank.

MODEL QUESTIONNAIRES

The basic approach of the DHS program is to collect data that are comparable across countries. To achieve this, standard model questionnaires have been developed, along with a written description of why certain questions or sections have been included. These model questionnaires—which have been reviewed and modified in each of the four phases of the DHS program—form the basis for the questionnaires that are applied in each country. Typically, a country is asked to adopt the model questionnaire in its entirety, but can add questions of particular interest. However, questions in the model can be deleted if they are irrelevant in a particular country.

DHS surveys are designed to collect data on marriage, fertility, family planning, reproductive health, child health, and HIV/AIDS. Due to the subject matter of the survey, women of reproductive age (15–49) are the focus of the survey. Women eligible for an individual interview are identified through the households selected in the sample. Consequently, all DHS surveys utilize a minimum of two questionnaires—a Household Questionnaire and a Women's Questionnaire.

The Household Questionnaire is used to list all the usual members and visitors in the selected households. Some basic information is collected on the characteristics of each person listed, including his/her age, sex, education, and relationship to the head of the household. The main purpose of the Household Questionnaire is to provide the mechanism for identifying women eligible for individual interview and children under five who are to be weighed, measured, and tested for anemia. In addition, information is collected about the dwelling itself, such as the source of water, type of toilet facilities, materials used to construct the house, ownership of various consumer goods, and use of iodized salt.

DHS questionnaires for women distinguish between countries with high and low contraceptive prevalence rates. The DHS Model "A" Women's Questionnaire is for use in the high contraceptive prevalence countries, while the DHS Model "B" Women's Questionnaire is for use in countries with relatively low contraceptive use. The main difference between these questionnaires is that the "A" core collects considerably more information on family planning than the "B" core. Both versions of the Women's Questionnaire cover the following sections:

- Background characteristics (age, education, religion, etc.)
- Reproductive history
- Knowledge and use of contraceptive methods
- Antenatal care, delivery care and postnatal care
- Breastfeeding and infant feeding practices
- Immunization, child health, and nutrition
- Marriage and recent sexual activity
- Fertility preferences
- Knowledge about HIV/AIDS and other sexually transmitted diseases
- Husband's background and respondent's work

It was also recognized that some countries have a need for special information not contained in the core questionnaires. To accommodate this need and to achieve some level of comparability across countries that applied them, optional questionnaire modules were developed on a series of topics, as follows:

- Female genital cutting
- Maternal mortality

- Pill-taking behavior
- Sterilization experience
- Consanguinity
- Verbal autopsy

- HIV/AIDS
- Children's education
- Women's status
- Domestic violence
- Malaria
- Health expenditures

OTHER BASIC DOCUMENTATION

To achieve comparable information across countries, it is necessary to ensure that the questionnaires and the survey procedures followed in each country are similar. Therefore, the DHS program has developed a set of basic documentation to go with the model questionnaires. The basic documentation consists of the following manuals, most of which are available in English, French, and Spanish:

- Interviewer's Manual
- Supervisor's and Editor's Manual
- Sampling Manual
- Household Listing Manual
- Guidelines for DHS Interviewer Training
- Data Processing Guidelines
- Guidelines for the Main Survey Report
- Communicating DHS Data

The first two manuals provide a detailed description of the survey questions, tips on conducting interviews and how to check completed questionnaires. The Sampling Manual presents the DHS approach to issues like optimum sampling frames, sample domains, stages, and sample selection. The companion Household Listing Manual describes how to locate selected sample points, how to draw a sketch map, and how to list the households and structures.

The Guidelines for DHS Interviewer Training is designed for survey managers to provide tips on how to organize and conduct training for field staff. It describes techniques of mock interviewing, demonstration interviews in front of the class, and field practice, and contains sample tests for trainees. It also describes how to interpret the field check tables that are produced during the fieldwork.

The Data Processing Guidelines were developed to assist DHS staff and local collaborators in developing data processing procedures. While the quality of the data is determined mainly by the quality of the fieldwork, following appropriate steps can enhance it significantly during data processing. Data entry and editing for inconsistencies are major steps in this process, as is the imputation of missing data.

The Guidelines for the Main Survey Report detail the tabulations that will be produced on the basis of the survey data and provides some insight into the interpretation of these tabulations. This manual also aids data processing staff in determining the exact tabulations that are required for the survey. Communicating DHS Data is a manual with tips on how to effectively present DHS data to various audiences through computerized slide presentations, reports, fact sheets, maps, etc.

REPORTS AND DATA

The survey results for each participating country are published in a brief preliminary report, a more detailed final report and a summary type report on key findings. The latter two reports are widely distributed and constitute the primary output of the project.

Further dissemination of survey data is achieved through the publication of analytical and other reports. Of particular relevance for program and policy purposes are the Comparative Report series. These reports provide information across survey countries and can contribute greatly to the policy debate through the exhaustive view they provide on a particular situation in a large number of countries. Analytical Reports are also published and provide rigorous analysis of survey data.

In the early 1980s, ORC Macro pioneered the development of an innovative survey research tool, the software package, Integrated System for Survey Analysis (ISSA). This software was specially designed to allow for hierarchical data entry, complex consistency checking, tabulation of survey data, and computation of sampling errors. Not only was ISSA used worldwide for all ORC Macro surveys, but it was also adopted for the processing of the Bolivia Integrated Household Survey funded by the World Bank, by the Rand Corporation for the processing of the Family Life Surveys in Indonesia, Bangladesh, and Malaysia, and by the Economic Commission for Europe for its Family and Fertility Surveys. A few years ago, ORC Macro joined with the U.S. Census Bureau to develop an integrated, Windows-based version of their software packages. The new program, CSPro, was launched recently and has already been used in the DHS surveys and censuses for several countries.

Another new tool that has been developed to facilitate the use of DHS data is the STATcompiler. This is a Web-based tool that allows users to build customized tables for DHS countries based on hundreds of indicators. The tool is also an excellent mechanism for policymakers to get quick access to specific information that may be needed for reference, presentations, speeches etc. The STATcompiler can be accessed through the DHS Web site (www.measuredhs.com).

DATA ARCHIVING

DHS believes that widespread access to survey data by responsible researchers has enormous advantages for the countries concerned and the international community in general. Therefore, DHS policy is to release survey data to researchers after the main survey report is published, generally within 12 months after the end of fieldwork. DHS maintains a data archive, with datasets available on the Internet through a process of electronic registration. The address is http://www.measuredhs.com.

The reformatting of each dataset into a standard recode file facilitates use of DHS data. This file standardizes the variable names, location, and value categories across countries and constructs many of the commonly used variables such as age in five-year groups. Because DHS surveys collect an enormous amount of information on different subjects for the household, women age 15–49, children under age five and often men age 15–59, these standard recode files are a particular advantage to cross-country analysis.

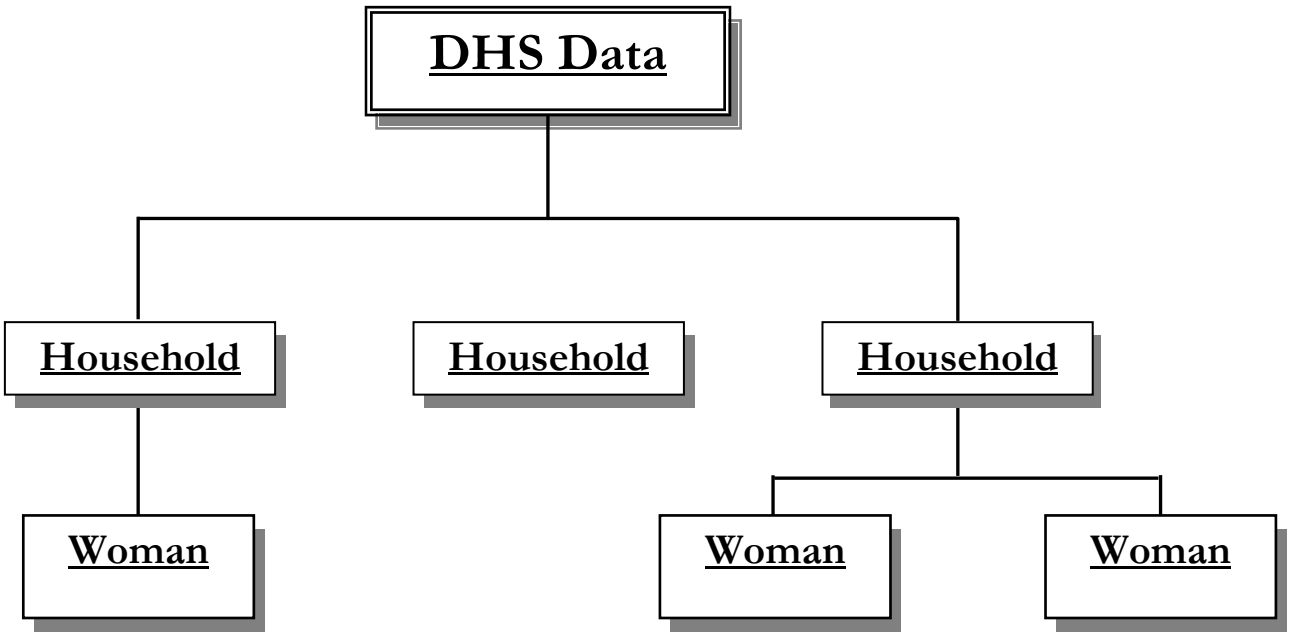
II. Overview of Key Aspects of the Organization of DHS Data

Structure of DHS Data Files

DHS uses a special software package, ISSA[1], to process its surveys. ISSA is specifically designed to meet the data processing needs of complex surveys such as DHS and one of its key features is its ability to handle hierarchical files. ISSA has been used at DHS in all steps of data processing with no need for another package or computer language. All steps, from entering the data into magnetic files to the production of statistics (including sampling errors) and tables published in DHS final reports, are done with ISSA. In addition, ISSA provides a mechanism to export data to the statistical packages SPSS, SAS and STATA. Data files exported using ISSA are stored using ASCII representation and they are provided in Flat or Rectangular formats. Along with the data file there is a syntax file describing the data file. The syntax files have extensions SPS, SAS or DO and DCT depending on the software the data file was exported to. DHS provides flat and rectangular files for households, women, children, and men and couples when the men questionnaire was included in a survey. These files are mainly in recode[2] format, but they can also be provided in raw data format.

Hierarchical Files

When requesting DHS files, users must specify the type of file(s) that they need. The following briefly describes the various file options.



[1] DHS is currently developing, in conjunction with the U.S. Census Bureau and SERPRO from Chile, a new software called CSPro. CSPro supports hierarchical files the same way ISSA does. CSPro has become the DHS standard data processing tool and it has already been used in several countries. Anything in this document referred to ISSA applies the same way to CSPro.

[2] The original country raw data is converted into a standardized format allowing easy comparison among countries or different DHS phases in the same country. Trevor Croft, *Description of the Demographic and Health Surveys Individual Recode Data File.* January 6, 1998.

The figure on the previous page illustrates the nature of the hierarchical structure of a typical DHS survey in which two separate questionnaires (household and woman) are used to collect information.

The example shows that, while a single questionnaire is always completed for each household in the DHS sample, the number of Women's Questionnaires that will be completed depends on the number of eligible women listed in the Household Questionnaire. In other words, for each Household Questionnaire there may be zero or several women questionnaires.

The hierarchical data file produced in ISSA has a two-level structure reflecting the relationship between the questionnaires; the Household Questionnaire is at Level 1 and the Women's Questionnaire(s) at Level 2. Within each ISSA level, there can be one or more different types of records. For example, records in the household level in the typical DHS file can be single (e.g., household characteristics) or multiple (e.g., household members listing). Thus, using the same file, it is possible to work with different units of analysis (households, household members, women and children). This makes the analysis of variables across different units easy.

The hierarchical structure defined by ISSA has several advantages and disadvantages. Among the advantages, the following can be highlighted:

- All the data is stored in just one ASCII file. Virtually all statistical packages can read ASCII files.
- Since all the data is stored in the same file, it is easy to maintain the integrity of the data in terms of data structure related to levels and records.
- The data file mirrors the paper questionnaire. Each section in the questionnaire can be defined as a record in the data file and only the information that is needed is present in the file.

The major disadvantage is that this structure can be easily handled only by ISSA, CSPro, or by a customized program written in low-level computer languages such as C, C++, FORTRAN, or Basic to name a few.

Flat Files

In a flat file there is one record for each case. All variables in each case are placed one after the other on the same record. The multiple or repeating records of the file are placed one after the other on the record, with the maximum number of occurrences of each section being represented in the data file. Each variable in a repeating section is placed immediately after the preceding variable of the same occurrence, such that all variables for occurrence 1 precede all variables for occurrence 2 of a section. The length of the record in the flat data file is fixed.

Rectangular Files

In a rectangular file each case in the data file contains a fixed number of records. Each record in the hierarchical file will have a representation in the rectangular file. For multiple or repeating sections there is a record for each occurrence of the section, with the maximum number of occurrences of each repeating section included in the data file. Essentially, the rectangular file looks the same as the hierarchical file but with blank records padded whenever the record does not exist in the hierarchical file or to complete the maximum number of records in repeating records.

Example 1 below shows how the structure of household data collected from a single household (Number 1) would vary in the hierarchical, flat and rectangular formats. As the example shows, the hierarchical data file would include two types of records for the household. Record type H1 is a single record with general household information, (e.g., on the source of water (SW) and toilet facilities (TF)). Record type H2 is a multiple record in which the same information (e.g., line number (L), sex (S) and age (A)) is entered for each members of the household. Additional records H2 can be added up to a set maximum.

Example 1. Data Representation for a Household Questionnaire in Selected File Formats

Data Structure	Records
Hierarchical ▪ One simple record H1 ▪ Three records H2	1H1 1 2
	1H2 1145
	1H2 2239
	1H2 31 8
Flat—One single record per case	1 1 2 1145 2239 31 8
Rectangular ▪ First record comes from H1 ▪ Following three records come from H2 ▪ Two blank records padded	1 1 2
	1 1145
	1 2239
	1 31 8
	1
	1

The flat file has only one record but blank space must be left at the end of the record to allow data to be entered for up to the maximum allowable number of members for each household. In the rectangular file, empty records also must be included to complete the maximum number of household members. In the above example, it is assumed that there would be a maximum of five members per household so that two blank records were included in the rectangular file.

Example 2 uses the household data described above in illustrating how the variable naming conventions differ in the various file formats. Variable names in repeating records of hierarchical data files need to be declared only once. In flat and rectangular files, since variables occupy different locations in the data file they need different names. When exporting data, ISSA adds the dollar ($) sign followed by the occurrence number to variables belonging to repeating records. If the maximum number of a repeating record is less than 10, ISSA will add just one digit. If it is 10 or more but less than 100, it will add two digits and so on.

Example 2. Variable Naming After Data Exporting

Data Structure	Records
Hierarchical	ID H1 WS TF
	ID H2 L S A
Flat	ID WS TF L$1 S$1 A$1 L$2 S$2 A$2 L$3 S$3 A$3 L$4 S$4 A$4 L$5 S$5 A$5
Rectangular	ID WS TF
	L$1 S$1 A$1
	L$2 S$2 A$2
	L$3 S$3 A$3
	L$4 S$4 A$4
	L$5 S$5 A$5

Before using an exported DHS data file, it is important to check for duplicate variable names. Software like SPSS allows only a maximum of eight characters for variable names. If variable Q1005A is defined in a record with a maximum number of occurrences of 20, it will be exported as Q1005A$01, Q1005A$02, Q1005A$20. When this syntax is read by SPSS it will chop the last character with the result of nine ($01 thru $09) variables with the same Q1005A$0 name.

Recode Files and Recode Variable Naming Conventions

There are three core questionnaires in DHS surveys: A Household Questionnaire, a Women's Questionnaire, and a Male questionnaire. There are also several standardized modules for countries with interest in those topics. Except for the Service Provision Assessment (SPA), which is a separate questionnaire in itself, all the other modules are applied as part of the household, women, or male questionnaires.

Since the very beginning of DHS a recode file was designed. There were two outstanding reasons to develop a recode file, to define a standardized file that would make cross-country analysis easier and to compare data with the World Fertility Surveys (WFS) to study trends. In DHS-I the recode was defined only for the Women's Questionnaire. The reason was that in WFS there was not a recode file defined for the Household Questionnaire and on the other hand there were very few countries that used a male questionnaire. The recode file proved to be very useful and as a result since DHS-II, a recode file, was introduced for the household and male questionnaires.

DHS questionnaires have changed extensively since the first phase. For this reason there is a different recode definition for each DHS phase. However, if a variable is present in one or more phases, that variable has the same meaning in wherever phase it is present. If a question is dropped from one phase to another, the name of the variable used for that question is not reusable. The variable will not be present in the recode definition of the phase were it was dropped. If a new question is added to the core questionnaire a new variable will be added to the recode definition.

Recode files are initially created using a hierarchical model and later exported to flat or rectangular files. There are two physical recode hierarchical data files. The first one includes the household and Women's Questionnaire and the second one is for the male questionnaire. The hierarchical data file is broken down into a number of records. The records were originally designed to map different sections of the model questionnaires, but because of changes among phases that may not be the case anymore. Some of these records are repeating or multiple-occurrence records while others are single-occurrence records. Single records contain simple, single-answer variables. Multiple records are used to represent sets of questions that are repeated for a number of events. The birth history is an example of a multiple section, where each child has an entry in the birth history. Multiple records are used for sets of questions where the number of occurrences may vary.[3]

There are special records to keep variables that are not part of the model questionnaires but were included in a particular country. These records are known as country-specific records and they can also be multiple or single depending on whether the question was added to a single or multiple section in the questionnaire.

[3] A full description of recode files can be found in the document *Description of the Demographic and Health Surveys Individual Recode Data File* by Trevor Croft.

Based on the two recode files mentioned above, DHS creates and distributes files in flat or rectangular formats with SPSS, SAS and STATA data definitions for several units of analysis. Of course, hierarchical data files are also distributed for users familiar with ISSA or CSPro. The data can be downloaded from the DHS Web site or can be requested by mail. The following table shows the files available along with the names that they are given for the Colombia 2000 Measure/*DHS*+ survey.

	File Types		
Unit of Analysis	**Hierarchical**	**Rectangular**	**Flat**
Household	COHR41	COHR41RT	COHR41FL
Women	COIR41	COIR41RT	COIR41FL
Men	COMR41	COMR41RT	COMR41FL
Children		COKR41RT	COKR41FL
Couples		COCR41RT	COCR41FL

In general, the first two digits of the file name correspond to the country code. The next two digits identify the unit of analysis (HR–Household, IR–Women, MR–Male, KR–Children, and CR–couples). The fourth digit identifies the DHS phase. The fifth digit identifies the data release number (when a new version of the country recode file is created a new release number is assigned). Finally, the last two digits identify whether it is a rectangular (RT) or flat (FL) file; for the hierarchical file they are left blank.

Variables in the recode file begin with one or two letters followed by one, two, or three digits and in some cases followed by a letter. Following is a list describing the general variable name conventions.

- HVxxx - Household standard variables
  - HAxx - Anthropometry and anemia for women
  - HCxx - Anthropometry and anemia for children
  - SHxxx - Household, country-specific.
- Vxxx - Women standard variables
  - Bxx - Birth history
  - Mxx - Pregnancy, postnatal care, and breastfeeding
  - Hxx - Immunization and health
  - HWxx - Anthropometry for children of eligible women
  - MMxx - Maternal mortality (optional)
  - DVxx - Domestic violence (optional)
  - Sxxx - Women, country-specific
- MVxxx - Men standard variables
  - SMxx - Men, country-specific

In this list "xx" represent digits, and the last one could be a letter. There are a handful of variables that do not fully adhere to this scheme (BORD, BIDX, MIDX, HWIDX, etc.), but at least the first letter will indicate where it belongs. The "xx" in the household, women, and men country-specific variables correspond to the actual question number in the country questionnaire. The maternal mortality (MMxx) and domestic violence (DVxx) variables will only be present if the modules were applied in the country.

If the calendar was used in the country, variables VCOL and VCAL will be present. These variables are part of a repeating record. VCOL is the calendar column number and VCAL is a string of 80 characters containing the actual calendar information.

Missing Values and Other Special Codes

In DHS, a missing value is defined as a variable that should have a response, but because of interview errors the question was not asked. For example, questions 111 and 112 must be answered by women with no education or only primary school education. If the interviewer incorrectly marked question 110 as secondary school or higher when in fact the respondent only completed primary school, questions 111 and 112 do not get asked. The general rule for DHS data processing is that "UNDER NO CIRCUMSTANCES AN ANSWER SHOULD BE MADE UP;" instead a missing value will be assigned. The data will be kept as missing in the data file and no imputation for this kind of questions is assigned. As will be discussed below, only the seven dates related to women's chronological events are imputed.

Missing values in general are codes 9, 99, 999, 9999, etc. depending on the variable number of digits. There are however some important background variables where the missing code is not accepted:

- Geographical variables such as Urban/rural (HV025, V025, MV025, V102, MV102), regions (HV024, V024, V101, MV101), and in general any other variable whose value can be established by the sample design
- Level of education for women and men in the individual questionnaire (V106, MV106)
- Current use of contraception for women (V312)
- Current marital status of women (V501)
- Variables related to the woman's birth history (V201 to V210, B0, B4, B5, B9).

It is possible that earlier phases of DHS may have missing values in some of these variables.

Another special code used for data cleaning purposes is code "Inconsistent." This code is generally used by people in the secondary editing group, when a value or code is not plausible. For example, dates for vaccinations recorded as having occurred before the birth of the child. The value is not missing but is not possible. The secondary editing group is instructed to find out what piece of information is wrong (day, month, or year) and assign code "Inconsistent" to it. This of course is done after checking for clues that could lead to correct the problem.

Inconsistent codes are 7, 97, 997, etc., depending on the variable number of digits.

Codes 8, 98, 998 are assigned to "don't know" responses. These codes are normally pre-coded in the questionnaires, but they are consistently used throughout the recode file.

Finally, BLANKS in a variable represent not applicable for the respondent either because the question was not asked in a particular country or because the question was not asked of this respondent due to the flow or skip pattern of the questionnaire.

Missing, inconsistent, and “don’t know” codes must be excluded when calculating statistics such as means or medians; otherwise they will be treated as real values. For example, if they are not excluded to calculate the mean age at first sex, eventually the mean will be inflated by ages 97, 98, and 99. Multiple Response Variables

There are two types of multiple response variables in DHS questionnaires. In the first type, all possible responses are read to the respondent; a code will be assigned to each one of them. Example, question 412 goes as in the table below:

During this pregnancy, were any of the following done at least once?	**Yes**	**No**
Were you weighed?	1	2
Was your height measured?	1	2
Was your blood pressure measured?	1	2
Did you give a urine sample?	1	2
Did you give a blood sample?	1	2

At data entry time there are separate fields (variables) for each response. This question and all other variables of this type get translated to the recode into separate variables:

- M42A—Were you weighed?
- M42B—Was your height measured?
- M42C—Was your blood pressure measured?
- M42D—Did you give a urine sample?
- M42E—Did you give a blood sample?

The translation is done on a one-to-one basis. Code 2 will be assigned; code 0 and codes 1 and missing are kept with no modifications. This type of recoding is basically done for all variables with responses 1–Yes and 2–No throughout the recode file.

In the second type of multiple response variables, the question is asked and the respondent spontaneously begins providing answers. The answers are not read, but the interviewer has to classify the response according to the options available for the question. The interviewer keeps probing for other responses until the respondent says that she has no more answers. Question 407 of the core questionnaire is a typical example of this type of question:

Did you see anyone for antenatal care for this pregnancy?		
IF YES: Whom did you see? Anyone else?	Health Professional:	
	Doctor	A
	Nurse/Midwife	B
PROBE FOR THE TYPE OF PERSON AND RECORD ALL PEOPLE SEEN	Auxiliary midwife	C
	OTHER PERSON	
	Traditional birth attendant	D
	Other	X
	No one	Y

At data entry time, contrary to the first type of multiple response questions, all responses are stored in just one field. This makes data entry easy to control, and the data entry operator just type the letters that are circled. However, dealing with alphabetic codes presents a challenge for analysis purposes. To solve this problem, each response for this type of question gets translated into codes "0–No, 1–Yes, 9–Missing" in the recode as follows:

- M2A—Prenatal care: Doctor
- M2B—Prenatal care: Nurse/Midwife
- M2C— Prenatal care: Auxiliary Midwife
- M2D— Prenatal care: CS Health Specialist
- M2E—Prenatal care: CS Health Specialist
- M2F—Prenatal care: Trained Birth Attendant
- M2G—Prenatal care: Traditional Birth Attendant
- M2H—Prenatal care: Relative
- M2I—Prenatal care: CS Other Person
- M2J—Prenatal care: CS Other Person
- M2K—Prenatal care: Other Response (uncoded)
- M2L—Prenatal care: CS Other
- M2N—Prenatal care: No One

There are several issues that need to be taken into consideration with this type of variable:

1. If the question was missing, all applicable variables will get assigned code missing.
2. If a response does not exist for a particular country, the variable will be left coded BLANK.
3. There are always provisions to include country-specific (CS) responses that are not part of the core questionnaire. For example, some countries include "health workers" as a possible response.

Sampling Weights

DHS users should be aware that, in many cases, the data must be weighted. The following describes how DHS weights are constructed and when they should be used.

Definition

Sampling weights are adjustment factors applied to each case in tabulations to adjust for differences in probability of selection and interview between cases in a sample, either due to design or happenstance. In the DHS surveys, many times the sample is selected with unequal probability *to* expand the number of cases available (and hence reduce sample variability) for certain areas or subgroups for which statistics are needed. In this case, weights need to be applied when tabulations are made of statistics *to* produce the proper representation. When weights are calculated because of sample design, corrections for differential response rates are also made.

There are two main sampling weights in DHS surveys: household weights and individual weights. The household weight for a particular household is the inverse of its household selection probability multiplied by the inverse of the household response rate of its household response rate group. The individual weight of a respondent's case is the household weight multiplied by the inverse of the individual response rate of her individual response rate group. There may be additional sampling weights for sample subsets, such as male surveys, anthropometry, biomarkers, etc. There is only a need for the additional sample weights if there is a differential probability in selecting the subsamples. For example, if one in five households is selected in the whole sample for doing biomarkers, then an additional sample weight is not necessary. However, if one in five households in urban areas and one in two households in rural areas are selected, then an additional sample weight is necessary when estimating national levels or for any group that includes cases from both urban and rural areas. Notwithstanding the foregoing, the DHS has customarily included both household weights and individual weights to the men's surveys (modules), normalizing the weights for the number of households in the subset for the men's surveys, and to the number of men's individual interviews even when no differential subselection has been used.

Response rate groups are groups of cases for which response rates are calculated. In DHS surveys, households and individuals are grouped into sample domains and response rates are calculated for each domain.

Household Response Rate

A. Coverage: Excluded are dwellings without a household (no household lives in the dwelling, address is not a dwelling, or the dwelling is destroyed).

B. Numerator: Number of households with a completed household interview.

C. Denominator: Sum of number of households with a completed household interview, households that live in the dwelling but no competent respondent was at home, households with permanently postponed or refused interviews, and households for which the dwelling was not found.

Women's Individual Response Rate

A. Coverage: Women eligible for interview, usually women who are between the ages of 15 and 49 who slept in the household the night before the survey. In ever-married samples, women are eligible for interview only if they have ever been married or lived in a consensual union. In some surveys, the age range of eligibility has differed, e.g., all ever-married women age 12–49.

B. Numerator: Number of eligible women with a completed individual interview.

C. Denominator: Sum of number of eligible women with a completed individual interview, eligible women not interviewed because they were not at home, eligible women with permanently postponed or refused interviews, eligible women with partially completed interviews, eligible women for whom an interview could not be completed due to incapacitation and for other reasons.

Men's Individual Response Rates

Coverage: The age ranges and eligibility criteria has varied for men. Check with survey documentation.

Calculation

Initial sample weights are produced by the DHS sampler using the sample selection probabilities of each household and the response rates for households and for individuals. The initial weights are then standardized by dividing each weight by the average of the initial weights (equal to the sum of the initial weight divided by the sum of the number of cases) so that the sum of the standardized weights equals the sum of the cases over the entire sample. The standardization is done separately for each weight.

Handling of Missing Values—Not applicable

Application

Sample weights are calculated to six decimals but are presented in the standard recode files without the decimal point. They need to be divided by 1,000,000 before use *to* approximate the number of cases.

In tabulation programs, sampling weights need to be applied through the use of special commands.

Examples:

a) In SPSS using the WEIGHT command with the weight variable:
COMPUTE rweight = V005/1000000
WEIGHT by rweight.

b) In ISSA using the weight parameter
rweight = V005/1000000
x = xtab(table1, rweight).

Notes and Considerations

1. The sum of the sample weights only equals the number of cases for the entire sample and not for subgroups such as urban and rural areas.

2. Where there are no differential probabilities, weights may not be calculated since weights based just on response rates usually make little difference in results.

3. Use of sample weights is appropriate when representative levels of statistics are desired, such as percentages, means, and medians.

4. Use of sample weights is inappropriate for estimating relationships, such as regression and correlation coefficients.

5. Use of sample weights biases estimates of confidence intervals in most statistical packages since the number of weighted cases is taken to produce the confidence interval instead of the true number of observations. For oversampled areas or groups, use of the sample weights will drastically overestimate sampling variances and confidence intervals for those groups.

Century Month Code (CMC)

Century month codes (CMC) are calculated by multiplying by 12 the difference between the year of an event and 1900. That year was chosen as the reference period because all of the DHS relevant events occurred during the twentieth or twenty-first centuries. The month of the event is added to the previous result.

$$CMC = (Year – 1900) * 12 + Month$$

For example, the CMC for June 2002 will be:

CMC = (2002–1900) * 12 + 6 = 1230In other words, 1,230 months have elapsed since January 1900 to June 2002. Based on CMC it is possible to calculate the month and year using the following formulas:

$$Year = int((CMC – 1)/12) + 1900 \qquad [int(x) \text{ is the integer part of } x]$$

$$Month = CMC – ((Year – 1900) * 12)$$

Century month codes are particularly important to check consistency of dates, to calculate intervals between events, and to impute dates when the information for an event is missing or partially complete. The main DHS events, along with their corresponding recode variable numbers in the recode file, are the following:

- Respondent's date of birth (DOB, V011)
- Date of first union (DOM, V509)
- Date of birth for each of the respondent's children (DBCx, B3(1), B3(2), … B3(n))
- Date of current pregnancy (DOP, V214)
- Date of sterilization (DOS, V317)
- Date of interview (DOI, V008).

Where appropriate, the DHS file also includes variables for the month and year in which the events occurred. For example, V009 and V010 are the woman's month and year of birth. If an event was not experienced by the respondent, variables for that event will not exist or will be not applicable.

Example:

Let us see the events for a married, sterilized respondent with three births and with event dates that occurred as shown in the following figure. If a horizontal line is drawn from the woman's date of birth to the date of interview, all the events can be depicted in the line.

VARS	V011	V509	B3(1)	B3(2)	B3(3)	V317	V008
DATES	04/79	10/97	07/98	10/99	11/00	04/01	08/01
CMC	952	1174	1183	1198	1211	1216	1220
EVENTS	DOB	DOM	DBC1	DBC2	DBC3	DOS	DOI

DHS uses century month codes extensively during the process of editing and imputing data. The advantages of the approach include the following:

- When checking for consistency, use of the century month codes makes it easy to check not only that the events occurred in chronological order, but also that there should be a minimum interval between them. For example B3(2)—B3(1) should be greater or equal to 9 months (the duration of a pregnancy).
- For imputation[4] purposes, if information were missing between two events, the random imputation would be quite reasonable. For example, if date of birth for the second child is unknown, that birth should have occurred between the first birth plus nine months, and nine months before the third birth. The lower and upper limits for a random number generator are plausible.

Random (B3(1) + 9, B3(3) – 9) , Random (1192, 1202)

The use of CMC at the analysis phase facilitates the calculation of intervals or ages at different events. Throughout all DHS analysis computer programs, instructions like the ones shown below are very common.

Respondent's age	=	int((DOI – DOB)/12) (1220 – 952)/12 = 22 years old
Age at first birth	=	int((DBC1 – DOB)/12) (1183 – 952)/12 = 19 years old
Age at sterilization	=	int((DOS – DOB)/12) (1216 – 952)/12 = 22 years old
Age of last child in months	=	DOI – DBC3 1220 – 1211 = 9 months
Interval between birth 1 and 2	=	DBC2 – DBC1 1198 – 1183 = 15 months
Months since sterilization	=	DOI – DOS 1220 – 1216 = 4 months

DHS recommends that analysts use the century month code variables when dealing with intervals or ages at different events.

Median Calculations

There are four different types of median calculations in DHS statistics, and results vary according to the type of variable being analyzed.

Medians for completed time periods. These are medians for variables such as intervals between events or ages calculated at different events; for example, current age, age at first union, and age at sterilization. Medians for this type of variables take into consideration that ages are given in completed years. A respondent who is currently 20 years old could be somewhere between 20 years and 1 day old or 20 years and 364 days old.

Medians for continuous variables. These are medians for variables such as children's weight at birth or any other type of measurement in the continuous scale.

[4] Croft, T. 1991. Date Editing and Imputation. *Demographic and Health Surveys World Conference Proceedings, II:* 1337–1356, Columbia, Maryland: IRD/ORC Macro.

Medians for discrete variables. These types of medians apply to variables such as number of children, number of prenatal care visits, or in general any discrete variable where the only possible values are integers. A respondent can only have one, two, or any integer number of children. It is not possible to have 2.3 children.

Medians using the current status data. These types of medians are calculated for variables where 100 or close to 100 percent of the population have that characteristic at the beginning of an event and the percentages diminish as time passes by. For example, 100 percent of children do not know how to walk at birth. As time progresses, some children begin to walk, and there is an age (in months) where 50 percent or more of the children learn to walk.

The calculations for the first three types of medians are calculated in a similar manner and vary basically in the final result. Medians using the current status data are calculated using a different methodology. To illustrate the approach used in calculating the first three types of medians, consider the following table.

Variable	Cumulative Percentage
21	18
22	25
23	43
24	52
25	58
27	63

If the variable in question is current age in completed years, the median would be calculated for a completed period. In this case the interpolation will take place between ages 23 and 24. The result of the interpolation is (50 – 43)/(52 – 43) + 23 = 23.8, however, one year has to be added because, according to the completed year definition of age, the cumulative value of 52 percent was reached before completing 25 years of age.

If the variable were a continuous one, for example, time to a health facility in minutes (for purposes of simplicity, time is truncated to the minute). The interpolation would also take place between ages 23 and 24, but no adjustment is needed. The median time to a health facility would then be 23.8 minutes.

If the variable were of type discrete, the median would be obtained at 24 when 50 percent or more was reached.

Current Status Data

Current status data are used in DHS to calculate the median duration of breastfeeding, post-partum amenorrhea, and post-partum abstinence. Looking at how the median duration of breastfeeding is calculated (the same principle applies for amenorrhea and abstinence), information is first obtained on the proportion of children currently being breastfed (for breastfeeding, according to the children's age in months). For purposes of providing some stability to the proportions, the birth data are grouped into two-month intervals. Before calculating the proportions, the distribution is smoothed by a moving average of three groups.

Age	Total Children	Total Children Smoothed	Current Breastfeeding Smoothed	Proportion Still Breastfeeding
0–1	166.3	166.3	139.9	84.1
2–3	223.2	226.9	147.5	65.0
4–5	291.3	244.7	113.0	46.2
6–7	219.6	248.4	65.1	26.2
8–9	234.4	223.1	16.8	7.5
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
34–35	118.5	118.5	.	.
Median				4.1

To smooth the distribution by a three-group moving average, take the previous, current, and following value of the distribution and divide it by 3. For example, the smoothed total children for the age group 2–3 comes from: (166.3 + 223.2 + 291.3)/3. The first (0–1) and last (34–35) age groups cannot be smoothed, so they remain with the original values. The number of children currently being breastfed is shown after the data were smoothed. With the distributions smoothed, the percentages of children in each group are calculated.

The first age (duration) for which the proportion falls below 50 percent will be used for the calculation of the median by linear interpolation between the midpoint of that age group and the next youngest midpoint age. The median in this example falls between the age groups 2–3 and 4–5. Between these two age groups is when the transition from more than 50 percent (65.0%) to less than 50 percent (46.2%) of children still being breastfed occurred. The interpolation is then done between the midpoints for the age groups, which are 2.5 and 4.5.

The reason for taking the midpoint is that age is calculated by subtracting the date of birth of the child from the date of the interview. As a result, kids reported to be age x are in fact between exact ages x and x+1. A reasonable assumption is that these kids are on the average age x. A category consisting of reported ages x to x+1 has a midpoint at exact age (x+x+1)/2.

It should be noted that the midpoint for the first age group is calculated in a somewhat different manner. On the average, there are only about half as many children born in the month of the interview than in any other regular month. A reasonable age for children born in the month of interview is 0.25, assuming that interviews are uniformly distributed. Thus, the age average for kids born 0 to 1 months is calculated as (0.5 * 0.25 + 1)/1.5 = 0.75.

All Women Factors

Definition

Factors for adjusting ever-married samples to estimate statistics based on all women are as follows.

A. Coverage: All women factors are specific for individual years of age and specific for the particular subgroup to be estimated (e.g., there are two sets of factors by type of area: one for urban areas and one for rural areas).

B. Numerator: Number of all de facto women

C. Denominator: Number of de facto ever-married women.

Calculation

A. Numerator: From the household survey, tabulation by single year of age and by subgroup of the number of women 15–49 years of age of any marital status who slept in the household the night before the interview.

B. Denominator: From the household survey, tabulation by single year of age and by subgroup of the number of women age 15–49 years who slept in the household the night before the interview who have ever been married.

C. Tabulations use the household sampling weights.

D. If there are no or very few women ever married for a single age for the subgroup, then neighboring ages are combined to calculate the all women factor, using the same value for each of the single ages.

Handling of Missing Values

Women with missing or unknown age are excluded from both the numerator and the denominator. Women with unknown or missing marital status are considered never-married.

Notes and Considerations

Women in consensual unions and women separated from consensual unions are considered as ever-married.

The all women factors for ever-married samples are the inverse of the proportions ever married at the time of the survey.

The all women factors come about because of the selection process used in the design of the sample. Therefore it is only appropriate to use internally generated factors and not to use information external to the survey. For each subgroup to be estimated, the same factor is applied to each woman irrespective of the time period to be estimated since it is based on sample selection.

Application of All Women Factors

In tabulations, all women factors are treated as adjustments to the respondent sampling weight variable for each woman, multiplying the weight variable for the woman by her appropriate all women factor. The appropriate all women factor is that of the woman's reported household single year of age at the time of the survey for the subgroup to be tabulated.

Example:

The following table shows the number of women age 15–49 for all women, currently married women and the resulting all women factors according to place of residence (urban/rural) and total. The distribution was obtained from the household schedule for *de facto* women (women who slept the previous night in the household). The data were weighted using the household weights.

Age	All Women			Ever-Married Women			All Women Factors x 100		
	Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
15	50.7	413.65	464.35	0	43.55	43.55	884	950	1066
16	48.18	466.25	514.44	0	127.09	127.09	884	367	404
17	41.07	381.55	422.61	15.84	254.55	270.4	884	150	156
18	42.16	380.85	423.01	20.01	301.43	321.44	211	126	132
19	58.19	377.12	435.31	35.83	318.74	354.57	162	118	123
20	57.33	387.45	444.78	40.69	351.82	392.51	141	110	113
21	53.13	293.66	346.78	41.45	263.32	304.77	128	112	114
22	50.88	318.5	369.38	37.06	305.28	342.34	137	104	108
23	35.92	383.58	419.51	28.06	359.32	387.38	128	107	108
48	17.46	175.2	192.66	17.46	172.85	190.31	100	101	101
49	14.18	152.22	166.4	13.51	152.22	165.73	105	100	100

The "all women factors" are calculated by dividing the total number of women by the number of ever-married women for each characteristic (urban/rural and total) within each single age. The factor for women 15 years old living in rural areas is 9.5. This value is calculated by dividing 413.65 by 43.55. What this factor is saying is that an ever-married woman age 15 living in a rural area represents 9.5 women of the total women population in that area for that survey.

It is important to note that there are no ever-married women for ages 15 and 16 in the urban areas. *To* calculate the factors for those ages, the total number of women needs to be accumulated until an age where ever-married women are found (age 17 in this case). The result of this accumulation is then divided by the number ever-married women found.

(50.7 + 48.18 + 41.07)/15.84 = 139.95/15.84 = 8.84

The same factor is applied to women ages 15, 16, and 17 because 15.84 ever-married women represent the 139.95 accumulated women for those age groups.

Although these factors are calculated at the household level, they are normally used at the women's level (or men's levels for ever-married male surveys). To do that, each woman is assigned a factor for each characteristic for which the factors are calculated. The links to assign the factors are age and the characteristic. In our example a woman 18 years old, living in an urban area, would be assigned the factors 1.32 for the total and 2.11 for place of residence. When establishing the values for the characteristics to be used as links, they must be taken from the household level and not from the woman's level. The reason is that there may be differences from the data recorded in the household compared with the data recorded in the Women's Questionnaire. A different person than the respondent to the Women's Questionnaire may report age and education in the Household Questionnaire. However, when factors are calculated they need to use the household variables (because not all women have an individual interview). Another important issue is that factors can only be calculated for characteristics present in the Household Questionnaire. For example, factors for media exposure or knowledge of HIV/AIDS cannot be calculated because they are collected at the women's level.

DHS women's recode files include four standard variables containing all women factors:

- AWFACTT—All women factor, for total population
- AWFACTU—All women factor, for education (none, primary, secondary, higher)
- AWFACE—All women factor, for place of residence (urban/rural)
- AWFACTR—All women factor, for region (according to the country's regions).

These variables are five-digit variables with two implicit decimals, so they have to be divided by 100. Any table that is based on total population for ever-married samples should use these factors. All tables based on all women (fertility rates, age at first union, age at first sex, etc.) must use ever-married factors. If the background variable to be used is not one of the standard or country-specific "all women factors" they have to be created as described in this example.

Matching Data Files

As mentioned earlier, DHS distributes separately Household, Women, Children under five, Men, and Couples files in flat or rectangular formats. Special care has been taken to include all variables that are deemed important for each of these files. For example, variables for household characteristics are included in the women, men, and children's files. However, there are instances when researchers have to merge or combine different files to obtain the variables that meet their analysis needs. This section discusses the variables and mechanisms that can be used to accomplish that task.

One of the advantages of processing complex surveys with a software capable of handling hierarchical files is that it allows to tightly control the case identifiers. DHS guarantees that their files can be matched seamlessly whenever a relationship is possible. To properly manipulate the files it is necessary to know what the variables or fields that identify the cases are. The following table shows those fields.

Unique Identifiers for Data Files

File	ID Variable	Cluster	HH Number	Line Number	Birth Order	Husband/ Wife
Household	HHID	HV001	HV002	HVIDX		
Women	CASEID	V001	V002	V003		V034
Men	MCASEID	MV001	MV002	MV003		MV034*i*
Children	CASEID	V001	V002	V003	BIDX	
Couples	CASEID	V001	V002	V003		

Column "ID Variable" is an alphabetical variable that uniquely identify a case. This variable is normally constructed by concatenating variables cluster and household number for the household file and cluster; household number; and line number for women, men, and couples. In the case of children, variable CASEID is the same as that of their mother plus a consecutive number to differentiate among children born in the last five years of the same mother. CASEID for couples is that of the woman (as opposed to MCASEID the man identifier) because in polygamous countries a man can be the partner for more than one woman.

When merging files it is probably easier to do so using the original variables rather than the ID variables. For example,

it is not possible to merge the household and women's files using HHID and CASEID because CASEID has three extra characters identifying the women's line number. The files can be more easily merged using variables HV001 with V001 and HV002 with V002.

Relationships

When merging data files it's important to know the type of relationship that exists between the files to be merged as well as the type of output file desired (unit of analysis). There are two types of relationships: The first is that of one entity related to many other entities [1 : 0-*N*] and the second is that of one entity related to just one other entity [1 : 0-1]. An example of a relationship of one to many entities can be found between households and women or men. There may exist zero or several women or men questionnaires for each household (see structure figure at the beginning of the document). An example of a relationship of one to one can be found in the relation existent between women and men. In a monogamous country, there may be zero or one man questionnaire for each woman if she is currently married.

All statistical packages (SPSS, SAS, STATA) have commands that allow merging files, but regardless of the package the following steps are necessary:

- Determine the common identifiers (identification variables).
- Sort both data files by the identification variables.
- Determine the base (primary) file. The base file essentially establishes the unit of analysis.
  - Normally, when the relationship is that of one too many [1:0-*N*], the base file is the one with the many entities. For example, if merging data from households and women, the base file should be the women's file. The reason is that you may want to assign to every woman the characteristics of her household. If the match is done the other way around, once the program matches the first woman it will not look for another woman or it will give an error for finding duplicate cases. In the case of matching women and children, the base file should be the children's file. That way, mothers' characteristics are assigned to children.
  - If the relationship is that of one to one [1:0-1], the base file is normally the one with the least number of cases. In DHS, male questionnaires are only applied to a subsample of households. This means that not all currently married women have a match with a male questionnaire. In this case, the base file should the male questionnaire and the resulting file (unit of analysis) will be the Couples file.
- Finally, using the right commands depending on the software to be used, the files will be merged.

The following table shows the variables required to match different files. In the rows, the base files are listed. In the columns, the secondary files along with the variables to be used as keys or matching variables are listed. In the cells intersecting the rows and columns, variables from the base files used to match the secondary file are listed.

Matching Variables

	Secondary Files	
Base	Match variables for Households HV001 + HV002	Match Variables for women V001 + V002 + V003
Women	V001 + V002	
Children	V001 + V002	V001 + V002 + V003
Men	MV001 + MV002	Couples MV001 + MV002 + MV034*i*

This table shows that household variables can be appended to women, men and children. Women variables can be appended to their children. They also can be appended to men, to create couples. Notice that there is no relationship between children and men because children come from the birth history, which is asked to women. It is important to mention that matching files is only necessary when variables required for the analysis are not present in the distributed file but are present in any other file.

Normalization of Household Members and Children

There are several other potential files that are not distributed directly as separated files, but their data are included as part of the household or women files. The first of these potential files is a household member file. Variables related to household members are included in the household flat or rectangular files. The information is stored with variables for person *i*, preceded by variables for person *i-1*, and followed by variables for person *i+1*. These set of variables are repeated for as many times as the maximum number of household members defined by variable HV009. Schematically, the way the variables are stored looks as follows:

General Variables	Variables Member 1	Variables Member 2	Variables Member *n*	Household Characteristics Variables

However, if the unit of analysis is the household members, this type of file is not the most appropriate. Instead the file has to be normalized, with each person occupying a separate record, as shown in the following figure:

General Variables	Variables Member 1	Household Characteristic Variables
	Member 2	
	Member *n*	

The way to normalize the file depends on the software to be used. In SPSS, one way to do it is by writing each member to a separate file and then appending the resulting files together. There are similar commands in other statistical packages that can be used to accomplish the same task.

Example:

- *Create a file for each member.
  - Get file = ‘hhflat.sav’.
  - Select if (not(missing(HVIDX$01))).
  - Save outfile = ‘member01.tms’
    - /keep = HV000 to HV044, HV101$01, HVIDX$01, HV102$01, HV103$01$01, V201 to HV234

- /rename = (HVIDX$01, HV101$01, HV102$01, HV103$01 = HVIDX, HV101, HV102, V103).

- Get file=‘hhflat.sav’.
- Select if (not(missing(HVIDX$02))).
- Save outfile = ‘member02.tms’
  - /keep = HV000 to HV044, HVIDX$02, HV101$02, HV102$02, HV103$02, HV201 to HV234
  - /rename = (HVIDX$02, HV101$02,HV102$02,HV103$02=HVIDX,HV101,HV102, HV103).
- Get file = ‘hhflat.sav’.
- Select if (not(missing(HVIDX$n))).
- Save outfile = ‘membern.tms’
  - /keep = HV000 to HV044, HVIDX$n, HV101$n, HV102$n, HV103$n, HV201 to HV234
  - /rename = (HVIDX$n, HV101$n, HV102$n, HV103$n = HVIDX, HV101, HV102, HV103).

- *Join the newly created files for each household member together*.
  - Get file = ‘member01.tms’.
  - Add files
    - /file = *
    - /file = ‘member02.tms’.
  - Add files
    - /file = *
    - /file = ‘membern.tms’.
- Save outfile = ‘MEMBERS.SAV’.
- * Delete working files
    - Erase file = ‘member01.tms’.
    - Erase file = ‘member02.tms’.
    - Erase file = ‘membern.tms’.

In this example, HV000 to HV044 are "General" variables and HV201 to HV234 correspond to "Household Characteristics". For each household member, a number of variables is selected (HVIDX, HV101, HV102, HV103. All household member variables can be included (if desired) along with the general and household characteristic variables. The member-specific variables are renamed so that later the files will be appended correctly. In this example, only three household members are included, but the different pieces of the program need to be repeated as many times as the maximum number of household members specified by variable HV009.

This type of file can now be used to determine the population distribution in the sample or any other statistic related to household composition. One special use is to calculate all women factors for ever-married samples as described in previous sections.

The same principle used to create household member files, can be used to create children files out of the women flat or rectangular files. As mentioned earlier, the children file distributed by DHS only includes children born in the last five years. The reason is that variables related to prenatal/postnatal care and children's health are only available for children born in the last five years. However, some analysis, such as mortality rates for the last 10 years, require the women's full birth history. Contrary to the household members file where the number of members is variable across countries, the maximum number of births in all DHS files is constant (20) for all women. Variables BIDX$01-B16$01 to BIDX$20-B16$20 are the birth history variables in DHS files and are always present in women files.

In previous DHS phases (DHS-I to DHS-III), data on anthropometry and anemia for children and women was collected in the Women's Questionnaire. In DHS-IV, that information is collected in the Household Questionnaire. Therefore, the anemia and height and weight data for women and children are distributed in the household flat or rectangular files. However, to facilitate the comparability across DHS phases the height and weight and anemia-related variables were brought down from the household level to the woman level. Variables V436 to V447 (women's height and weight) and HWIDX to HW58 (children's height and weight) continue to be part of the DHS-IV woman's file and are comparable with previous phases. Anthropometry and anemia variables for foster children and children whose mothers have died, are only present at the household level. The same is true for never-married women in ever-married samples.

Similarly to how the members file was created, it is also possible to create files for anthropometry and anemia for women and children out of the household flat or rectangular files. Once those files are created, the relationships shown in the table below can be established. In that table, the women and men are the original files distributed by DHS. The children file can be either the DHS-distributed file or a children file created using the birth history.

	Secondary Files Based on Household Data		
	Members Match Variables	**Height, Weight, and Anemia for Women**	**Height, Weight, and Anemia for Children**
Base	HV001 + HV002 + HVIDX	HV001 + HV002 + HA0	HV001 + HV002 + HC0
Women	V001 + V002 + V003	V001 + V002 + V003	
Children	V001 + V002 + B16		V001 + V002 + B16
Men	MV001 + MV002 + MV003		
Members		HV001 + HV002 + HVIDX	HV001 + HV002 + HVIDX

III. Guide to Individual Statistics

FERTILITY

CURRENT FERTILITY RATES

Statistics: ***Age-Specific Fertility Rate (ASFR) Also Known As Age-Period Fertility Rate in the Three Years That Precede the Date of the Survey, Presented as an Annual Rate***

Definition

A. Coverage: All women age 15–49 years in seven five-year age groups (15–19, 20-24, 45–49 years).

B. Numerator: Number of births that occurred in the 1–36 months before the survey to women in the age group at the time of the birth.

C. Denominator: Number of women-years of exposure in the 1–36 months before the survey of women in the age group.

D. Rate is the quotient of the numerator divided by the denominator.

Calculation

A. **Numerator:** Births are tabulated according to period of birth and the age of mother at the time of the birth.

1. **Period of birth:** The period of birth is calculated as the difference in months between the date of interview and the date of birth, both in century-month code format (CMC). Births are included in the tabulation if they occur 1–36 months before the survey.

2. **Age of mother at the time of the birth:** The difference in months between the date of birth of the child and the date of birth of the mother both in CMC. The difference is then divided by 60 and truncated to whole numbers to form the age groups. Births are tabulated by age group.

B. **Denominator:** Women-years of exposure are calculated as the sum of the number of months exposed in the five-year age group during the time period divided by 12. A woman can contribute exposure to just two five-year age groups during the 36-month period.

1. Higher age group: A woman's age at the end of the period determines the higher age group. The high age group is calculated by subtracting the women's date of birth from the date of interview minus one (in CMC), dividing the difference by 60 and truncating to a whole number. The number of months spent in the higher age group is the difference in months between her age at the end of the period of exposure (date of interview less one month) and the lower age limit of the age group plus one month. If the number of months in the age group is less than the duration of the time period (36 months), then the woman contributes exposure to both the higher age group and the next lower age group.

2. Lower age group: The contribution to the lower age group is 36 less the number of months in the higher age group. If the number of months in the higher age group is greater than or equal to the duration of the time period (i.e., ≥ 36 months), then the exposure in the higher group is the duration and the exposure in the lower age group is zero.

    a) **Tabulation:** Each woman is tabulated twice, once according to her higher age group accumulating the exposure she contributes to that group and once in the lower age group accumulating lower age group exposure. (In ISSA, the same table is used, effectively summing the accumulations within each age group.) For ever-married samples, the exposure is adjusted to represent all women by multiplying by the woman's "all-woman factor" (AWFACTOR), which is derived from the proportion of ever-married women from the household data file. See section on all women factors for details in their calculation.

3. Examples

    a) **Example 1:** A woman interviewed in December 2001, born in May 1970. Her CMC date of interview is 12*(2001 – 1900) + 12 = 1224. The date of the end of the period of exposure is 1224 – 1 = 1223. Her CMC date of birth is 12*(1970 – 1900) + 5 = 845. Her age in months at the end of the period is 1223 – 845 = 378. The age group at the end of the period is 378/60 = 6.3, truncated to 6. This represent age group 30–34 years (30 = 6*5 years interval). The number of months in this age group is 378 – 6*60 + 1 = 19 months. Since this is less than the total number of months during the period (36 months), she contributed 19 months to age group 30–34 during the period and 36 – 19 = 17 months to the age group 25–29 during the period.

    b) **Example 2:** A woman interviewed in December 2001, born in March 1967. Her CMC date of interview is 12*(2001 – 1900) + 12 = 1224. The date of the end of the period of exposure is 1224 – 1 = 1223. Her CMC date of birth is 12*(1967 – 1900) + 3 = 807. Her age in months at the end of the period is 1223 – 807 = 416. The age group at the end of the period is 416/60 = 6.93, truncated to 6. This represent age group 30–34 years (30 = 6*5 years interval). The number of months in this age group is 416 – 6*60 = 56 months. Since the number of months in this age group is greater than 36 months, she contributed 36 months of exposure to age group 30–34 during the period and no exposure to the next lower age group during the period.

    c) **ASFR:** The age-specific fertility rate is calculated as the quotient of the numerator divided by the denominator for each age group, multiplied by 1000. The result is an average rate over the 36-month period, expressed as an annual rate.

Handling of Missing Values

The total number of children to which a woman has given birth is recorded obligatorily by the interviewer; no unknown numbers children are allowed. There are three values involved in the calculation of ASFR, interview date, birth date of woman and birth dates of children. The interview date is always known from fieldwork dates. If missing or unknown, the birth dates of interviewed women and her children are imputed before formation of the standard recode file. See Croft, 1991 on date imputation.

Notes

Births to women at ages less than 15 years or more than 49 years at the time of the birth are not generally included. In a few specific countries, births to girls 10–14 are included.

Births in the month of interview are excluded. This exclusion is because this month does not represent a full month but is censored by the date of interview.

A three-year (36 month) time period is taken for calculating current AFSR. This period is a compromise between the need for recency and reduction of sampling variation. This time period was selected during the World Fertility Survey, when sample sizes were on usually about 5,000 women. For comparability over time and across surveys, this period has been maintained by DHS.

No adjustment is made for truncation by age. (Women who are at most 49 years at the time of interview were 48 years the year before and 47 years two years before.) The reason no adjustment is made is that the tiny probability of giving birth by women 48 and 49 years of age outweighs the complication of doing the adjustment by single years of age.

In line with general DHS policy, no adjustment is made for possible omission or date misreporting of the dates of birth of children or misreporting of the date of birth of the woman.

For ever-married samples, it is assumed that never-married women have not had any births. Only the denominator of the rates is adjusted to estimate the number of all women.

References

Croft, T. 1991. Date Editing and Imputation. *Demographic and Health Surveys World Conference Proceedings, II*: 1337–1356, Columbia, Maryland: IRD/ORC Macro.

Statistics: *Trends in Age-Specific Fertility Rates*

Definition

Age-specific fertility rates are calculated for five- or four-year periods of time preceding the interview.

Calculation

The calculation of ASFRs for trends is the same as the calculation of current ASFRs, except that it is done for either five-year periods preceding the survey (1–60, 61–120, 121–180 and 181–240 months before the interview date) or four-year periods (1–48, 49–96, 97–144, 145–192, and 193–240 months before the interview date).

Handling of Missing Values

See current age-specific fertility rates.

Notes and Considerations

Since the time periods are five years or less in duration, the calculation of women-years of exposure is the same as that for current age-specific rates with the total duration within each period per woman of 60 or 48 months for five-year and four-year periods, respectively.

For five-year time periods, the following ASFRs will be truncated:

Period	Age Group
0–4 years	45–49 years
5–9	40–44
10–14	35–39
15–19	30–34

For four-year time periods, the following ASFRs will be truncated:

Period	Age Groups
0–3 years	45–49 years
4–7	40–44, 45–49
8–11	35–39, 40–44
12–15	35–39
16–19	30–34, 35–39

The four-year period is used where birth displacement at five years of age is thought to be important.

Statistics: *Total Fertility Rate*

Definition

The total fertility rate (TFR) is an age-period fertility rate for a synthetic cohort of women. It measures the average number of births a group of women would have by the time they reach age 50 if they were to give birth at the current age-specific fertility rates. The TFR is expressed as the average number of births per woman. Unless otherwise specified, the TFR is for all women. For current fertility rates, the DHS survey uses the period 1–36 months before the survey.

Calculation

The TFR is the sum of the age-specific fertility rates for all women multiplied by five. The ASFRs are those for the seven five-year age groups from 15–19 to 45–49.

Handling of Missing Values

Not applicable.

Notes and Considerations

Variants:

The TFR for 15–44 is also presented. It uses the 6 five-year age groups from 15–19 to 40–44.

For countries with very early fertility, the ASFR for age group 10–14 years may be included.

Statistics: *General Fertility Rate*

Definition

The general fertility rate (GFR) is the average number of children currently being born to women of reproductive age.

A. Coverage: All women 15–44 years of age during the three years preceding the survey.

B. Numerator: Total number of births in period.

C. Denominator: Total number of women-years of exposure during period.

Calculation

A. Numerator: The sum over all the five-year age groups of the births tabulated for the age-specific fertility rates.

B. Denominator: The sum over all the five-year age groups from 15–19 to 40–44 of the women-years of exposure tabulated for the age-specific fertility rates.

Handling of Missing Values

Not applicable.

Notes and Considerations

All births in the period are included.

For ever-married samples, all women factors are applied in the tabulation of the denominators.

References

Age-specific fertility rates.

Statistics: *Crude Birth Rate*

Definition

The crude birth rate (CBR) is the annual average number of births per thousand of the whole population.

A. Numerator: Total number of births in the three years preceding the survey.

B. Denominator: Total number of person-years of exposure during the three years preceding the survey.

Calculation

A. The crude birth rate is calculated by summing the product of the current age-specific fertility rates by the proportion of women in the specific age group out of the total de facto population.

B. Proportion of women in the specific age group.

C. Numerator: From the household survey, the number of women who slept in the household the night before the interview tabulated by five-year age group (according to age reported in the household schedule) for groups 15–19, …, 45–49.

D. Denominator: From the household survey, the total number of people of both sexes and all ages who slept in the household the night before the interview.

Handling of Missing Values

Sex of household member is always assigned. It may be possible that in early DHS surveys sex had missing values. In that case the household member was included in the denominator.

Women with unknown age are assumed to be outside the 15–49 year range; thus they are not tabulated for numerators.

People of either sex with unknown age are included in the denominators.

Notes and Considerations

Assumptions—Due to the lack of birth dates for the household population, the proportion of women in each age group at the time of the survey is assumed to be the same as the proportion at the midpoint of the time period.

Statistics: *Percentage Currently Pregnant*

Definition

Proportion of all women 15–49 years of age who report themselves as pregnant at the time of interview expressed as a percentage.

A. Coverage: All women 15–49 years of age.

B. Numerator: Number of women who report themselves pregnant at the time of interview.

C. Denominator: Number of all women.

Calculation

A. Numerator: Tabulation of interviewed women who say they are pregnant at the time of the survey.

B. Denominator: In all women samples, tabulation of all women. In ever-married samples, tabulation of women using all women factors.

Handling of Missing Values

Women with missing and "don't know" responses are considered not pregnant. Women who do not know or are unsure whether they are pregnant are excluded from the numerator but are included in the denominator. Women with missing data on whether or not they are pregnant are excluded from the numerator but are included in the denominator.

Notes and Considerations

The percentage of women currently pregnant is underreported because women who are in their early stage of pregnancy may not yet know if they are pregnant and because some women may not want to declare that they are pregnant. In ever-married samples, never-married women are assumed not to be pregnant.

Statistics: *Percent Distribution by Number of Children Ever Born*

Definition

Percentage of women with specified number of children ever born.

A. Coverage:

1. All women.

2. Currently married women—Include formal and married or living in a consensual union women only.

B. Numerator: Number of women who in their lives have given birth to the specified number of children.

C. Denominator: Number of women in sample.

Calculation

For each specified number of children, 100 times the quotient of the numerator divided by the denominator. In ever-married samples, the denominator is adjusted by the all woman factor.

Handling of Missing Values

Not applicable—Number of children and marital status are not allowed to have missing values.

Notes and Considerations

Percentages add up to 100 percent.

Statistics: *Mean Number of Children Ever Born and Mean Number of Living Children*

Definition

Arithmetic average of children, both surviving and dead, to whom women in the sample have ever given birth.

A. Coverage:

1. All women.

2. Currently married women—Include formal and married or living in a consensual union women only.

B. Numerators:

1. Mean number of children ever born: Sum of number of children surviving and dead at the time of the survey who were given birth by women in the sample.

2. Mean number of living children: Sum of number of children alive at the time of the survey who were given birth by women in the sample.

C. Denominator: Number of women in sample. Same for both statistics.

Calculation

Numerator divided by denominator. In ever-married samples, the denominator is adjusted by the all woman factor.

Handling of Missing Values

Not applicable—Number of children and marital status of women and children's survival status are not allowed to have missing values.

Notes and Considerations

Expressed to two decimal places.

BIRTH INTERVALS

Statistics: *Percent Distribution by Months Since Preceding Birth*

Definition

Percentage of births in the five years preceding the survey by specified grouped number of months since the preceding birth.

A. Coverage: All second- and higher-order births that occurred in the 0–59 months preceding the interview.

B. Numerator: Number of births whose interval with the preceding birth is within the specified group.

C. Denominator: Total number of births.

Calculation

A. Covered births: Children both surviving and dead who have a difference between the date of interview and their birth date (both in CMC format) of 0–59 months, excluding first-born children.

B. Preceding birth interval: Difference between birth date of child and birth date of preceding child (in CMC format) grouped into categories. For children of multiple births, the birth date of the preceding child is the number of months since the end of the preceding pregnancy that ended in a live birth.

C. Numerators for each birth interval category are divided by the same denominator and multiplied by 100 to obtain percentages.

Handling of Missing Values

Not applicable—Birth dates of children are imputed if missing; therefore birth dates do not have missing values.

Notes and Considerations

Percentages add up to 100 percent.

First-order births (and their twins) are excluded from both numerators and denominators.

Births in the month of interview are included and births 60 months before the interview are excluded.

Statistics: *Median Number of Months Since Preceding Birth*

Definition

Median number of months since the preceding birth for births in the five years preceding the survey.

A. Coverage: All second- and higher-order births that occurred in the 0–59 months preceding the interview.

B. Median: Interpolated retrospective median.

Calculation

A. Covered births: Children both surviving and dead who have a difference between the date of interview and their birth date (both in CMC format) of 0–59 months, excluding first-born children.

B. Preceding birth interval: Difference between birth date of child and birth date of preceding child (in CMC format) grouped into categories. For children of multiple births, the birth date of the preceding child is the number of months since the end of the preceding pregnancy that ended in a live birth.

C. Median is calculated from single month percent distributions of durations of preceding birth intervals. Median is linearly interpolated between the duration month values by which 50 percent or more of the preceding intervals were closed by a birth.

Handling of Missing Values

Not applicable—Birth dates of children are imputed if missing; therefore birth dates do not have missing values.

Notes and Considerations

1. First-order-birth children (and their twins) are excluded from calculation.

2. Births in the month of interview are included and births 60 or more months before the interview are excluded.

3. Since birth dates at both ends of the birth interval are in CMC, difference between them is taken as an exact difference. No adjustment is made to the interpolated median.

AGE AT FIRST BIRTH

Statistics: *Percent Distribution by Years of Age at First Birth*

Definition

Percentages of women who had a first birth in the specified age groups.

A. Coverage: Women of all marital statuses.

B. Numerator: Number of women whose first birth occurred in the time they were in the specified age group.

C. Denominator: Total number of women including those without a birth.

Calculation

A. Age at first birth: Difference between birth date of first-born child and birth date of woman (in CMC format) grouped into categories of age in years.

B. Numerators: Number of women within specified age group categories. Women with no births are included in a separate category.

C. Denominator is the number of women of all marital statuses. Ever-married sample denominators are adjusted by the all women factors.

D. Numerators for each age group category are divided by the same denominator and multiplied by 100 to obtain percentages.

Handling of Missing Values

Not applicable—Birth dates of children and women are imputed if missing; therefore birth dates do not have missing values.

Notes and Considerations

1. Percentages add up to 100 percent.

2. Percentages are not calculated for the first-birth age categories by cohorts of women where the youngest member of the cohort has not yet completed the oldest age of the category. (For example, percentages for the cohort of women 20–24 at the time of the survey will not be calculated for age group categories 20–21, 22–24, and 25+ because some women without births are still in or below those age ranges and could still have first births within the age group categories).

Statistics: *Median Age at First Birth*

Definition

Median age in years at birth of first child.

A. Coverage: Women of all marital statuses.

B. Median: Interpolated calculated median.

Calculation

A. Age at first birth: Difference between birth date of first-born child and birth date of woman (in CMC format), truncated to single years of age.

B. Numerators: Number of women within single age group categories. Women with no births are included in a separate category.

C. Denominator is the number of women of all marital statuses. Ever-married samples denominators are adjusted by the all women factors.

D. Numerators for each age group category are divided by the same denominator and multiplied by 100 to obtain percentages.

E. Medians are calculated from cumulated single year of age percent distributions of age at first birth. Median is linearly interpolated between the age values by which 50 percent or more of the women had a first birth.

Handling of Missing Values

Not applicable—Birth dates of children are imputed if missing; therefore birth dates do not have missing values.

Notes and Considerations

1. Since the difference in mother's and child's birth dates are truncated to years, difference between them is taken as including up to next completed year of age (cumulated percentage for women with a difference of 19 years actually includes all women below 20 years of age at first birth). Therefore an adjustment is made to the interpolated median by increasing the interpolated value by one year.

2. Since the median is based on all women including those without a birth, there may not be a median for younger cohorts of women (since fewer than 50 percent of the cohort may have had a birth).

TEENAGE PREGNANCY AND MOTHERHOOD

Statistics: *Percentages Who Are Mothers, Pregnant with Their First Child, and Have Begun Childbearing*

Definition

A. Coverage: Women of all marital statuses between 15 and 19 years of age at interview.

B. Numerator:

1. Percentage of women who are mothers: Number of women who have had a birth.
2. Percentage of women who are pregnant with first child: Number of women who have not had a birth but are pregnant at the time of interview.
3. Percentage of women who have begun childbearing: Number of women who either have had a birth or who are pregnant at the time of interview.

C. Denominator: Total number of women including those without a birth.

Calculation

A. Denominator is the number of women of all marital statuses. Ever-married samples denominators are adjusted by the all women factors.

B. Percentage of women who have begun childbearing is the sum of the percentage who are mothers and the percentage who are pregnant with their first child.

C. Numerators for percentages are divided by the same denominator and multiplied by 100 to obtain percentages.

Handling of Missing Values

Women who have missing data for current pregnancy or are unsure whether they are pregnant are considered not pregnant at the time of interview.

CONTRACEPTION

KNOWLEDGE OF CONTRACEPTIVE METHODS

EVER USE OF CONTRACEPTIVE METHODS

Statistics: ***Percentages Who Know of Any Method, Any Modern Method, Any Traditional Method and Specific Methods***

Percentages Who Ever Used Any Method, Any Modern Method, Any Traditional Method and Specific Methods

Definition

A. Coverage:

1. Population base: There are several bases used—

   a) All women (all interviewed women between ages 15 and 49 years).

   b) Currently married women (married women and women living in a consensual union).

   c) Unmarried sexually active women: Includes women who are not currently married or in a consensual union (single, divorced, widowed, and separated) and who had sexual intercourse within the last 30 days.

   d) Unmarried sexually inactive women: Includes women who are not currently married or in a consensual union (single, divorced, widowed, and separated) and who had sexual intercourse at least once in their lives but not within the last 30 days.

   e) Unmarried women who never had sexual intercourse: Includes women who are not currently married or in a consensual union (single, divorced, widowed, and separated) and who never had sexual intercourse.

2. Time period: Lifetime of woman interviewed.

B. Knowledge Numerators (within each coverage category):

1. Specific methods: The number of women who say they know (or have heard) of the specific method, whether through a spontaneous response or after describing the method.

2. Any method: The number of women who say they know of at least one specific method, including self-reported other ways.

3. Modern methods: The number of women who say they know at least one of the following methods: female sterilization (tubal ligation, laparectomy, voluntary surgical contraception for women), male sterilization (vasectomy, voluntary surgical contraception for men), the contraceptive pill (oral contraceptives), interuterine contraceptive device (IUD), injectables (Depo-Provera), implants (Norplant), female condom, male condom (prophylactic, rubber), diaphragm, contraceptive foam and contraceptive jelly, lactational amenorrhea method (LAM), emergency contraception (double dose of contraceptive pill twice in 24 hours for two days and specific dosage “emergency pills,” does NOT include abortion, menstrual regulation), country-specific modern methods and other modern contraceptive methods respondent mentioned (including cervical cap, contraceptive sponge, and others).

4. Traditional methods: The number of women who say they know at least one of the following methods: periodic abstinence (rhythm, calendar method), withdrawal (coitus interruptus) and country-specific traditional methods of proven effectiveness.

5. Folk methods: The number of women who say they know at least one locally described method and/or spiritual method of unproven effectiveness, such as herbs, amulets, gris-gris, etc.

6. In the latest version of the DHS tabulations, both traditional methods and folk methods have been grouped under the heading traditional methods.

C. Ever Use Numerators (within each coverage category):

1. Specific methods: The number of women who say they have used the specific method. (Women who do not know a method are not asked about its use and are assumed not to have used it).

2. Any method: The number of women who say they have used at least one of any specific method.

3. Modern methods: The number of women who say they have used at least one of the following methods: female sterilization (tubal ligation, laparectomy, voluntary surgical contraception for women), male sterilization (vasectomy, voluntary surgical contraception for men), the contraceptive pill (oral contraceptives), interuterine contraceptive device (IUD), injectables (Depo-Provera), implants (Norplant), female condom, male condom (prophylactic, rubber), diaphragm, contraceptive foam and contraceptive jelly, lactational amenorrhea method (LAM), emergency contraception (double dose of contraceptive pill twice in 24 hours for two days and specific dosage “emergency pills”, does NOT include abortion, menstrual regulation), country-specific modern methods and respondent mentioned other modern contraceptive methods (including cervical cap, contraceptive sponge, and others).

4. Traditional methods: The number of women who say they have used at least one of the following methods: periodic abstinence (rhythm, calendar method), withdrawal (coitus interruptus) and country-specific traditional methods of proven effectiveness.

5. Folk methods: The number of women who say they have used at least one locally described method and spiritual methods of unproven effectiveness, such as herbs, amulets, gris-gris, etc.

6. In the latest version of the DHS tabulations, both traditional methods and folk methods have been grouped under the heading traditional methods.

D. Denominator: The denominators are the numbers of women in the coverage categories.

Calculation

Within each coverage category, the numerator divided by the denominator, expressed as a percentage.

Handling of Missing Values

A. Coverage categorization: Missing value in whether or not currently married is allowed in the data. Missing value in sexual activity treated as had sexual relations (non-virgin). Missing value in time since last intercourse treated as greater than 30 days, i.e., not sexually active.

B. Numerators: Treated as does not know or has not used method for individual methods and grouped methods.

C. Denominators: All women in coverage category included, even if missing values on all methods.

Notes and Considerations

Breastfeeding, prolonged breastfeeding, prolonged abstinence are NOT contraceptive methods in themselves. The lactational amenorrhea method is based on four criteria: Woman is amenorrheic since last birth; last birth occurred within six months; woman is exclusively or predominately breastfeeding; other form of contraception is necessary if any of the foregoing criteria does not hold. In the DHS description of LAM, only the following is used: "Up to 6 months after childbirth, a woman can use a method that requires that she breastfeeds frequently, day and night, and that her menstrual period has not returned." This description varies from the official LAM criteria by not including exclusive or predominant breastfeeding (which is based on whether or not the child received complementary liquids and foods), substituting frequency of breastfeeding (it is not a requirement that frequent night-time feeding occurs), and by not including the criterion that the woman knows that another form of contraception is necessary. The DHS description may therefore include women who say yes even though they had never heard of the term LAM or of programs that instruct in the method, thus overestimating knowledge and ever use.

Changes over Time

The list of specific methods and their categorization has changed.

In DHS I and II surveys, modern methods included pill, IUD, injection, vaginal methods, condom, female sterilization, and male sterilization. The vaginal methods included in a single group diaphragm, foam and jelly. Traditional methods included periodic abstinence (of any kind), withdrawal and all respondent-mentioned other methods.

In DHS III surveys, modern methods included pill, IUD, injection, vaginal methods, condom, female sterilization, male sterilization, and implants. Traditional methods included periodic abstinence (of any kind), withdrawal, and lactational amenorrhea. Folk methods included respondent-mentioned other methods and were categorized separately from traditional methods.

CURRENT USE OF CONTRACEPTIVE METHODS

Statistics: *Percentages Who Currently Use Any Method, Any Modern Method, Any Traditional Method and Specific Methods*

The percentage of currently married women who currently use any method of contraception is called the contraceptive prevalence rate (CPR).

Definition

A. Coverage: Population base—There are several bases used:

1. All women (all interviewed women between ages 15 and 49 years).

   a) Currently married women (married women and women living in a consensual union).

   b) Unmarried sexually active women—includes women who are not currently married or in a consensual union (single, divorced, widowed and separated) and who had sexual intercourse within the last 30 days.

   c) The following two groups are not included in the current use of contraception table:

      i. Unmarried sexually inactive women—includes women who are not currently married or in a consensual union (single, divorced, widowed and separated) and who had sexual intercourse at least once in their lives but not within the last 30 days.

      ii. Unmarried women who never had sexual intercourse—includes women who are not currently married or in a consensual union (single, divorced, widowed and separated) and who never had sexual intercourse.

B. Time period: Current use is defined by the respondent.

1. Numerators (within each coverage category):

   a) Specific methods: The number of women who say they currently use the specific method, after being asked whether they are doing something or using something to delay or avoid becoming pregnant. Women who declared that they ever used female sterilization are directly coded as currently using female sterilization. Women who say they are pregnant are coded as not currently using any method.

   b) Any method: The number of women who say they use any specific method, including ever-users of female sterilization.

   c) Modern methods: The number of women who say they use one of the following methods: female sterilization (tubal ligation, laparectomy, voluntary surgical contraception for women), male sterilization (vasectomy, voluntary surgical contraception for men), the contraceptive pill (oral contraceptives), interuterine contraceptive device (IUD), injectables (Depo-Provera), implants (Norplant), female condom, male condom (prophylactic, rubber), diaphragm, contraceptive foam and contraceptive jelly, lactational amenorrhea method (LAM), country-specific modern methods and respondent mentioned other modern contraceptive methods (including cervical cap, contraceptive sponge, and others), does NOT include abortions and menstrual regulation.

d) Traditional methods: The number of women who say they currently use one of the following methods: periodic abstinence (rhythm, calendar method), withdrawal (coitus interruptus) and country-specific traditional methods of proven effectiveness.

e) Folk methods: The number of women who say they currently use at least one locally described methods and spiritual methods of unproven effectiveness, such as herbs, amulets, gris-gris, etc.

In the latest version of the DHS tabulations, both traditional methods and folk methods have been grouped under the heading traditional methods.

C. Denominator: The denominators are the numbers of women in the coverage categories.

Calculation

Within each coverage category, the numerator divided by the denominator, expressed as a percentage.

Handling of Missing Values

A. Coverage categorization: Missing value in whether or not currently married is allowed in the data. Missing value in sexual activity treated as had sexual relations (non-virgin). Missing value in time since last intercourse treated as greater than 30 days (not sexually active).

B. Numerators: Treated as does not use any method for individual methods and grouped methods.

C. Denominators: All women in coverage category included, even if missing values on current use of methods.

Notes and Considerations

Breastfeeding, prolonged breastfeeding, and prolonged abstinence are NOT contraceptive methods in themselves. The lactational amenorrhea method is based on four criteria: Woman is amenorrheic since last birth; last birth occurred within six months; woman is exclusively or predominately breastfeeding; other form of contraception is necessary if any of the foregoing criteria does not hold. In the DHS description of LAM only the following is used: “Up to 6 months after childbirth, a woman can use a method that requires that she breastfeeds frequently, day and night, and that her menstrual period has not returned.” This description varies from the official LAM criteria by not including exclusive or predominant breastfeeding (which is based on whether or not the child received complementary liquids and foods), substituting frequency of breastfeeding (it is not a requirement that frequent night-time feeding occurs) and by not including the criterion that the woman knows that another form of contraception is necessary. The DHS description may therefore include women who say yes even though they had never heard of the term LAM or of programs that instruct in the method, thus overestimating knowledge and ever use.

Each respondent defines current use of contraception. While some methods such as sterilization, pill, IUD, injectables, implants, and LAM involve continuous protection, other methods are coital-specific, requiring use during intercourse, such as condoms, vaginal methods, periodic abstinence, and withdrawal. Current use for coital-specific methods is a difficult concept since it may mean use at last intercourse, which could have been a long time before the interview or intention to use at next intercourse.

Changes over Time

The list of specific methods and their categorization has changed.

In DHS I and II surveys, modern methods included pill, IUD, injection, vaginal methods, condom, female sterilization, and male sterilization. The vaginal methods included in a single group diaphragm, foam and jelly. Traditional methods included periodic abstinence (of any kind), withdrawal, and all respondent-mentioned other methods.

In DHS III surveys, modern methods included pill, IUD, injection, vaginal methods, condom, female sterilization, male sterilization, and implants. Traditional methods included periodic abstinence (of any kind), withdrawal and lactational amenorrhea. Folk methods included respondent-mentioned other methods and were categorized separately from traditional methods.

In Measure *DHS+* (DHS IV) surveys, emergency contraception was added to the list of contraceptive methods but is not included as a separate method for current use (included in "others"). The questionnaire allows for more than one method to be currently used. For specific methods, the following hierarchy is used to tabulate current use: female sterilization, male sterilization, contraceptive pill, intrauterine contraceptive device (IUD), contraceptive injection, contraceptive implants (Norplant), condoms, vaginal methods (foam, jelly, suppository), lactational amenorrhea method (LAM), periodic abstinence, withdrawal, other methods.

NUMBER OF CHILDREN AT FIRST USE OF CONTRACEPTION

Statistics: *Percent distribution of Women Who Have Ever Used Contraception by Number of Living Children at First Use of Contraception*

Median number of living children at first use of contraception for women who have ever used contraception.

Definition

A. Coverage:

1. Population base.

2. All women (all interviewed women between ages 15 and 49 years) who have ever used contraception.

3. Time period: Lifetime of the respondent.

B. Numerators: Women who have ever used contraception, by declared number of living children at the time of first use of contraception.

C. Denominator: The denominator is the number of women who have ever used contraception.

Calculation

A. Percent distribution: The numerators divided by the denominator, expressed as a percentage. Total is 100 percent.

B. Median: The interpolated number of living children by which the accumulated percentages of the distribution reaches 50 percent.

Example: Suppose the distribution of living children at first use is the following:

0 children: 10 percent, 1 child: 25 percent, 2 children: 35 percent, 3 children: 20 percent, and 4 children: 10 percent.

Then the accumulated distribution is 0 children: 10 percent, 1 child: 35 percent, 2 children: 70 percent, 3 children: 90 percent, 4 children: 100 percent. The median then would be interpolated between 1 child and 2 children as (50% – 35%)/(70% – 35%) + 2 = 2.4 living children at first use. However, see discussion below.

Handling of Missing Values

Numerators: Missing value for number of living children at first use is included as a separate category in the percent distribution.

Denominator: Because respondents are asked about ever use of contraceptive methods separately, to have a missing value in ever use of contraception requires that all answers be missing. Respondents with all answers missing are classified as never-users in the standard recode file.

Median: Women with a missing value for number of living children at first use are excluded from both the numerator and the denominator of the percentages used to calculate the median number of living children at first use.

Notes and Considerations

Use of an interpolated median value for integer response categories is illogical since no respondent can have a non-integer value (i.e., no woman can have 2.4 children). For these types of calculations a mean value would likely be preferred to a median.

Changes over Time

In the latter tabulations for Measure *DHS+* (DHS IV), the median has been dropped.

Note that when tabulating by current age of respondent or variables correlated with current age of respondent, the results may be misleading since ever use of contraception varies by age as well as by number of living children, changing the denominator in addition to the numerator.

USE OF SOCIAL MARKETING BRAND OF PILL

Statistics: *Percentage of Women Currently Using Oral Contraceptives Who Use a Socially Marketed Brand*

Definition

A. Coverage:

1. Population base.

2. All women (all interviewed women between ages 15 and 49 years).

3. Time period: Current as defined by the respondent.

B. Numerator: Number of women who are currently using a socially marketed brand of oral contraceptive pill (country-specific).

C. Denominator: Number of women who are currently using oral contraceptives.

Calculation

Percentage—The numerator divided by the denominator, expressed as a percentage.

Handling of Missing Values

Women who did not know or have a missing value for the brand of oral contraceptive are excluded from the numerator (assumed not using brand).

Notes and Considerations

KNOWLEDGE OF THE FERTILE PERIOD

Statistics: *Percent Distribution of Women, by Knowledge of the Fertile Period During the Ovulatory Cycle*

Definition

A. Coverage:

1. Population base.

   a) All women (all interviewed women between ages 15 and 49 years).

   b) Women who currently use periodic abstinence.

   c) Women who do not currently use periodic abstinence.

2. Time period: Current as defined by the respondent.

B. Numerators: Number of women within each base by category of knowledge of the fertile period.

C. Denominator: Number of women who are in each population base.

Calculation

Percent distribution—The numerator divided by the denominator, expressed as a percentage. Categories add to 100 percent total.

Handling of Missing Values

Women who did not know or have a missing value for knowledge of the fertile period are included as separate categories in the distribution. Women who have a missing value for current method are included as nonusers of periodic abstinence.

Notes and Considerations

AGE AT STERILIZATION

Statistics: *Percent Distribution, by Grouped Years of Age at Contraceptive Sterilization*

Definition

Percentages of women who had contraceptive sterilization in the specified age groups.

A. Coverage: Women of all marital statuses.

B. Numerator: Number of women whose contraceptive sterilization occurred in the time they were in the specified age group.

C. Denominator: Total number of women who have had contraceptive sterilization.

Calculation

A. Age at sterilization: Difference between date of sterilization operation and birth date of woman (in CMC format), truncated to single years of age.

B. Numerators: Number of women within specified age group categories. Women who have not been sterilized are excluded.

C. Denominator is the number of women of all marital statuses who have been sterilized.

D. Numerators for each age group category are divided by the same denominator and multiplied by 100 to obtain percents.

Handling of Missing Values

Women with missing values for age at sterilization are imputed in the data file.

Notes and Considerations

Percentages add up to 100 percent.

Statistics: *Median Age at Contraceptive Sterilization*

Definition

Median age in years when women had contraceptive sterilization operation.

A. Coverage: Women of all marital statuses who were sterilized at less than 40 years of age.

B. Median type: Interpolated calculated median.

Calculation

A. Age at sterilization: Difference between date of sterilization operation and birth date of woman (in CMC format), truncated to single years of age.

B. Numerators: Number of women within single age group categories. Women not sterilized are not included.

C. Denominator is the number of women of all marital statuses who were sterilized at less than 40 years of age.

D. Numerators for each age group category are divided by the same denominator and multiplied by 100 to obtain percentages.

E. Medians are calculated from cumulated single year of age percent distributions of age at sterilization. Median is linearly interpolated between the age values by which 50 percent or more of the women had been sterilized.

Handling of Missing Values

Missing values for age at sterilization have been imputed in the data set.

Notes and Considerations

Women who were sterilized at age 40 years or over are excluded from the calculation of the median to minimize problems of censoring.

Since the difference between the date of sterilization and the respondent's birth date is truncated to years, the difference is taken as including up to next completed year of age (the cumulated percentage for women with a difference of 29 years actually includes all women below 30 years of age at sterilization). Therefore an adjustment is made to the interpolated median by increasing the interpolated value by one year.

Since the median is based only on women who have been sterilized by the time of the survey, the median age is not calculated for periods of time of sterilization 10 years and more before the survey because of censoring (the oldest possible age in the data set would be 39 years at ten years prior and less for earlier time periods).

SOURCE OF CONTRACEPTION

Statistics: *Percent Distribution of Current Users of Modern Methods, by Most Recent Source of Supply*

Definition

A. Coverage: Women of all marital statuses who use a modern method of contraception except for the lactational amenorrhea method (LAM).

B. Numerators: Number of women, by declared most recent source of contraception.

C. Denominator: Total number of women who currently use a modern method of contraception, excluding LAM.

Calculation

Numerators for category are divided by the same denominator and multiplied by 100 to obtain percentages.

Handling of Missing Values

Women who did not know or with missing values for most recent source are included as separate categories.

Notes and Considerations

Percentages add up to 100 percent.

Modern methods include female sterilization, male sterilization, pill, IUD, injectables, implants, male condom, female condom, diaphragm, and foam or jelly.

INFORMED CHOICE

Statistics:

A. Percentage of current users of selected contraceptive methods who were informed about side effects or problems of the method used.

B. Percentage of current users of selected contraceptive methods who were informed of what to do if they experienced side effects or problems with the method used.

C. Percentage of current users of selected contraceptive methods who were informed of other methods of contraception that could be used.

D. Percentage of women who have had contraceptive sterilization who were told that they would not be able to have any more children.

Definition

Percentages of women who had contraceptive sterilization in the specified grouped ages.

A. Coverage:

1. Statistics 1 and 2: Women of all marital statuses who use female sterilization, pill, IUD, injectables, and implants.

2. Statistic 3: Women of all marital statuses who use female sterilization, pill, IUD, injectables, implants, female condom, diaphragm, foam or jelly, and LAM.

3. Statistic 4: Women of all marital statuses who have had contraceptive sterilization.

B. Numerators: Number of women who say they were informed at the start of the current episode of use of the method. Limited to method episodes that began in the five years preceding the survey.

C. Denominators: Total number of women who are in the coverage category for each statistic. Limited to method episodes that began in the five years preceding the survey.

Calculation

Numerators for each statistic are divided by the denominators and multiplied by 100 to obtain percentages.

Handling of Missing Values

Women who did not know whether they were informed are considered not informed (“NO”). Women with missing values on whether they were informed are excluded from both denominators and numerators.

Notes and Considerations

Percentages add up to 100 percent.

Modern methods include female sterilization, male sterilization, pill, IUD, injectables, implants, male condom, female condom, diaphragm, and foam or jelly.

FIRST-YEAR CONTRACEPTIVE DISCONTINUATION RATES

Statistics: *Percentage of Episodes of Contraceptive Use Where the Specific Method is Discontinued Within 12 Months After Beginning Its Use, by Reason for Discontinuation, According to Specific Method*

The discontinuation rates are based on episodes of use of particular methods. An individual woman may contribute more than one episode to the calculation. The calculation procedure is based on life table methods.

Definition

A. Coverage:

1. Population base.

2. Contraceptive use episodes of all interviewed women who used a contraceptive method in the five years preceding the survey.

3. Time period: Users who began an episode of use 3–59 months before the interview.

B. Discontinuation rate: Users who discontinue using a contraceptive method within 12 months of beginning use during a specific episode of use. Users who switch to another method are considered to have discontinued the previous method at the time of switching.

C. Exposure: Duration of use of a specific method within one episode of use. Exposure begins with initial month of use and ends with discontinuation or with the month of interview if method was still being used at the time of the interview.

Calculation

A. In the DHS standard recode file, the data for contraceptive discontinuation rates come from the reproductive calendar. The reproductive calendar in the questionnaire consists of two or more columns of boxes, where each box represents a specific calendar month. The reproductive calendar usually begins with the first month of the fifth calendar year before the date of the start of fieldwork. For example, if the fieldwork began in July 2000, the calendar would start in January 1995. In the first column, episodes of use of contraception are indicated by placing a method code in the boxes that correspond to the calendar months when used. Pregnancies, births, and nonlive birth terminations are also represented in this column by placing the corresponding codes in the appropriate months. Months with codes “0” are those in which the woman did not use contraception, was not pregnant, did not give birth, or did not have a fetal loss or stillbirth. In the second column, the reason for contraceptive discontinuation is noted in the box that corresponds to the last month of use.

B. In the standard recode file, the reproductive calendar is represented by character strings of fixed length. Each position within the character string represents a calendar month. Thus the third position may represent March 1995, while the fifth position represents May 1995. To calculate the durations of the episodes of use, each position is examined in chronological order for a contraceptive code. The first code following a position without that code indicates the start of a new episode of use. The first position is ignored in this examination, since a code in that position may represent an episode of use that began before the calendar start date. The number of contiguous positions with the same contraceptive code indicates the number of months of use in the episode. An episode ends if the following position does not have the same contraceptive code (a discontinuation) or corresponds to the month of interview (a censored duration). The episodes are then tabulated by duration and type of ending for each contraceptive method and for all methods combined. Standard life table calculations are then applied to the terminations to calculate months of exposure and number of discontinuations by month of episode. The cumulative proportion that discontinued by 12 months is taken as the 12-month discontinuation rate.

C. The discontinuation rate is categorized by type of discontinuation, which is noted in the second column of the reproductive calendar in the box that corresponds to the month of discontinuation. Discontinuation, by reason of contraceptive failure, is given if the woman became pregnant while using contraception. In this case, the box in the first column corresponding to the month following the termination should include a "P" for pregnancy or a "T" for pregnancy termination (very unlikely to occur). Discontinuation to switch to another method is determined if the box for the month following the discontinuation of the specific method contains another contraceptive method. Discontinuation for switching is also indicated by a code in the second calendar column that indicates that the woman wanted a more effective method and that the new method began within two months of discontinuation (i.e., only one month with a '0,' indicating no contraceptive use, between episodes of use). Discontinuation of a method, by reason of desire to become pregnant, is indicated in column 2 by the appropriate code.

D. In formulas, the monthly rate of discontinuation, q_{ij}, where i is the number of months since the start of the episode and j is the reason for discontinuation, is calculated by dividing the number of episodes discontinued in month i, d_{ij}, by the total number of episodes that reached duration i,e.;

$$q_{ij} = \frac{d_{ij}}{e_i}$$

and the cumulative probability of discontinuing by 12 months duration for reason j is

$$Q_{12j} = 1 - \prod_{i=1}^{12} (1 - q_{ij})$$

Handling of Missing Values

Use of contraception is not allowed to be missing in any month in the calendar. Missing and unknown reasons for discontinuation are treated as "Other" reasons.

Notes and Considerations

The life table calculated is a true multiple decrement table producing net discontinuation rates. The various reasons for discontinuation are treated as competing risks, and the monthly probabilities of discontinuation are additive across the reasons for discontinuation.

In the DHS tables, only episodes that began within the calendar period and ended three months before the interview are included. Episodes that began before the beginning of the calendar are excluded. Episodes that ended in the month of interview or the two months prior are treated as censored at three months before the interview rather than terminated afterwards to avoid bias due to unrecognized pregnancies.

For methods that are not followed by another method or a pregnancy, it is assumed that the method episode started on average in the middle of the first month of use and ended in the middle of the month after the last noted month of use. If the month following the last noted method indicates a pregnancy or a different method, then it is assumed that the episode ended on average in the middle of that following month. Thus, the duration of exposure is taken as the difference between the month of first use and the month of last use (i.e., equal to the number of months during that episode with a notation for the method).

Methods that have less than 125 months of exposure (unweighted) are not shown because of large sampling variance.

Changes over Time

The list of specific methods and their categorization has changed.

In DHS I and II surveys, modern methods included pill, IUD, injection, vaginal methods, condom, female sterilization, and male sterilization. The vaginal methods included in a single group diaphragm, foam, and jelly. Traditional methods included periodic abstinence (of any kind), withdrawal, and all respondent-mentioned other methods.

In DHS III surveys, modern methods included pill, IUD, injection, vaginal methods, condom, female sterilization, male sterilization, and implants. Traditional methods included periodic abstinence (of any kind), withdrawal, and lactational amenorrhea. Folk methods included respondent-mentioned other methods and were categorized separately from traditional methods.

In Measure *DHS+* (DHS IV) surveys, emergency contraception was added to the list of contraceptive methods but is not included as a separate method for current use (i.e., included in “other”). The questionnaire allows for more than one method to be currently used. For specific methods, the following hierarchy is used to tabulate current use: female sterilization, male sterilization, contraceptive pill, intrauterine contraceptive device (IUD), contraceptive injection, contraceptive implants (Norplant), condoms, vaginal methods (foam, jelly, suppository), lactational amenorrhea method (LAM), periodic abstinence, withdrawal, and other methods.

REASONS FOR DISCONTINUING CONTRACEPTION

Statistics: *Percent Distribution of Discontinued Episodes of Contraceptive Use, by Reason for Discontinuation, According to Specific Method*

The distributions, by reason of discontinuation, are based on episodes of use of particular methods. An individual woman may contribute more than one episode to the calculation.

Definition

A. Coverage:

1. Population base.

2. Episodes of contraceptive use that were discontinued for all interviewed women who used a contraceptive method in the five years preceding the survey.

3. Time period: Users who discontinued an episode of use 0–59 months before the interview.

B. Numerator: Number of episodes of use that were discontinued, by reason of discontinuation, according to contraceptive method and for all methods together.

C. Denominator: Total number of episodes of use that were discontinued, according to contraceptive method and for all methods.

Calculation

A. In the DHS standard recode file, the data for contraceptive discontinuation rates come from the reproductive calendar. The reproductive calendar in the questionnaire consists of two or more columns of boxes where each box represents a specific calendar month. The reproductive calendar usually begins with the first month of the fifth calendar year before the date of the start of fieldwork. For example, if the fieldwork began in July 2000, the calendar would start in January 1995. In the first column, episodes of use of contraception are indicated by placing a method code in the boxes that correspond to the calendar months when used. Pregnancies, births, and non-live birth terminations are also represented in this column by placing the corresponding codes in the appropriate months. Months with no codes are those in which the woman did not use contraception, was not pregnant, did not give birth or did not have a fetal loss or stillbirth. In the second column, the reason for contraceptive discontinuation is noted in the box that corresponds to the month of discontinuation.

B. In the standard recode file, the reproductive calendar is represented by character strings of fixed length. Each position within the character string represents a calendar month. Thus the third position may represent March 1995, while the fifth position represents May 1995. Each position is examined in chronological order for a contraceptive code. The first occurrence of a contraceptive code following a position without that code indicates the start of a new episode of use. Subsequent positions are examined until a different code is found. An episode of use ends if the following position does not have the same contraceptive code (a discontinuation). The episodes are then tabulated by type of ending for each type of contraceptive method and for all methods combined.

C. The reason for discontinuation is noted in the second column of the reproductive calendar in the box that corresponds to the month of discontinuation.

Handling of Missing Values

Information on use of contraception is not allowed to be missing in any month in the calendar. Missing reasons for discontinuation are treated as "Other" reasons.

Notes and Considerations

The distribution of reasons for discontinuation is a little different than that obtained in the calculation of discontinuation rates. In the distribution, all discontinuations in the five years preceding the survey are included, whereas in the calculation of the discontinuation rates, only those that ended within the first 12 months of use are included. Also, censored episodes are not included in the rate calculations.

Methods that have less than 25 discontinued episodes (unweighted) are not shown because of large sampling variance.

Changes over Time

The list of specific methods and their categorization has changed.

In DHS I and II surveys, modern methods included pill, IUD, injection, vaginal methods, condom, female sterilization, and male sterilization. The vaginal methods included in a single group diaphragm, foam, and jelly. Traditional methods included periodic abstinence (of any kind), withdrawal, and all respondent-mentioned other methods.

In DHS III surveys, modern methods included pill, IUD, injection, vaginal methods, condom, female sterilization, male sterilization, and implants. Traditional methods included periodic abstinence (of any kind), withdrawal, and lactational amenorrhea. Folk methods included respondent-mentioned other methods and were categorized separately from traditional methods.

In Measure *DHS+* (DHS IV) surveys, emergency contraception was added to the list of contraceptive methods but is not included as a separate method for current use (i.e., included in "others"). The questionnaire allows for more than one method to be currently used. For specific methods, the following hierarchy is used to tabulate current use: female sterilization, male sterilization, contraceptive pill, intrauterine contraceptive device (IUD), contraceptive injection, contraceptive implants (Norplant), condoms, vaginal methods (foam, jelly, suppository), lactational amenorrhea method (LAM), periodic abstinence, withdrawal, other methods.

FUTURE USE OF CONTRACEPTION

Statistics: *Percent Distribution of Currently Married Women Who Are Not Using Contraception, by Intention to Use in the Future*

Definition

A. Numerator: Number of currently married women who do not currently use a method of contraception, by whether they intend to use contraception at any time in the future.

B. Denominator: Total number of currently married women who do not currently use a method of contraception.

Calculation

Numerators for each response are divided by the same denominator and multiplied by 100 to obtain percentages.

Handling of Missing Values

Women who are unsure or with missing values for intention to use are included as separate categories.

Notes and Considerations

Percentages add up to 100 percent.

REASON FOR NOT INTENDING TO USE CONTRACEPTION

Statistics: *Percent Distribution of Currently Married Women Who Are Not Using Contraception and Who Do Not Intend to Use At Any Time in the Future, by Main Reason for Not Intending to Use*

Definition

A. Numerators: Number of currently married women who do not currently use a method of contraception and who do not intend to use at any time in the future, by main reason for not intending to use at any time in the future.

B. Denominator: Total number of currently married women who do not currently use a method of contraception and who do not intend to use at any time in the future.

Calculation

Numerators for each response are divided by the same denominator and multiplied by 100 to obtain percentages.

Handling of Missing Values

Women who are unsure or with missing values for intention to use are not included in either the numerator or the denominator.

Women with missing values for reasons for not intending to use are not included in either the numerator or the denominator.

Women who do not know the reasons for not intending to use are included as a separate category.

Notes and Considerations

Percentages add up to 100 percent.

PREFERRED METHOD OF CONTRACEPTION FOR FUTURE USE

Statistics: *Percent Distribution of Currently Married Women Who Are Not Using Contraception and Who Intend to Use at Some Time in the Future, by Preferred Method in Future*

Definition

A. Numerators: Number of currently married women who do not currently use a method of contraception and who intend to use at some time in the future by method would prefer to use in the future.

B. Denominator: Total number of currently married women who do not currently use a method of contraception and who intend to use at some time in the future.

Calculation

Numerators for each response are divided by the same denominator and multiplied by 100 to obtain percents.

Handling of Missing Values

Women who are unsure or with missing values for intention to use are not included in either the numerator or the denominator.

Women with missing values for preferred method are included as a separate category.

Women who are unsure of preferred method are included as a separate category.

Notes and Considerations

Percentages add up to 100 percent.

In many countries, women state a method that they would prefer to use that previously they declared not knowing.

EXPOSURE TO FAMILY PLANNING MESSAGES

Statistics: *Percentages of All Women, by Background Characteristic, Who Heard Or Saw a Family Planning Message on the Radio or Television or in a Newspaper or Magazine in the Last Few Months or in None of the Preceding Media*

Definition

A. Numerators for each indicator for each background category:

1. Number of women who heard a family planning message on the radio in the last few months.
2. Number of women who saw a family planning messages on television in the last few months.
3. Number of women who read a family planning message in the newspaper or in a magazine in the last few months.
4. Number of women who did not hear a family planning message on the radio, did not see a family planning message on television, and did not read a family planning message in a newspaper or a magazine in the last few months.

B. Denominator for all indicators:

1. For each background category.
2. Total number of women of all marital statuses.

Calculation

For each background category, the numerator for each indicator is divided by the denominator of that background category and multiplied by 100 to obtain percentages.

Handling of Missing Values

Women who are unsure or with missing values for heard, saw, or read a family planning message are considered as did not hear, see, or read a message. These women are included in the denominator of the percentage.

CONTACT OF NONUSERS WITH FAMILY PLANNING PROVIDERS

Statistics: *Percentages of Women Who Are Not Using Contraception Who Were Visited by a Fieldworker Who Discussed Family Planning, Who Visited a Health Facility and Discussed Family Planning, and Who Visited a Health Facility but Did Not Discuss Family Planning During the 12 Months Preceding the Survey*

Definition

A. Coverage: Women who are not using a contraceptive method.

B. Numerators for each indicator for each background category:

1. Number of women who were visited by a fieldworker in the preceding 12 months who discussed family planning.

2. Number of women who visited a health facility in the preceding 12 months and who discussed family planning at the facility.

3. Number of women who visited a health facility in the preceding 12 months but who did not discuss family planning at the facility.

4. Number of women who did not discuss family planning with either a field worker or at a health facility in the preceding 12 months.

C. Denominator for all indicators:

1. For each background category.

2. Total number of women of all marital statuses.

Calculation

For each background category, the numerator for each indicator is divided by the denominator of that background category and multiplied by 100 to obtain percentages.

Handling of Missing Values

Women who are unsure or with missing values are considered as not having been visited by a field worker or visited a health facility or had discussed family planning. These women are included in the denominator of the percentages.

DISCUSSION OF FAMILY PLANNING WITH HUSBAND

Statistics: *Percent Distribution of Currently Married Women Who Know a Contraceptive Method by Number of Times They Discussed Family Planning with Their Husbands in the Preceding 12 Months*

Definition

A. Coverage: Currently married women who know at least one contraceptive method.

B. Numerators: Number of women distributed by number of times they discussed family planning with their husbands in the preceding 12 months.

C. Denominator: Number of currently married women who know at least one contraceptive method.

Calculation

The numerator for each category is divided by the denominator and multiplied by 100 to obtain percentages.

Handling of Missing Values

Women with missing values on number of times discussed family planning with husband are included as a separate category. These women are included in the denominator.

Notes and Considerations

Percentages add up to 100 percent.

Question on discussion of family planning with husband is only asked in the model B core questionnaire (for countries with low contraceptive prevalence) in *DHS+*.

Statistics: *Percent Distribution of Currently Married Women Who Know a Contraceptive Method, by Approval of Family Planning and Their Perception of Husband's Attitude toward Family Planning*

Definition

A. Coverage: Currently married women who know at least one contraceptive method.

B. Numerators: Number of women distributed, by approval of family planning and within approval categories and by husband's perceived approval of family planning.

C. Denominator: Number of currently married women who know at least one contraceptive method.

Calculation

The numerator for each category is divided by the denominator and multiplied by 100 to obtain percentages.

Handling of Missing Values

Women unsure of whether they approve or with missing values on approval included as a separate category. Men's approval is not broken down within this category. Women with unknown and missing values on perceived husband's approval are included as separate categories within women's approval categories. Women who are unsure, and with missing or unknown values for either themselves or for their husbands are included in the denominator.

Notes and Considerations

Percentages add up to 100 percent.

In *DHS+*, the question on perception of husband's approval of family planning is only asked in low-contraceptive prevalence countries (using the model B core questionnaire) without a male survey. Question on respondent's approval of family planning is only asked in countries using the model B core questionnaire.

NUPTIALITY

CURRENT MARITAL STATUS

Statistics: *Percent Distribution of Women, by Current Marital Status*

Definition

Percentage of women according to current status of marriage or cohabitation.

A. Coverage: All women.

B. Numerator categories: Numbers of women who are 1) currently legally or formally married, 2) not married but living with a man in a consensual union, divorced from a legal or formal marriage, 3) separated from a marriage or consensual union, 4) widowed from a marriage or consensual union and not remarried or not in a consensual union, and 5) who never married nor lived in a consensual union.

C. Denominator: Number of women in sample.

Calculation

For each category, 100 times the quotient of the numerator divided by the denominator. In ever-married samples, the denominator is adjusted by the all woman factor.

Handling of Missing Values

Not applicable—Marital status is not allowed to have missing values.

Notes and Considerations

Percentages add up to 100 percent.

The definition of a consensual union may vary between countries. No proof of marital status is shown; categorization is done by the respondent.

NUMBER OF CO-WIVES

Statistics: *Percent Distribution of Currently Married Women by Number of Co-Wives*

Definition

Percentage of women who are either legally or formally married or who are living in a consensual union, by the number of other wives that her partner has.

A. Coverage: Currently married women.

B. Numerator categories: Number of women, by number of other wives of her current partner.

C. Denominator: Number of currently married women in sample.

Calculation

For each number of co-wives, 100 times the quotient of the numerator divided by the denominator.

Handling of Missing Values

If number of co-wives is unknown by respondent or missing, respondent is excluded from both denominator and numerator.

Notes and Considerations

Percentages add up to 100 percent.

Questions on number of co-wives are normally only asked in questionnaires using the B core (low prevalence), primarily in sub-Saharan Africa.

AGE AT FIRST MARRIAGE

Statistics: *Percent Distribution by Years of Age at First Marriage, Total, and According to Current Age Group*

Definition

Percents of women who first married or lived with a man before attaining the specified ages.

A. Coverage: Women of all marital statuses.

B. Numerator: Number of women whose first marriage or consensual union occurred before they attained the specified age group.

C. Denominator: Total number of women, including those who have never been married or in a consensual union.

Calculation

A. Age at first marriage or first union: Difference between date when women began living with first husband or consensual partner and date of birth of woman in completed single years (in the DHS standard recode, this difference is already calculated—V511).

B. Numerators: Number of women below specified age categories. Women who have never married nor lived in a consensual union are included in a separate category.

C. Denominator is the number of women of all marital statuses. Ever-married sample denominators are adjusted by the all women factors.

D. Numerators for each age group category are divided by the same denominator and multiplied by 100 to obtain percentages.

Handling of Missing Values

Not applicable—birth dates of women and ages at first marriage were imputed if missing.

Notes and Considerations

Percentages add up to 100 percent.

Percents are not calculated for the age-at-first-marriage categories by cohorts of women where the youngest member of the cohort has not yet completed the oldest age of the category. (For example, percents for the cohort of women 15–19 at the time of the survey will not be calculated for age categories 18, 20, 22, and 25 because some women who have never been married are still below those ages and could still marry or start to live in a consensual union before reaching those age categories).

MEDIAN AGE AT FIRST MARRIAGE

Definition

Median age in years when the women first lived with husband or consensual partner.

A. Coverage: Women of all marital statuses.

B. Median: Interpolated calculated median.

Calculation

A. Age at first marriage or first union: Difference between date when woman began living with first husband or consensual partner and date of birth of woman in completed single years (in the DHS standard recode, this difference is already calculated—V511).

B. Numerators: Number of women within single year of age categories. Women who have never been married or lived in a consensual union are included in a separate category.

C. Denominator is the number of women of all marital statuses. Ever-married samples denominators are adjusted by the all women factors.

D. Numerators for each age category are divided by the same denominator and multiplied by 100 to obtain percents.

E. Medians are calculated from cumulated single-year percent distributions of age at first marriage or consensual union. Median is linearly interpolated between the age values by which 50 percent or more of the women had married for the first time or begun living in a consensual union.

Handling of Missing Values

Not applicable—Dates of birth and age at first marriage or consensual union were imputed if missing.

Notes and Considerations

Since the difference in woman's birth dates and date she began living with first husband or partner are truncated to years, the difference between them is taken as including up to next completed year of age (e.g., cumulated percent for women with a difference of 19 years actually includes all women below 20 years of age at first marriage). Therefore, an adjustment is made to the interpolated median by increasing the interpolated value by one year.

Since the median is based on all women, including those who have never been married or lived in a consensual union, there may not be a median for younger cohorts of women (since fewer than 50 percent of the cohort may have been married or lived in a consensual union).

SEXUAL INTERCOURSE

AGE AT FIRST SEXUAL INTERCOURSE

Statistics: *Percent Distribution, by Years of Age at First Sexual Intercourse, Total, and According to Current Age Group*

Definition

Percents of women who first had sexual intercourse before attaining the specified ages.

A. Coverage: Women of all marital statuses.

B. Numerator: Number of women who had their first sexual intercourse before they attained the specified age group.

C. Denominator: Total number of women, including those who never had sexual intercourse.

Calculation

A. Age at first sexual intercourse: Reported by woman as age in completed years or when she began living with first husband or consensual partner. The latter response has been converted to age in completed years in the DHS standard recode.

B. Numerators: Number of women below specified age categories. Women who have never had sexual intercourse are included in a separate category.

C. Denominator is the number of women of all marital statuses. Ever-married samples denominators are adjusted by the all women factors. However, see note below.

D. Numerators for each age group category are divided by the same denominator and multiplied by 100 to obtain percentages.

Handling of Missing Values

Women with unknown and missing ages at first sexual intercourse are excluded from both the denominator and the numerators.

Notes and Considerations

Percentages add up to 100 percent.

Percents are not calculated for the age at first sexual intercourse categories by cohorts of women where the youngest member of the cohort has not yet completed the oldest age of the category. (For example, percents for the cohort of women 20–24 at the time of the survey will not be calculated for age group categories 20–21, 22–24, and 25+ because some women who have never had intercourse are still in or below those age ranges and could still have their first intercourse within the age group categories).

In most of the ever-married DHS surveys, age at first sexual intercourse was not asked. In the other ever-married surveys, never-married women are assumed not to have had sexual intercourse.

Statistics: *Median Age at First Sexual Intercourse*

Definition

Median age in years at first sexual intercourse.

A. Coverage: Women of all marital statuses.

B. Median: Interpolated calculated median.

Calculation

A. Age at first sexual intercourse: Reported by woman as age in completed years or when she began living with first husband or consensual partner. The latter response has been converted to age in completed years in the DHS standard recode.

B. Numerators: Number of women within single year of age categories. Women who never had sexual intercourse are included in a separate category.

C. Denominator is the number of women of all marital statuses. Ever-married samples denominators are adjusted by the all women factors.

D. Numerators for each age category are divided by the same denominator and multiplied by 100 to obtain percents.

E. Medians are calculated from cumulated single year percent distributions of age at first sexual intercourse. Median is linearly interpolated between the age values by which 50 percent or more of the women had their first sexual intercourse.

Handling of Missing Values

Women with unknown and missing ages at first sexual intercourse are excluded from both the denominator and the numerators.

Notes and Considerations

The reported age at first sexual intercourse is in completed years (e.g., women reporting 19 years of age at first sexual intercourse actually includes all women who had their first sexual intercourse under 20 years of age). Therefore an adjustment is made to the interpolated median by increasing the interpolated value by one year.

Since the median is based on all women, including those who never had sexual intercourse, there may not be a median for younger cohorts of women (since fewer than 50 percent of the cohort may have had sexual intercourse).

In most of the ever-married DHS surveys, age at first sexual intercourse was not asked. In the other ever-married surveys, never-married women are assumed not to have had sexual intercourse.

POSTPARTUM EXPOSURE

POSTPARTUM AMENORRHEA, ABSTINENCE, AND INSUSCEPTIBILITY

Statistics: ***Percentage of Births in the Three Years Preceding the Survey for Which Mothers Are a) Postpartum Amenorrheic, b) Abstaining from Sexual Intercourse and c) Insusceptible to Pregnancy, by Number of Months Since Birth***

Definition

A. Coverage:

1. Population base: All births within 0 to 35 months before the survey.

2. Time period: Current status at the time of the survey (interview date).

B. Numerators:

1. Number of births at specified times before the survey for which the mother is postpartum amenorrheic (i.e., her menstrual period has not resumed since the birth).

2. Number of births at specified times before the survey for which the mother is postpartum abstaining (i.e., mother has not resumed sexual intercourse since the birth)

3. Number of births at specified times before the survey for which the mother is either postpartum amenorrheic or abstaining.

C. Denominator: Number of births at specified times before survey.

Calculation

A. Numerators:

1. Number of last births, grouped by two-month intervals before the survey, as determined by difference in interview date and birth date (in CMC), where the mother is amenorrheic since the birth. For last birth, amenorrhea status is from the mother's report of her status at the time of the interview. (For all other births, the mother is assumed not to be amenorrheic since the birth.)

2. Number of last births, grouped by two-month intervals before the survey, as determined by difference in interview date and birth date (in CMC), where the mother has abstained from sexual intercourse since the birth. For last birth, abstinence status is from the mother's report of her status at the time of the interview. (For all other births, the mother is assumed not to have been abstaining since the birth.)

3. Number of last births, grouped by two-month intervals before the survey, as determined by difference in interview date and birth date (in CMC), where the mother is either amenorrheic or has abstained from sexual intercourse since the birth. For last birth, insusceptibility is from the mother's report of her amenorrhea and sexual abstinence status at the time of the interview. (For all other births, the mother is assumed not to have been insusceptible since the birth.)

B. Denominator: Number of all births (last and other) grouped by two-month intervals before survey, as determined by difference in interview date and birth date (in CMC).

C. Percentages: Numerators are divided by the denominator and multiplied by 100.

Handling of Missing Values

1. Birth dates were imputed if missing. Dates of interview are not allowed to have missing data.

2. Women with missing reports of amenorrhea or sexual abstinence are considered to be not amenorrheic or not abstaining, respectively.

Notes and Considerations

Assumptions

Only the last of multiple births are considered. The other births are ignored in both numerators and denominator.

Births of both living and dead children are included. Survival status is ignored.

Statistics: *Median and Mean Durations of Postpartum Amenorrhea, Abstinence, and Insusceptibility*

Definition

A. Coverage:

1. Population base: All births in the time period 0 to 35 months before the survey.

2. Time period: Current status at the time of the survey (interview date).

Calculation

Utilize the numerators and denominators calculated for the proportions amenorrheic, abstaining and insusceptible by time since birth, as given above. Numerators and denominators are each smoothed by a three-group moving average. For example, the value of the numerator and denominators for births that occurred 4–5 months before the interview is the average of groups 2–3, 4–5, and 6–7 months. Groups < 2 months and 34–35 months are not smoothed. Smoothed proportions are calculated by dividing the smoothed numerator values by the smoothed denominator values of each two-month group.

Median

Beginning with the lowest-time-since-birth group, each group is examined to see whether the proportion amenorrheic, abstaining, or insusceptible is less than 0.5. Value of median is determined by linear interpolation of percentage of first group below 0.5 and previous group percentage using the following formula:

$$\textbf{median} = m_{i-1} + (p_{i-1} - 0.5)/(p_{i-1} - p_i) \times (w_i),$$

where p_i is the proportion amenorrheic, abstaining, or insusceptible for the first group, where the proportion is below 0.5, p_{i-1} is the proportion amenorrheic, abstaining, or insusceptible for the preceding group, m_{i-1} is the midpoint value for the preceding group, and w_i is the time width of the group taken as the difference between the midpoint value of the current group and the preceding group.

For all groups except the first, the midpoint values are ½ month above the lower limit. This midpoint value comes about because the time since birth is calculated as the difference between birth dates and date of interview that are accurate for month only (day of month is not ascertained).

For example, if the date of interview were April 2001, the interview could have occurred at any time during the month, from the 1st to the 30th. The same holds true for a birth that occurred in January 2001, at any time between the 1st and the 31st of the month. Thus, the difference in time between the date of birth and the date of interview could be between 60 days and 120 days. Assuming a constant distribution by day of month for interviews and for births, the midpoint is 90 days or three months, which is the value of the difference in the century-month codes of the dates. The midpoint value for the group of the difference of 2 months and 3 months together is therefore 2.5.

The first group is special. The value of the previous group is assumed to be 100 percent since all women are assumed to be amenorrheic and abstaining on the day of birth. Moreover, births that occur in the month of interview can only come before the date of interview, rather than on any day of the month of interview. Based on simulations, 0.75 is chosen as the midpoint value for the group < 2 months (i.e., 0–1). The midpoint of the previous group is 0 (day of interview).

An example of the median calculation is given below:

Example:

Group	P_i	Midpoint	Width
<2	0.90	.75	0.75
2–3	0.70	2.5	1.75
4–5	0.55	4.5	2.00
6–7	0.43	6.5	2.00

Median duration = 4.5 + (.55–.5)/(.55–.43)*(6.5–4.5) = 5.33 months

Mean

The mean duration is the accumulation over all groups of the proportions amenorrheic, abstaining, or insusceptible multiplied by the width of the time-since-birth group.

$$\textbf{mean duration} = \Sigma\, p_i\, w_i$$

Handling of Missing Values

1. Birth dates were imputed if missing. Dates of interview are not allowed to have missing data.

2. Women with missing reports of amenorrhea or sexual abstinence are considered to be not amenorrheic or not abstaining, respectively.

Notes and Considerations

Assumptions

Medians and means are based on current status of mothers of the births. The distributions of the proportions of births by month of birth of the child are analogous to the l_x column of the synthetic life table. The basic assumption is that there has been no change in the proportions amenorrheic, abstaining, and insusceptible over time so that proportions that are calculated from births with small times since birth are the same as those of births that occurred earlier when those earlier births had those durations. For the short period of 36 months, this assumption is very likely to be approximately correct. The proportions are assumed to be 1 at the time of birth and to decrease monotonically with time since birth. The time at which the proportions decline to 0.5 is taken as the median. The mean is taken similarly from the l_x column of the life table, which in this case is the series of proportions by time since birth. Because the sum of proportions equals 1, there is no need to divide the sum of the proportions times the width of the interval.

Because of the relatively small number of births in each month before the interview, the months are first grouped and then smoothed to reduce random fluctuations due to sampling variance.

Mean

Truncated Mean: Because of the limitation to births that occurred within the three years preceding the survey, the mean is truncated if there are mothers who are amenorrheic or abstaining longer than three years after their last birth. It is very unlikely that the proportions are more than negligible after 35 months since birth, except for those women who are not really postpartum amenorrheic or abstaining (more likely for older women whose postpartum amenorrhea blended into menopause or whose postpartum sexual abstinence blended into terminal abstinence). Therefore, it is felt that the truncated mean is very close to the full mean.

The measures are based on all births that occurred within the three years preceding the interview, including last and other births, surviving or not.

Decisions on Alternatives

Current status medians and means are used instead of measures based on the recall of durations because of the severe heaping (digit preference) on multiples of 3 and 6 months in the recall data.

In the current calculation of median and mean durations, a woman can contribute more than once if she had more than one birth in the three years preceding the survey. An alternative calculation would be to base the medians and means on women rather than births. In this alternative, each woman is represented only once, which is equivalent to durations based on only the last birth. Estimates of durations based on last births (open interval durations) are thought to overestimate average durations of amenorrhea and abstinence since they almost always exceed those based on recall after other births (closed interval durations).

MENOPAUSE

Statistics: *Percentage of Women Who Are Menopausal*

Definition

A. Coverage:

1. Population base: Women of all marital statuses, between 30 and 49 years of age.

2. Time period: Date of interview.

B. Numerator: Number of women whose last menstruation occurred six or more months before the survey or before the last birth, who have had a hysterectomy, who declare that they are menopausic, or who never menstruated. Excludes women who are currently pregnant or postpartum amenorrheic.

C. Denominator: Number of women of all marital statuses

Calculation

Numerator is divided by denominator and multiplied by 100 to get percentages.

Handling of Missing Values

Women with inconsistent or missing values and “don’t know” on time since last menstrual period are not considered menopausic.

Notes and Considerations

Since the purpose of the statistic is to measure biological non-exposure to the risk of pregnancy, which increases with age, women who had a hysterectomy and who never had a period are included, as well as those who are truly menopausal. The lack of a period for six months (not postpartum) is taken as a prime indication of menopause in older women.

FERTILITY PREFERENCES

NEED FOR FAMILY PLANNING—CURRENTLY MARRIED WOMEN

Statistics: ***Percentages of Currently Married Women with a) Unmet Need for Family Planning, b) with Met Need for Family Planning, and c) by Total Demand for Family Planning, According to Whether for Spacing, for Limiting, and Total***

Definition

A. Coverage:

1. Population base: Currently married women.

2. Time period: Time of survey.

Fecundity—A woman is assumed to be fecund unless she declares that she is infecund, had a hysterectomy, or is menopausic. Also considered infecund is a woman who is neither pregnant nor postpartum amenorrheic but who has not had a menstruation for six or more months. Also infecund are women who, while married and not using contraception during the past five years, have not had a birth and are not currently pregnant.

B. Numerators—Number of women with an unmet need for contraception

1. For spacing—Number of women who are not using a method of contraception, and

   a) are fecund and say they want to wait two or more years for their next birth, or

   b) are fecund and are unsure whether they want another child, or

   c) are fecund, want another child, but are unsure when to have the birth, and say that a pregnancy in the next few weeks would be a problem, or

   d) are pregnant not due to a contraceptive failure and who say that the pregnancy was mistimed (would have wanted to wait), or

   e) are postpartum amenorrheic within six months of the birth and say that the birth was mistimed (would have wanted to wait).

2. For limiting—Number of women who are not using a method of contraception, and

   a) are fecund and say they do not want any more children, or

   b) are pregnant not due to a contraceptive failure and say that they did not want to become pregnant, or

   c) are postpartum amenorrheic within six months of the birth and say that they did not want the birth.

3. Number of women currently using a method of contraception.

4. Number of women not using contraception who are currently pregnant or postpartum amenorrheic (within six months of birth) due to a failure of contraception and who did not want to become pregnant at the time of conception.

5. Total demand for family planning: sum of numbers of women with an unmet need for contraception, using contraception, or an unwanted pregnancy or postpartum amenorrhea due to a contraceptive failure.

C. Denominator: Number of currently married women.

Calculation:

First select only currently married women for numerators and denominator.

A. Numerators:

Unmet Need

1. Select women who are not currently using a contraceptive method.

2. For surveys following the "A" core questionnaire, exclude women who are pregnant or postpartum amenorrheic for less than six months, whose pregnancy or last birth was the result of a failure of contraception (given as a reason for discontinuing contraception in the reproductive calendar). For surveys following the "B" core questionnaire, there is no way to determine contraceptive failure.

3. Exclude women who are infecund. A woman is infecund if

   a) She declares that she had a hysterectomy or is menopausic in responses to the questions on why she does not intend to use contraception in the future or when her last period occurred.

   b) She is neither pregnant nor postpartum amenorrheic and has not had a menstruation for at least six months, or never menstruated, or her last menstruation occurred before her last birth.

   c) In surveys following the "A" core questionnaire, she has not had a birth in the preceding five years, has not used contraception during that time, and has been continuously married during that time (derived from the reproductive calendar).

   d) In surveys following the "B" core questionnaire, she has not had a birth in the preceding five years, has never used contraception, and was first married more than five years preceding the survey.

4. Unmet need for spacing

   a) Include women who are neither pregnant nor less than six months postpartum amenorrheic and who

      i. Say they want to wait at least two years for their next birth or,

ii. Say they do not know whether they want another child or when they want the next child, and say that it would be a problem if they became pregnant within the next few weeks.

b) Include women who are pregnant (not due to a contraceptive failure) who say that the current pregnancy was wanted later.

c) Include women who are less than six months postpartum amenorrheic and who say that the last birth was wanted later.

5. Unmet Need for Limiting

a) Include women who are neither pregnant nor less than six months postpartum amenorrheic and say they do not want another birth.

b) Include women who are pregnant (not due to a contraceptive failure) who say that they did not want another child before they became pregnant.

c) Include women who are less than six months postpartum amenorrheic and who say that they did not want another child before they became pregnant.

6. For total unmet need

Include women classified in the numerator of unmet need for spacing or for limiting (sum of the two numerators).

Met Need

1. Select women who are currently using a contraceptive method.

2. Met need for limiting—Include women who want no more children, are using sterilization (female or male), or declare themselves as infecund.

3. Met need for spacing—Include women using contraception who are not in the category of met need for limiting.

Contraceptive Failure (only in surveys using "A" core questionnaires)

1. Select women who are not currently using a contraceptive method, who are either pregnant or less than six months postpartum amenorrheic, and whose pregnancy or birth was the result of a contraceptive failure (from the reproductive calendar in "A" core questionnaire-based surveys).

2. Failure to Space

a) Include women who are pregnant (not due to a contraceptive failure) who say that the current pregnancy was wanted later.

b) Include women who are less than six months postpartum amenorrheic who say that the last birth was wanted later.

3. Failure to Limit

   a) Include women who are pregnant who say that they did not want another child before they became pregnant.

   b) Include women who are less than six months postpartum amenorrheic who say that they did not want another child before they became pregnant.

4. For Total Failure

   Include women classified in the numerator for failure to space or failure to limit (sum of the two numerators).

Total Demand for Family Planning

In "A" core questionnaire surveys, sum of numerators for unmet need, met need, and contraceptive failure. In "B" core questionnaire surveys, sum of numerators for unmet need and met need (failure not measured and assumed zero).

A. Denominator: Number of currently married women.

B. Divide all numerators by denominators and multiply by 100 to get percentages.

C. Percentage of demand satisfied: Divide percentage for total met need by percentage total demand for family planning and multiply by 100.

Handling of Missing Values

Several variables make up the construction of the standard recode variable for need for family planning. Only if a respondent cannot be classified in any category due to missing information required for a category, then the recode variable takes on a missing value for that respondent. Respondents with missing values on the need recode variable are excluded from both numerators and denominator.

Notes and Considerations

The classification of need for family planning is included as a standard variable (V626) in the DHS standard recode data sets. The information on calculation given above is for reference and does not need to be recalculated to produce tables.

Assumptions

In "B" core questionnaire surveys, there is no information on pregnancies and births due to contraceptive failure, which is therefore assumed to be zero in the calculation of total demand.

Changes over Time

Several changes occurred in the calculation of need for family planning. The latest are the following: For later phase III and phase IV surveys, declarations of infecundity were included, as well as the lack of births in the preceding five years, as an indication of infecundity. Also, the question on whether it would be a problem if a woman found out that she were pregnant in the next few weeks was added in the determination of unmet need for women who are undecided about whether they want another child or when they want another child. Due to these changes, comparisons of unmet need and demand for family planning between surveys may not be valid if based on country reports.

NEED FOR FAMILY PLANNING—NOT CURRENTLY MARRIED WOMEN AND ALL WOMEN

Statistics: ***Percentages of Women Who Are Not Currently Married and Women of All Marital Statuses with a) Unmet Need for Family Planning, b) With Met Need for Family Planning, and c) By Total Demand for Family Planning, According to Whether for Spacing, for Limiting, and Total***

Definition

A. Coverage:

1. Population base: Women who are not currently married and all women.

2. Time period: Time of survey.

Fecundity—A woman is assumed to be fecund unless she declares that she is infecund, had a hysterectomy, or is menopausic. Also considered infecund is a woman who is neither pregnant nor postpartum amenorrheic but who has not had a menstruation for six or more months. Also infecund are women who while married and not using contraception during the past five years have not had a birth and are not currently pregnant.

B. Numerators: Number of women with an unmet need for contraception.

1. For spacing—Number of women who are not using a method of contraception, had sexual intercourse in the 30 days or 4 weeks preceding the survey (depending on whether response is in days or weeks), and

   a) Are fecund and say they want to wait two or more years for their next birth, or

   b) Are fecund and are unsure whether they want another child, or

   c) Are fecund, want another child but are unsure when to have the birth, and say that a pregnancy in the next few weeks would be a problem, or

   d) Are pregnant not due to a contraceptive failure and say that the pregnancy was mistimed (would have wanted to wait), or

   e) Are postpartum amenorrheic within six months of the birth and say that the birth was mistimed (would have wanted to wait).

2. For limiting—Number of women who are not using a method of contraception, have had sexual intercourse in the 30 days or 4 weeks preceding the survey (depending on whether response is in days or weeks), and

   a) Are fecund and say they do not want any more children, or

   b) Are pregnant not due to a contraceptive failure and say that they did not want to become pregnant, or

c) Are postpartum amenorrheic within six months of the birth and say that they did not want the birth.

3. Number of women currently using a method of contraception.

4. Number of women not using contraception who are currently pregnant or postpartum amenorrheic (within six months of birth) due to a failure of contraception who did not want to become pregnant at the time of conception.

5. Total demand for family planning: sum of numbers of women with an unmet need for contraception, using contraception, or an unwanted pregnancy or postpartum amenorrhea due to a contraceptive failure.

C. Denominators: Number of women who are not currently married and number of all women.

Calculation

For not currently married women, first select only women who are not currently married for numerators and denominator. For all women, do not select.

A. Numerators:

Unmet Need

1. Select women who are not currently using a contraceptive method.

2. Select women who have had sexual intercourse in the 30 days (or four weeks) preceding the survey (depending on whether response is in days or weeks).

3. For surveys following the "A" core questionnaire, exclude women who are pregnant or postpartum amenorrheic for less than six months and whose pregnancy or last birth was the result of a failure of contraception (given as a reason for discontinuing contraception in the reproductive calendar). For surveys following the "B" core questionnaire, there is no way to determine contraceptive failure.

4. Exclude women who are infecund. A woman is infecund if

   a) She declares that she had a hysterectomy or is menopausic in responses to the questions on why she does not intend to use contraception in the future or when her last period occurred.

   b) She is neither pregnant nor postpartum amenorrheic and has not had a menstruation for at least six months, never menstruated questionnaire, or her last menstruation occurred before her last birth.

   c) In surveys following the "A" core questionnaire, she has not had a birth in the preceding five years, has not used contraception, during that time, and has been continuously married during that time (derived from the reproductive calendar).

   d) In surveys following the "B" core questionnaire, she has not had a birth in the preceding five years, has never used contraception, and was first married more than five years preceding the survey.

5. Unmet need for spacing.

   a) Include women who are neither pregnant nor less than six months postpartum amenorrheic and who

      i. Say they want to wait at least two years for their next birth or

      ii. Say they do not know whether they want another child or when they want the next child and say that it would be a problem if they became pregnant within the next few weeks.

   b) Include women who are pregnant (not due to a contraceptive failure) who say that the current pregnancy was wanted later.

   c) Include women who are less than six months postpartum amenorrheic who say that the last birth was wanted later.

6. Unmet need for limiting.

   a) Include women who are neither pregnant nor less than six months postpartum amenorrheic and say they want do not want another birth or

   b) Include women who are pregnant (not due to a contraceptive failure) who say that they did not want another child before they became pregnant.

   c) Include women who are less than six months postpartum amenorrheic and who say that they did not want another child before they became pregnant.

7. For total unmet need.

   Include women classified in the numerator of unmet need for spacing or for limiting (sum of the two numerators).

Met Need

1. Select women who are currently using a contraceptive method.

2. Met need for limiting—Include women who want no more children, are using sterilization (female or male) or declare themselves as infecund.

3. Met need for spacing—Include women using contraception who are not in category of met need for limiting.

Contraceptive Failure (only in surveys using "A" core questionnaires)

Select women who are not currently using a contraceptive method and who are either pregnant or less than six months postpartum amenorrheic and whose pregnancy or birth was the result of a contraceptive failure (from the reproductive calendar in "A" core questionnaire-based surveys).

1. Failure to space.

   a) Include women who are pregnant (not due to a contraceptive failure) who say that the current pregnancy was wanted later.

b) Include women who are less than six months postpartum amenorrheic who say that the last birth was wanted later.

2. Failure to limit.

a) Include women who are pregnant who say that they did not want another child before they became pregnant.

b) Include women who are less than six months postpartum amenorrheic who say that they did not want another child before they became pregnant.

3. For total failure.

Include women classified in the numerator of failure to space or failure to limit (sum of the two numerators).

Total Demand for Family Planning

In "A" core questionnaire surveys, sum of numerators for unmet need, met need, and contraceptive failure. In "B" core questionnaire surveys, sum of numerators for unmet need and met need (failure not measured and assumed zero).

B. Denominators: Number of women who are not currently married and number of women of all marital statuses. Women are included in both denominators regardless of whether or not they had sexual intercourse in the 30 days or 4 weeks preceding the survey.

C. Divide all numerators by denominators and multiply by 100 to get percentages.

D. Percentage of demand satisfied: Divide percentage for total unmet need by percentage total demand for family planning and multiply by 100.

Handling of Missing Values

Several variables make up the construction of the standard recode variable for need for family planning. Only if a respondent cannot be classified in any category due to missing information required for a category, then the recode variable takes on a missing value for that respondent. Respondents with missing values on the need recode variable are excluded from both numerators and denominator. Not currently married respondents with missing values on date of the last sexual intercourse are considered not to have had sexual intercourse in last 30 days.

Notes and Considerations

The classification of need for family planning is included as a standard variable (V626) in the DHS standard recode data sets. The information on calculation given above is for reference and does not need to be recalculated to produce tables.

In ever-married surveys, need for family planning for women who are not currently married is assumed to be zero and cannot be properly calculated since time since last sexual intercourse is not calculated. For this same reason, need for family planning for women of all marital statuses is unenlightening.

Assumptions

Women who are not currently married and who have not had sexual intercourse within the 30 days preceding the survey (or four weeks if reported in weeks) are assumed to have no need for family planning.

In “B” core questionnaire surveys, there is no information on pregnancies and births due to contraceptive failure, which is therefore assumed to be zero in the calculation of total demand.

Changes over Time

Several changes occurred in the calculation of need for family planning. The latest are the following: For later phase III and phase IV surveys, declarations of infecundity were included, as well as the lack of births in the preceding five years, as an indication of infecundity. Also, the question on whether it would be a problem if a woman found out that she were pregnant in the next few weeks was added in the determination of unmet need for women who are undecided on whether they want another child or when they want another child. Due to these changes, comparisons of unmet need and demand for family planning between surveys may not be valid if based on country reports.

In previous published reports, the denominator for women not currently married was limited to women who had had sexual intercourse in the 30 days (or four weeks) preceding the survey.

WANTED FERTILITY RATES

Statistics: *Age-Specific Wanted Fertility Rate (ASWFR) and Total Wanted Fertility Rate (TWFR): Based on Age-Period Rates*

Definition

A. Coverage: All women aged 15–49 years in seven five-year age groups (15–19, 20–24, 45–49 years).

B. Numerators: Number of wanted births that occurred in the 1–36 months-period before the survey to women in the age group at the time of the birth.

C. Denominators: Number of women-years of exposure in the 1–36 month-period before the survey of women in the age group.

D. The age-specific wanted fertility rates (ASFWRs) are the quotients of the numerators divided by the denominators. The total wanted fertility rate (TWFR) is the sum of the seven age-specific wanted fertility rates multiplied by five.

Calculation

A. Numerator: Births are tabulated according to period of birth, age of mother at the time of the birth, and wantedness of the birth.

1. Period of birth—The period of birth is calculated as the difference in months between the date of interview and the date of birth, both in century-month code format (CMC). Births are included in the tabulation if they occur 1 to 36 months before the survey.

2. Age of mother at the time of the birth—The difference in months between the date of birth of the child and the date of birth of the mother, both in CMC. The difference is then divided by 60 and truncated to whole numbers to form the age groups. Births are tabulated by age group.

3. Wantedness—A birth is considered wanted if the number of living children at the time of conception of the birth is less than the ideal number of children as reported by the respondent. For the calculation of the time of conception, nine months are subtracted from the date of birth. A preceding child is considered living at the time of conception if it was born before the date of conception and a) is living at the time of the survey or b) died at or before the date of conception (calculations based on century-month codes). The date of death of preceding children who died is calculated by adding the age at death in months to the child's birth date in century-month code. If a child's age at death is reported in years, then the child is assumed to have died at ages 6 months higher than the number of years (30 months for children who died at 2 years of age, 42 months for children who died at 3 years of age, etc.).

B. Denominator: Woman-years of exposure is calculated as the sum of the number of months exposed in the five-year age group during the time period. A woman can contribute exposure to just two five-year age groups during the 36-month period.

1. Higher age group: A woman's age at the end of the period determines the higher age group. The higher age group is calculated by subtracting the women's date of birth from the date of interview (in CMC), dividing the difference by 60, and truncating to a whole number. The number of months spent in the higher age group is the difference in months between her age at the end of the period of exposure (date of interview less one month) and the lower age limit of the age group plus one month. If the number of months in the age group is less than the duration of the time period (36 months), then the woman contributes exposure to both the higher age group and the next lower age group.

2. Lower age group: The contribution to the lower age group is 36 less the number of months in the higher age group. If the number of months in the higher age group is greater than or equal to the duration of the time period (≥ 36 months), then the exposure in the higher group is the duration and the exposure in the lower age group is zero.

C. Tabulation: Each woman is tabulated twice, once according to her higher age group accumulating the exposure she contributes to that group and once in the lower age group accumulating lower age group exposure. (In ISSA, the same table is used, effectively summing the accumulations within each age group.) For ever-married samples, the exposure is adjusted to represent all women by dividing by the woman's "all-woman factor" (AWFACTOR), which is derived from the proportion of ever-married women from the household data file. See section on all women factors for details in their calculation.

D. Examples

Example 1—A woman interviewed in December 2001, born in May 1970. Her CMC date of interview is 12*(2001 – 1900)+12 = 1224. The date of the end of the period of exposure is 1224 – 1 = 1223. Her CMC date of birth is 12*(1970 – 1900)+5 = 845. Her age in months at the end of the period is 1223 – 845 = 378. The age group at the end of the period is 378/60 = 6.3, truncated to 6. This represents age group 30–34 years (30 = 6*5 years interval). The number of months in this age group is 378 – 6*60 + 1 = 19 months. Since this is less than the total number of months during the period (36 months), she contributed 19 months to age group 30–34 during the period and 36 – 19 = 17 months to the age group 25–29 during the period.

Example 2—A woman interviewed in December 2001, born in March 1967. Her CMC date of interview is 12*(2001 – 1900)+12 = 1224. The date of the end of the period of exposure is 1224 – 1 = 1223. Her CMC date of birth is 12*(1967 – 1900)+3 = 807. Her age in months at the end of the period is 1223 – 807 = 416. The age group at the end of the period is 416/60 = 6.93, truncated to 6. This represent age group 30–34 years (30 = 6*5 years interval). The number of months in this age group is 416 – 6*60 = 56 months. Since the number of months in this age group is greater than 36 months, she contributed 36 months of exposure to age group 30–34 during the period and no exposure to the next lower age group during the period.

E. ASWFR**—**The age-specific wanted fertility rate is calculated as the quotient of the numerator divided by the denominator for each age group, multiplied by 1000. The result is an average rate over the 36-month period, expressed as an annual rate.

F. TWFR**—**The total wanted fertility rate is calculated by summing the ASWFRs, multiplying by 5, and dividing by 1,000.

Handling of Missing Values

The total number of children to which a woman has given birth is recorded obligatorily by the interviewer; no unknown numbers of children are allowed. There are six values involved in the calculation of ASWFR: interview date, birth date of woman, birth dates of children, survival of children, ages at death of dead children, and ideal number of children. The interview date is always known from fieldwork dates. If missing or unknown, the birth dates of interviewed women and her children are imputed before the formation of the standard recode file. See Croft, 1999 on date imputation. Children's survival statuses are not allowed to be missing. For dead children with missing age at death, the age at death in months is imputed in the standard recode dataset, using a hot deck technique based on birth order. See Croft, 1999 for the imputation procedure.

For ideal number of children, non-numeric and "don't know" responses are considered to be high numbers, so that all births are considered wanted. Births to women with missing information on ideal number of births are considered as unwanted.

Notes and Considerations

The calculation of age-specific and total wanted fertility rates is the same as age-specific and total fertility rates with the addition classification of births to wanted or not wanted at the time of conception.

Births to women younger than 15 years or older than 49 years at the time of the birth are not generally included. In a few specific countries, births to girls 10–14 are included.

Births in the month of interview are excluded. This exclusion is because this month does not represent a full month, but is censored by the date of interview.

A three-year (36 month) time period is taken for calculating current AWFSR. This period is a compromise between the need for recency and reduction of sampling variation. This time period was selected during the World Fertility Survey, when sample sizes usually comprised about 5,000 women. For comparability over time and across surveys, this period has been maintained by DHS.

No adjustment is made for truncation by age. (Women who are at most 49 years old at the time of the interview were 48 years old the year before and 47 years old two years before.) The reason no adjustment is made is that the tiny probability of giving birth by women 48 and 49 years of age outweighs the complication of doing the adjustment by single year of age.

In line with general DHS policy, no adjustment is made for possible omission or date misreporting of the dates of birth of children or misreporting of the date of birth of the woman.

For ever-married samples, it is assumed that never-married women have not had any births. Only the denominator of the rates is adjusted to estimate the number of all women.

Reference

Croft, T. 1991. Date Editing and Imputation. *Demographic and Health Surveys World Conference Proceedings, II:* 1337–1356, Columbia, Maryland: IRD/ORC Macro.

INFANT AND CHILD MORTALITY

EARLY CHILDHOOD MORTALITY RATES

Methodology of DHS Mortality Rates Estimation

There are two principal categories of estimation methods for calculating infant and child mortality rates: direct and indirect. Direct methods of calculation use data on the date of birth of children, their survival status, and the dates of death or ages at death of deceased children. Indirect methods use information on survival status of children to specific age cohorts of mothers.

The direct methods require data that are usually obtained only in specifically designed surveys with birth histories or from vital statistics systems (which are generally deficient in less-developed countries). The indirect methods can utilize data that are commonly collected in censuses and many general surveys: the number of children ever born and the number living to women and age. Unlike the direct methods, the indirect methods are very dependent upon several assumptions that may or may not hold true: little or no change in fertility levels and age patterns, no change or a linear decline in mortality, and a pattern of mortality by age that conforms to known "families," basically derived from European experience.

Both types of methods can suffer from errors in data. Both methods suffer, probably equally, from the omission of deceased children. Estimation of infant mortality, using direct methods, depends on the correct reporting of age at death as under or over one year. The heaping of deaths at age 12 months is common, and to the extent that it causes a transfer of deaths across the one-year boundary, infant mortality rates may be somewhat underestimated. On the other hand, under two and under-five mortality rates are little affected by heaping. The misreporting of birth dates can also affect direct estimates, but unless the misreporting occurs much more for deceased children, mortality rates are little affected.

Estimates of indirect methods can suffer if women do not know their ages, as is common in many less-developed countries. The bias that results can be even greater if age is estimated on the basis of characteristics linked directly or indirectly to mortality levels; for example, number of children ever born. Censuses and surveys not specifically designed to gather data for mortality estimation have also been shown to be very prone to suffer from omission of dead children. On the other hand, there have also been some cases where both still and live births have been included in the answer to the question on children ever born, thus leading to overestimates of mortality rates. Violations of the assumptions of the indirect methods also commonly occur.

An implicit assumption of the indirect methods is that the births of a cohort of women represent the children born in a time period. Recent and ongoing work show that births to women 20–24 (and in some cases to women 25–29) have more elements of high risk of mortality than do all children born within the last five years of a survey.

Another problem with indirect methods is the location of the estimate in time. In truth, the indirect methods estimate the probability of dying on the basis of experience that can extend over many years, resulting in an average over that period. Depending on changes in fertility and trends in mortality, the methods used to place the mortality estimate in time can be more or less in error.

Because of the number of assumptions that may or may not hold, but that need to be made to use the indirect methods and the limited amount of information that they provide, DHS has chosen to use direct estimation methods. There are three variants of direct estimation methods for estimating infant mortality rates and other rates:

1. A vital statistics approach in which the numbers of deaths to children under age 12 months in a particular period are divided by the numbers of births in the same period. What is estimated is a rate of mortality but not a probability; a variation in the number of births with time will change the rate without changes in the underlying probabilities. To correct this, separation factors would need to be used, which would have to come from the other variants.

2. A true cohort life table approach in which deaths to children under age 12 months of a specific cohort of births are divided by the number of births in that cohort. This procedure gives true probabilities of death, but has the drawback that all children in the cohort must have been born at least 12 months before the survey to be fully exposed to mortality, thus not taking into account the most recent experience. This requirement of full exposure becomes more limiting the higher the age segment of interest: For under-five mortality rates, only the information on children born five or more years before the survey can be utilized. Another drawback is that true cohort rates are not specific to a particular period at death, but instead relate to the date of birth of the cohort. Therefore the effects of events that affect several cohorts at the same time, for example, a famine appears to be spread out over time.

3. A synthetic cohort life table approach in which mortality probabilities for small age segments based on real cohort mortality experience are combined into the more common age segments. This approach allows full use of the most recent data and is also specific for time periods. It is the method that the DHS has chosen to use.

In this latter approach, there are still a couple of decisions to be made: the length of the individual age segments, how to handle the heap of deaths at 12 months, and deaths with incomplete information on age at death. The information on age at death is recorded on the DHS questionnaire in days for less than 0–30 days, in months for 1–23 months, and in years for two years or higher. The DHS adopted the following age segments 0, 1–2, 3–5, 6–11, 12–23, 24–35, 36–47, 48–59 months (completed ages) for the calculation of the individual probabilities of dying. A different approach could have been to use monthly segments up to two years. This approach would have required much more computer memory, and studies done during the World Fertility Survey have shown the difference to be negligible. No adjustment for heaping on age at death 12 months was done for the rates presented in the survey reports. Because heaping is likely to come unequally from both sides of the one-year boundary and the correct number of deaths to reallocate is unknown, the rates are presented using the information as reported. Indeed, the extent of rounding up probably varies by country and within country, and in some cases all or most of the heaping may be due to reporting only whole year ages for children dying in the second year of life (rounding down or truncation of age). For children with missing information, the ages at death have been assigned according to a “hot deck” technique in which the information of the child of the same birth order and form of reporting (day, month, or year), if available, that precedes in the data file is assigned to the child for whom age at death is missing. This is a quasi-random technique that preserves the variance of responses in the data set.

Statistics: *Neonatal Mortality Rate, Post-Neonatal Mortality Rate, Infant Mortality Rate, Child Mortality Rate (Ages 1–4) and Under-five Mortality Rate*

Definition

A. Coverage:

1. Population base: Live births to respondents.

2. Time period: Five-year periods of time preceding the survey

3. Live birth: The complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of the pregnancy, which, after such separation, breathes or shows any other evidence of life, such as beating of 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. (WHO, 1950 and 1992).

B. Numerators: Number of deaths to live-born children during specified age range and specified time period.

1. Neonatal mortality—Deaths at age 0–30 days (also includes deaths reported at age zero month)

2. Post neonatal mortality—Deaths at ages 1 to 11 months (also includes deaths reported as age 31 to 59 days)

3. Infant mortality—Deaths at ages 0 to 11 months (also includes deaths reported as age zero years)

4. Child mortality—Deaths at age 1–4 years (also includes deaths reported as age 12–59 months)

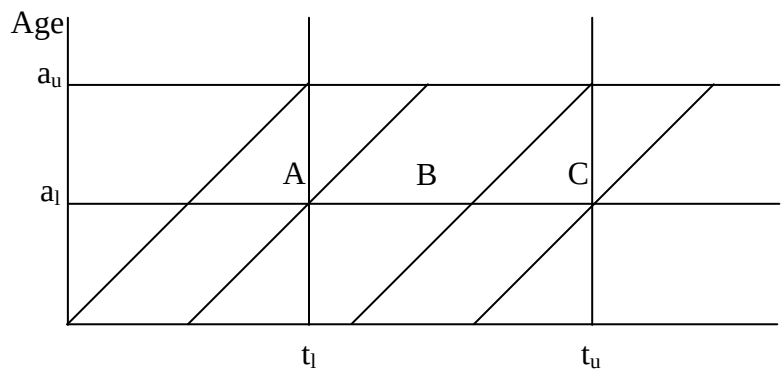
5. Under-five mortality—Deaths at age 0–4 years (also includes deaths reported as age 0–59 months and 0–59 days).

C. Denominators: Number of surviving children at beginning of specified age range during the specified time period.

Calculation

A. Component death probabilities are first tabulated. Then the component death probabilities are combined into the mortality rates. The component death probabilities are calculated for age segments 0, 1–2, 3–5, 6–11, 12–23, 24–35, 36–47, and 48–59 months of completed age.

B. Each component death probability is defined by a time period and an age interval. Within these two parameters, three birth cohorts of children are included, as indicated in the figure below.



C. One cohort of children is completely included and two are partially included. If the upper and lower limits of the age interval are given by a_l and a_u, respectively, and the upper and lower limits of the time period are given by t_u and t_l, respectively, then the three cohorts are defined as children born between dates $t_l - a_u$ and $t_l - a_l$ (cohort A), $t_l - a_l$ and $t_u - a_u$ (cohort B) and $t_u - a_u$ and $t_u - a_l$ (cohort C).

D. Cohorts A and C are only partially exposed to mortality between ages a_l and a_u during time period t_l to t_u. Therefore, account needs to be taken of the partial exposure. Because of the small age intervals of the component probabilities, the assumption is made that the exposure to mortality and deaths of birth cohorts A and C are well represented by taking one-half of the total exposure and one-half of the deaths (with the exception noted below).

E. Numerators: The sum of one-half of the deaths between ages a_l and a_u to children of cohort A, plus all of deaths between ages a_l and a_u to children of cohort B, plus one-half of the deaths between ages a_l and a_u to children of cohort C.

F. Denominators: The sum of one-half of the survivors at age a_l of children of cohort A, plus all of the survivors at age a_l of children of cohort B, plus one-half of the survivors at age a_l of children of cohort C.

G. Component death probabilities are calculated by dividing the numerator for each age range and time period by the denominator for that range and period.

H. Special exception: For the time period that ends with the date of the survey, numerators are calculated as the sum of one-half of the deaths between ages a_l and a_u to children of cohort A, plus all of deaths between ages a_l and a_u to children of cohort B, plus <u>all</u> of the deaths between ages a_l and a_u to children of cohort C. This change is because all of the deaths reported in the survey for cohort C for this time period represent one-half of the deaths that will have occurred to the cohort between ages a_l and a_u.

MORTALITY RATE CALCULATION

Neonatal mortality rate is the component death probability for 0 months of age multiplied by 1000.

A. Infant mortality rate:

1. Calculate the component survival probability by subtracting the component death probability from one.

2. Calculate the product of the component survival probabilities for 0, 1–2, 3–5, and 6–11 months of age.

3. Subtract the product from 1 and multiply by 1000 to get the infant mortality rate. Post neonatal mortality rate: Subtract the neonatal mortality rate from the infant mortality rate.

B. Child mortality rate:

1. Calculate the component survival probability by subtracting the component death probability from 1.

2. Calculate the product of the component survival probabilities for 12–23, 24–35, 36–47, and 48–59 months of age.

3. Subtract the product from 1 and multiply by 1000 to get the child mortality rate.

C. Under-five mortality rate:

1. Calculate the component survival probability by subtracting the component death probability from 1.

2. Calculate the product of the component survival probabilities for 0, 1–2, 3–5, and 6–11, 12–23, 24–35, 36–47, and 48–59 months of age.

3. Subtract the product from 1 and multiply by 1,000 to get the child mortality rate.

Handling of Missing Values

Four variables are used in the calculation of infant and child mortality: date of interview, date of birth of the child, survival of the child, and age at death of the child. The date of interview is not allowed to be missing on the questionnaire. Survival status for children is not allowed to be missing. If missing or unknown, the date of birth and age at death are imputed before the creation of the standard recode. See Croft, 1999 for the imputation procedures.

Notes and Considerations

Typically, mortality rates are calculated for five-year periods preceding the date of the survey. To provide stability in estimates for smaller subgroups, the ten-year period before the survey is used. To calculate the component death probabilities for the ten-year period, the numerators (deaths) for the 2 five-year periods are summed, as are the denominators (survivors) before dividing the numerators by the denominators.

Reference

Croft, T. 1991. Date Editing and Imputation. *Demographic and Health Surveys World Conference Proceedings*, *II*: 1337–1356, Columbia, Maryland: IRD/ORC Macro.

Statistics: *Perinatal Mortality Rate*

Definition

A. Coverage:

1. Population base: Pregnancies of seven or more months to women age 15–49 at time of survey.

2. Time period: Five-year period preceding the survey.

B. Numerator: Number of fetal deaths in pregnancies of seven or more months plus number of deaths of live-born children in the 0–6 days following birth.

C. Denominator: Number of pregnancies of seven or more months that terminated in a fetal death plus pregnancies that ended with a live birth.

Calculation

A. Numerator:

1. Number of stillbirths—From the reproductive calendar, number of pregnancies that lasted seven or more months and terminated in a fetal death in the five years preceding the survey.

2. Number of early neonatal deaths—Number of children who died at age 0–6 days after birth in the five years preceding the survey.

3. Numerator is the sum of (1) plus (2) above.

B. Denominator:

1. Number of stillbirths, given in numerator "A" above.

2. Number of live births in the five years preceding the survey.

3. Denominator is the sum of (1) plus (2) above.

Perinatal Rate—Quotient of numerator divided by denominator multiplied by 1,000.

Handling of Missing Values

Missing values are not allowed for any of the variables that make up the rate. Age at death for living children is imputed if unknown or missing.

Notes and Considerations

The perinatal mortality rate is defined by dividing the number of perinatal deaths (stillbirths and early neonatal deaths) by either the number of live births or by the sum of live births and stillbirths. Both definitions are prevalent in the literature.

The definition of the pregnancy duration for stillbirth in general has changed over time. Originally, it was the product of pregnancies lasting 28 weeks that ended in a fetal death. The duration limit was subsequently lowered to 24, 22, and even 20 weeks. For the purpose of calculating perinatal mortality, however, the definition remains at 28 weeks. DHS asks and records pregnancy duration in months so that the equivalent of seven months is used. The durations of pregnancy are taken as reported by the respondents and do not necessarily have a clinical basis.

Changes over Time

Earlier DHS reports used a somewhat different basis for perinatal rates. Deaths of live-born children were considered early neonatal deaths if they occurred 0 to 7 days after birth, given the large amount of heaping on day 7 and the consideration of the likelihood of rounding “in the first week of life” by interviewers to seven days, due to questionnaire coding requirements. Also, the denominator for the rates was the number of live births rather than the sum of live births and stillbirths.

HIGH-RISK FERTILITY BEHAVIOR RISK RATIOS

Statistics: *Ratio of Children in Elevated Fertility Behavior Risk Categories to Children in Non-Elevated Risk Category*

Definition

A. Coverage:

1. Population base: Children born to all women.
2. Time period: Births in the five years preceding the survey.

B. Numerator: Proportion of dead children born in the five years preceding the survey by risk category.

C. Denominator: Proportion dead of children in non-elevated risk category.

D. Elevated-behavior fertility risk categories:

1. Children born to mothers under age 18 years.
2. Children born to mothers 35 years and older.
3. Children of birth order 4 or higher.
4. Children born less than 24 months since a preceding birth.

E. Unavoidable elevated risk category—First birth order children born to mothers between ages 18 and 34 years.

F. Children not in any elevated risk category

1. Second- and third-birth-order children born to mothers between ages 18 and 34 years
2. The elevated risk categories are also combined into two-way and three-way combinations of elevated risk.

Calculation

A. To calculate relative risk ratios by categories of elevated risk due to fertility behavior, first the proportions dead for each category need to be calculated. For each proportion—

1. Numerator: Number of children born in the five years preceding the survey who died at any age.
2. Denominator: Number of children born in the five years preceding the survey.
3. Proportion dead: Numerator divided by denominator.

B. Then the ratio of the proportions need to be calculated:

1. Risk ratio Numerator: Proportion dead among children in each risk category.

2. Risk ratio denominator: Proportion dead among children not in any elevated risk category (see definition above).

3. Relative risk ratios: Risk category numerators divided by the non-elevated risk category denominator.

Handling of Missing Values

None of the information needed to calculate relative risk ratios is allowed to have missing or unknown data. Note that age at death is not used, only survival status at the time of the survey.

Notes and Considerations

To provide a more robust estimate of the elevated relative risk, proportions dead are used instead of period mortality rates in the calculation of relative risk, since period mortality rates have higher sampling errors. In DHS reports, risk ratios are not shown where the proportion dead is based on fewer than 25 births. Risk ratios where the proportion dead is based on 25 to 49 births are placed in parentheses to indicate large variability due to sampling.

Assumptions

It is assumed that the number of births and death rates do not vary substantially by time period within the preceding five years.

Changes over Time

Originally, the denominator for the risk ratio included the unavoidable risk category of first-order births to mothers between 18 and 24 years of age.

Statistics: *Percent Distribution of Women by Provider of Antenatal Care*

Definition

Percentage of women with a birth in the last five years, distributed by highest type of provider of antenatal care.

A. Coverage: Women who have had one or more births in the five years preceding the survey.

B. Numerators: Number of women who were attended for antenatal care for their last birth, distributed according to the type of provider with the highest level of qualification.

C. Denominator: Number of women with a birth in the last five years.

Calculation

For each type of provider, 100 times the quotient of the numerator divided by the denominator.

Handling of Missing Values

Included in distribution.

Notes and Considerations

Percentages distribution adds up to 100 percent.

The category "Trained nurse/midwife" includes only medically trained and licensed personnel. Traditional birth attendants (also sometimes called midwives) are not included, whether trained or untrained.

The category "Traditional birth attendant/other" includes auxiliary health personnel and cases where the respondent did not know the level of qualification.

NUMBER OF ANTENATAL CARE VISITS AND TIMING OF FIRST VISIT

Statistics: *Percentage of Women with a Birth in the Last Five Years, Distributed by Number of Antenatal Care Visits*

Percentage of women with a birth in the last five years, distributed by number of months pregnant at time of first antenatal care visit.

Median number of months pregnant at time of first antenatal care visit.

Definition

A. Coverage:

Population base: Women who have had one or more births in the five years preceding the survey.

B. Numerators

1. Numbers of women who received antenatal care for their last birth, according to grouped number of visits.

2. Numbers of women who were attended for antenatal care for their last birth, according to grouped number of months they were pregnant at time of first visit.

3. Numbers of women who were attended for antenatal care for their last birth according to the single number of months they were pregnant at time of first visit

C. Denominators:

1. A and B—Number of women with a birth in the last five years.

2. C—Number of women with a birth in the last five years who received antenatal care for their last birth.

Calculation

A. For percentages A and B, 100 times the quotient of the numerators divided by the same denominator.

B. For the median C, first calculate percentages of single months pregnant at first visit by

1. Dividing the numerators B by the denominator C.

2. Cumulate the percentages by single months starting with the lowest value.

3. Linearly interpolate between the number of months immediately before and after where the cumulated distribution exceeds 50 percent to determine the median.

Handling of Missing Values

"Don't know" and missing values included in percent distributions A and B. "Don't know" and missing values excluded from numerators and denominator of percentages for median calculation.

Notes and Considerations

Percent distributions add up to 100 percent.

ANTENATAL CARE

Percentages of mothers who received various components of antenatal care, iron supplements and antimalarials

Statistics: *Percentages of Women With a Birth in the Last Five Years Who Received Key Components of Antenatal Care: Informed of Signs of Pregnancy Complications, Weight Measured, Height Measured, Blood Pressure Measured, Urine Sample Taken, Blood Sample Taken*

Percentages of women with a birth in the last five years who received iron supplements and anti-malaria drugs.

Definition

A. Coverage:

1. Population base: Women who had one or more births in the five years preceding the survey.

2. Time period: Most recent birth in last five years.

B. Numerators:

1. Number of women who received antenatal care for their last birth and who received each of the key components of antenatal care.

2. Number of women who were given iron tablets or syrup and anti-malaria drugs.

C. Denominators:

1. Number of women with a birth in the last five years who received antenatal care for their last birth.

2. Number of women with a birth in the last five years.

Calculation

For percentages, 100 times the quotient of the numerators divided by the denominator.

Handling of Missing Values

"Don't know" and missing values on key components, iron supplement, and anti-malaria drugs excluded from numerators but included in denominators. (That is, assumed did not get.)

Notes and Considerations

Percent distribution adds up to 100 percent.

Prenatal tetanus toxoid injections

Statistics: *Percentage of Women with a Birth in the Last Five Years, Distributed by Number of Tetanus Toxoid Injections Received During the Pregnancy of the Most Recent Birth*

Definition

A. Coverage:

1. Population base: Women who had one or more births in the five years preceding the survey.

2. Time period: Five years preceding the survey.

B. Numerators: Numbers of women distributed, according to the number of tetanus toxoid injections received during the pregnancy of the most recent birth.

C. Denominator: Number of women with a birth in the last five years.

Calculation

For percentages, 100 times the quotient of the numerators divided by the same denominator.

Handling of Missing Values

Included in distribution.

Notes and Considerations

Percent distribution adds up to 100 percent.

DELIVERY CARE

PLACE OF DELIVERY

Statistics: *Percent Distribution of Live Births by Place of Delivery*

Definition

A. Coverage:

Population base: Live births to interviewed women who had one or more births in the five years preceding the survey.

B. Numerators: Numbers of live births distributed according to whether the delivery took place in a public health facility, private health facility, at home, or in another place.

C. Denominator: Number of live births in the last five years.

Calculation

For percents, 100 times the quotient of the numerators divided by the same denominator.

Handling of Missing Values

Included in distribution.

Notes and Considerations

Percent distribution adds up to 100 percent.

Statistics: *Percent Distribution of Live Births, by Type of Person Providing Assistance During Delivery*

Definition

A. Coverage:

1. Population base: Live births to interviewed women who had one or more births in the five years preceding the survey.

2. Time period: Five years preceding survey.

B. Numerators: Numbers of live births distributed according to type of person providing delivery assistance: doctor, nurse/midwife/auxiliary, traditional birth attendant, relative/other, no one.

C. Denominator: Number of live births in the last five years.

Calculation

For percents, 100 times the quotient of the numerators divided by the same denominator.

Handling of Missing Values

Included in distribution.

Notes and Considerations

Percent distribution adds up to 100 percent.

The category "Nurse/midwife/auxiliary" includes only medically trained and licensed personnel, including auxiliary health personnel. Traditional birth attendants (also sometimes called midwives) are not included, whether trained or untrained.

DELIVERY CARE

Statistics:

A. Percentage of live births delivered by caesarean Section.

B. Percent distribution of live births, by birth weight.

C. Percent distribution of live births, by birth size.

Definition

A. Coverage:

1. Population base: Live births to interviewed women who have had one or more births in the five years preceding the survey.

2. Time period: Five years preceding the survey.

B. Numerators:

1. Number of live births delivered by caesarean section.

2. Numbers of live births distributed according to birth weight and whether weighed at birth.

3. Numbers of live births distributed according to mother's estimate of baby's size at time of birth.

C. Denominator: Number of live births in the last five years.

Calculation

For percentages and percent distribution, 100 times the quotient of the numerators divided by the same denominator.

Handling of Missing Values

Excluded from numerator of percentage delivered by Caesarean section (assumed not Caesarean section).

Included in distributions of birth size but no longer included in distribution of birth weight.

Notes and Considerations

Same denominator used in percentage and percent distribution.

Percent distributions add up to 100 percent.

The category "Not weighed" is no longer included in the distribution of birth weight.

Statistics: *Percent Distribution of Women, by Time Since Delivery at First Postnatal Checkup*

Definition

A. Coverage:

1. Population base: Women who had a live birth outside of a health facility in the five years preceding the survey.

2. Time period: Five years preceding the interview.

B. Numerators: Number of women with a delivery outside of a health facility, distributed by the number of days since delivery at time of first postnatal checkup.

C. Denominator: Number of women with a live birth in the last five years, which was not delivered in a health facility.

Calculation

For percent distribution, 100 times the quotient of the numerators divided by the same denominator.

Handling of Missing Values

Included in distribution.

Notes and Considerations

Women whose last birth was delivered in a health facility are excluded from both numerators and denominators. They are assumed to have received a postnatal checkup soon after delivery.

Percent distribution adds up to 100 percent.

The category "Did not received prenatal checkup" is included in the distribution. This category also includes women whose first postnatal checkup occurred 42 or more days after delivery.

VACCINATION RATES

Statistics: *Percentage of Children Age 12–23 Months Who Received Specified Vaccines at Any Time Before the Survey*

- Percentage of children age 12–23 months who received all specified vaccines at any time before the survey.
- Percentage of children age 12–23 months who received no specified vaccines at any time before the survey.
- Percentage of children age 12–23 months who received specified vaccines by 12 months of age.
- Percentage of children age 12–23 months who received all specified vaccines by 12 months of age.
- Percentage of children age 12–23 months who received no specified vaccines by 12 months of age.
- Percentage of children age 12–23 months who received specified vaccines by 12 months of age.

Definition

A. Coverage:

1. Population base: Living children age 12–23 months.
2. Time period:
   a) Any time before survey.
   b) By age 12 months of child.

B. Numerator: Number of children receiving specified vaccines, all specified vaccines, and no specified vaccine, according to information on vaccination card or report by respondent.

C. Denominator: Number of living children.

Specified Vaccines

1. BCG—Anti-tuberculosis to be received shortly after birth.
2. DPT—Triple vaccine for diphtheria, pertussis and tetanus, received in three doses at 6, 10, and 14 weeks after birth; may also be given as a pentavalent vaccine with Hepatitis B and Hib.
3. Polio—Received in three doses usually given at same time as DPT, plus at birth in some countries.
4. Measles—Recommended to be given at 9 months, except in Pan American Health Organization (PAHO) countries where it is recommended to be given between 12 and 15 months of age.

5. All vaccinations—child was given BCG, three doses of DPT, three doses of polio (excluding. the dose given shortly after birth), and measles.

6. No vaccinations—none of the specified vaccinations was given.

Calculation

A. Numerators

1. Number of living children between 12 and 23 months of age at the time of the survey who received the specified vaccine. Where the information is present on a vaccination card shown to the interviewer, the record of the vaccination is used. Where no card was shown to the interview or there was no record of the vaccination, the mother's report of the vaccination is used. For DPT, where not asked of the mother, if the mother reported a polio vaccination given not shortly after birth, then the equivalent dose of DPT is assumed. Care needs to be taken not to confuse the polio vaccine dose given at birth (called polio 0) from the doses given later (polio 1, polio 2, and polio 3).

2. Number of living children between 12 and 23 months of age who received the specified vaccine by 12 months of age. Where the information is present on a vaccination card shown to the interviewer, the date of the vaccination is used. The difference between the date of the vaccination and the birth date of the child is used to determine whether the vaccination was given between age 0 and 11 months. Where no card is shown to the interview, there is no record of the vaccination, or there is no date of vaccination on the card, the mother's report of the vaccination is used. For DPT, where not asked of the mother, if the mother reported a polio vaccination different from that given shortly after birth, then the equivalent dose of DPT is assumed. Care needs to be taken not to confuse the polio vaccine dose given at birth (called polio 0) from the doses given later (polio 1, polio 2, and polio 3). It is assumed that the proportion of vaccinations reported by the mother has the same distribution by age as that indicated by the vaccination card. Therefore, the number of children with the specific vaccination dose from mother's report is multiplied by the proportion of that dose given before 12 months of age as determined by information taken from the vaccination card.

B. Denominator: Number of living children between 12 and 23 months of age.

Handling of Missing Values

Missing vaccination card: Use mother's report of vaccination, if any.

Missing vaccination on card: Use mother's report of vaccination, if any.

Missing or invalid date on vaccination card: Impute whether before or after 12 months of age according to distributions on vaccination cards with valid dates.

Missing mother's (respondent's) report of vaccination or mother does not know if vaccination given: Treat as vaccine not received when no valid information is on the vaccination card.

Notes and Considerations

Changes: In surveys before those using the DHS IV core questionnaire, questions on DPT vaccination were not asked of the mother. In this case, see the first assumption below.

Assumptions

When a child does not have a vaccination card that was shown to the interviewer, two assumptions are made:

1. The vaccinations for DPT 1, 2, and 3 were given at the same time as those for Polio 1, 2, and 3, respectively.

2. The ages at vaccinations for children whose mothers reported the vaccination are the same as those derived from the dates on the vaccination cards.

Decisions on Alternatives

In some countries, especially those of Latin America, the recommended age for measles vaccination has changed from below age 12 months to age 12 months or above. Where this is the case, the age range of the children in both numerator and denominator needs to be changed from 12–23 months to 18–29 months. Additionally, age 18 months instead of 12 months should be used as the age by which vaccinations should have taken place.

Care must be taken in comparisons over time and between countries that the same age ranges are used; otherwise, underestimation of vaccination rates may occur where vaccines are recommended to be given at older ages than before.

CHILD MORBIDITY PREVALENCE AND TREATMENT: ACUTE RESPIRATORY ILLNESS AND FEVER

Statistics: *Percentage of Children Under Five Years of Age with Symptoms of Acute Respiratory Illness (ARI)*

- Percentage of children under five years of age with fever.
- Percentage of children under five years of age for whom treatment for ARI and/or fever was sought at a health facility or from a health provider.
- Percentages of children under five years of age, by anti-malarial drug taken for fever.

Definition

A. Coverage:

1. Population base: Children under five years of age.
2. Time period: Two weeks preceding the survey.

B. Numerators:

1. Number of children ill with a cough accompanied by short, rapid breathing at any time during the two weeks preceding the interview.
2. Number of children ill with a fever at any time during the two weeks preceding the interview.
3. Number of children under five years of age ill with a cough accompanied by short, rapid breathing and/or with a fever, at any time during the two weeks preceding the interview, for whom treatment was sought at a health facility or from a health provider.
4. Number of children under five years of age ill with a fever, at any time during the two weeks preceding the interview, who took fansidar, who took chloroquine, and who took any anti-malarial drug.

C. Denominators:

1. Number of children under five years of age.
2. Number of children under five years of age.
3. Number of children under five years of age who were ill with ARI or a fever in the two weeks preceding the interview.
4. Number of children under five years of age who were ill with a fever in the two weeks preceding the interview.

Calculation

For percentages, 100 times the quotient of the numerators divided by the denominators.

Handling of Missing Values

Excluded from numerators for (1) and (2) (assumes no illness).

Excluded from both numerator and denominator for (3) (assumes no illness or no treatment sought from a health facility/provider).

Excluded from both numerator and denominator for (4) (assumes no anti-malarial drugs taken or not ill with fever).

Notes and Considerations

Percent distribution adds up to 100 percent.

Pharmacies are not included in health facilities.

CHILD MORBIDITY PREVALENCE AND TREATMENT: DIARRHEAL DISEASE

Statistics: *Percentage of Children Under Five Years of Age with Diarrhea*

- Percentage of children under five years of age with diarrhea taken to a health provider.
- Percentage of children under five years of age with diarrhea given various forms of oral rehydration therapy.
- Percentage of children under five years of age with diarrhea given other treatment and no treatment.
- Percent distribution of children under five years of age with diarrhea, by amount of liquids and amount of foods offered.

Definition

A. Coverage:

1. Population base: Children under five years of age.

2. Time period: Two weeks preceding the interview.

B. Numerators:

1. Number of children ill with diarrhea (as defined by the respondent—child's mother) at any time during the two weeks preceding the interview.

2. Number of children ill with diarrhea at any time during the two weeks preceding the interview who were taken to be seen by a health provider.

3. Number of children ill with diarrhea at any time during the two weeks preceding the interview who received oral rehydration salts (ORS) packets; recommended home fluids (RHF); either ORS or RHF; increased fluids; and ORS, RHF, or increased fluids.

4. Number of children under five years of age ill with diarrhea at any time during the two weeks preceding the interview who received treatment other than ORT: pill/syrup, injection, intravenous solution, and home remedy/other), and number who received no treatment.

5. Number of children under five years of age ill with diarrhea at any time during the two weeks preceding the interview in categories of amount of liquids and foods offered: same as usual, more than usual, somewhat less, much less or none.

C. Denominators:

1. Number of children under five years of age.

2. Number of children under five years of age who were ill with diarrhea in the two weeks preceding the interview.

3. Number of children under five years of age who were ill with diarrhea in the two weeks preceding the interview.

4. Number of children under five years of age who were ill with diarrhea in the two weeks preceding the interview.

5. Number of children under five years of age who were ill with diarrhea in the two weeks preceding the interview.

Calculation

For percentages, 100 times the quotient of the numerators divided by the denominators.

Handling of Missing Values

Excluded from numerators for a. (assumes no illness).

Excluded from both numerator and denominator for b, c, and d (assumes no illness, no treatment).

Included in distributions of amounts of liquids and food for e as separate categories.

Notes and Considerations

Percent distribution adds up to 100 percent.

Pharmacies, shops, and traditional practitioners are not included in health providers.

Amounts of liquids and foods offered are compared to normal practice.

CHILD MORBIDITY PREVALENCE AND TREATMENT: DIARRHEAL DISEASE

Statistics: *Percentage of Mothers Who Know About ORS Packets*

Definition

A. Coverage:

1. Population base: Mothers with births in the five years preceding the survey
2. Time period: Time of survey

B. Numerator: Number of mothers who know about oral rehydration salts (ORS) packets.

C. Denominator: Number of mothers.

Calculation

For percentage, 100 times the quotient of the numerator divided by the denominator.

Handling of Missing Values

Excluded from numerator (assumes does not know).

CHILDHOOD MORBIDITY PREVENTION: MALARIA

Statistics:

A. Percentage of households with bednets.

B. Percent distribution of households with a child under five years of age, by use of bednets.

Definition

A. Coverage:

1. Population bases:

   a) All households.

   b) Households with a child under five years of age who slept in the household the night before the interview.

2. Time period: Night before interview.

B. Numerators:

1. Number of households with mosquito bednets.

2. Number of households, by whether children under five years of age distributed by use of bednets the night before the interview: All children slept under bednets, some children slept under bednets, no child slept under a bednet (including households without bednets).

C. Denominators:

1. Number of households.

2. Number of households where a child under five years of age slept the night before the interview.

Calculation

For percentage and percent distribution, 100 times the quotient of the numerators divided by the denominators.

Handling of Missing Values

Excluded from numerator for percentage with a bednet (assumes does not have).

Included in distribution of use of bednets as a separate category.

Notes and Considerations

CHILDHOOD MORBIDITY PREVENTION: DIARRHEA

Statistics:

A. Percentage of households with hand-washing materials.

B. Percent distribution of mothers with a child under five years of age, by form of disposal of child's stools.

Definition

A. Coverage:

B. Population base:

1. All households.

2. Women with a child under five years of age.

C. Numerators:

1. Numbers of households with hand-washing materials: water/tap, soap/ash/other cleansing agent, basin; all three hand-washing materials; no hand-washing materials.

2. Numbers of women with a child under five years of age distributed by way in which youngest child's fecal matter is disposed: Child always uses toilet/latrine; fecal matter thrown into toilet/latrine, buried in yard, thrown outside dwelling, thrown outside yard, rinsed away, not disposed of; child uses disposable diapers; child uses washable diapers; other.

D. Denominators:

1. Number of households.

2. Number of mothers with a child under five years of age living with her.

Calculation

For percentages and percent distribution, 100 times the quotient of the numerators divided by the denominators.

Handling of Missing Values

Excluded from numerators for household percentages.

Included in distribution of disposal of child's stools as a separate category.

WOMEN'S ACCESS TO HEALTH CARE

Statistics: *Percentages of Women with Specific Problems in Accessing Health Care for Themselves*

Definition

A. Coverage: Population base: All interviewed women.

B. Numerators: Numbers of women with specific problems in accessing health care for themselves, such as:

- Knowing where to go for treatment
- Getting permission to go for treatment
- Getting money for treatment
- Distance to health facility
- Having to take transportation
- Not wanting to go alone
- Concern there may not be a female provider
- Any of the specified problems.

C. Denominator: Number of interviewed women

Calculation

For percentages, 100 times the quotient of the numerators divided by the denominator.

Handling of Missing Values

Excluded from numerators (assumed not a problem).

Notes and Considerations

Researchers should not assume that there is only one possible response for each of the response categories. For example, problem of getting permission to go for treatment may refer to employer, insurance company, or HMO, in addition to husband, father, or other family member.

WOMEN'S USE OF SMOKING TOBACCO

Statistics:

A. Percentage of women who use tobacco.

B. Percent distribution of women who smoke cigarettes, by number of cigarettes in preceding 24 hours.

Definition

A. Coverage:

1. Population bases:

   a) All interviewed women.

   b) Women who smoke cigarettes.

2. Time periods:

   a) Time of survey.

   b) Twenty-four hours preceding survey.

B. Numerators:

1. Number of women who use tobacco: cigarettes, pipe, and other tobacco.

2. Number of women distributed by number of cigarettes smoked in preceding 24 hours.

C. Denominators:

1. Number of interviewed women.

2. Number of women who smoke cigarettes.

Calculation

For percentages and distribution, 100 times the quotient of the numerators divided by the denominator.

Handling of Missing Values

Excluded from numerators of percentages.

Included in distribution as a separate category. Excluded from denominator (assumes does not smoke cigarettes).

Notes and Considerations

Frequency of use of tobacco other than cigarettes is not usually collected.

INFANT FEEDING

INITIAL BREASTFEEDING

Statistics:

A. Percentage of children ever breastfed.

B. Percentage of children who started breastfeeding within one hour and one day of birth.

C. Percentage of children who received a prelacteal feeding.

Definition

A. Coverage:

Population bases:

1. Children born in the five years preceding the survey, surviving and dead.
2. Children born in the five years preceding the survey, surviving and dead, who were ever breastfed.

B. Numerators:

1. Number of children ever breastfed.
2. Numbers of children who started breastfeeding within one hour of birth and within one day of birth.
3. Number of children given something other than breast milk during the first three days of life before they started breastfeeding regularly.

C. Denominators:

1. Number of children born in the five years preceding the survey to interviewed women.
2. Number of children born in the five years preceding who were ever breastfed.

Calculation

Numerators divided by denominator and quotient multiplied by 100 to get percentages.

Handling of Missing Values

Excluded from numerators in all percentages.

Included in denominators.

Notes and Considerations

Percentage of children who started breastfeeding within one day of birth includes children who started within one hour of birth.

INFANT FEEDING

BREASTFEEDING AND COMPLEMENTARY FEEDING

Statistics: *Percent distribution of children exclusively breastfeeding, fully breastfeeding, predominantly breastfeeding, and receiving complementary foods*

Definition

A. Coverage:

Population base:

1. Living children born 0 to 35 months before the survey.

2. Time period: 24 hours before the survey.

B. Numerators:

1. Exclusive breastfeeding: Children who were given nothing but breast milk in the 24 hours preceding the interview.

2. Full breastfeeding: Children who were given either breast milk alone or breast milk plus plain water in the 24 hours preceding the interview.

3. Predominant breastfeeding: Children who were given breast milk, with optionally plain water, water-based liquids (such as soft drinks), and/or juices in the 24 hours before the survey. Children given non-breast milk are excluded.

4. Complementary foods: Children who were breastfed and given solid or semi-solid foods in the 24 hours before the survey. Children may also be given non-breast milk.

C. Denominator: Number of living children.

Calculation

Numerators divided by denominator and quotient multiplied by 100 to get percentages.

Handling of Missing Values

Missing data on breastfeeding: treated as not breastfeeding.

Missing data on foods given: treated as not given.

Notes and Considerations

Assumptions

Children who are not the last-born are assumed not to have been breastfed in the 24 hours preceding the survey. However, all children of multiple births are treated the same.

Decisions on Alternatives

An alternative formulation of exclusive breastfeeding would be to consider only children who have never received anything but breast milk. The required information was collected only in the second phase of the DHS, when respondents were asked about the age of introduction of liquids and foods. These questions were dropped when they were found to produce older ages at introduction of foods than that estimated from current status.

Changes over Time

Surveys based on the DHS phase IV core questionnaire only ask about the feeding status of the youngest living child who lives with the mother. Other children are assumed not to be breastfeeding.

Statistics: *Median and Mean Durations of Any Breastfeeding, Exclusive Breastfeeding and Predominant Breastfeeding*

Definition

A. Coverage:

1. Population base.

2. All births in the time period 0 to 35 months before interview.

B. Time period: Breastfeeding and other liquids and foods given in the 24 hours preceding the survey.

Calculation

A. Utilizes the numerators and denominators calculated for the proportions any breastfeeding, exclusive breastfeeding and predominant breastfeeding by time since birth, as given above. Numerators and denominators are each smoothed by a three-group moving average. For example, the value of the numerators and denominators for births that occurred 4–5 months before the interview is the average of groups 2–3, 4–5, and 6–7 months. Groups < 2 months and 34–35 months are not smoothed. Smoothed proportions are calculated by dividing the smoothed numerator values by the smoothed denominator values of each two-month group.

B. Median

1. Beginning with the lowest time-since-birth group, each group is examined to see whether the proportion breastfeeding, exclusively breastfeeding, or predominantly breastfeeding is less than 0.5. Value of median is determined by linear interpolation of percentage of first group below 0.5 and previous group percentage using the following formula:

$$\textbf{median} = m_{i-1} + (p_{i-1} - 0.5)/(p_{i-1} - p_i) \times (w_i),$$

where p_i is the proportion breastfeeding, exclusively breastfeeding, or predominantly breastfeeding for the first group where the proportion is below 0.5, p_{i-1} is the proportion breastfeeding, exclusively breastfeeding, or predominantly breastfeeding for the preceding group, m_{i-1} is the midpoint value for the preceding group, and w_i is the time width of the group taken as the difference between the midpoint value of the current group and the preceding group.

2. For all groups except the first, the midpoint values are one-half month above the lower limit. This midpoint value comes about because the time since birth is calculated as the difference between birth dates and the date of interview, which are accurate for month only (day of month is not ascertained).

3. For example, if the date of interview were April 2001, the interview could have occurred at any time during the month, from the 1st to the 30th. The same holds true for a birth that occurred in January 2001, at any time between the 1st and the 31st of the month. Thus, the time difference between the date of birth and the date of interview could be between 60 days and 120 days. Assuming a constant distribution by day of month for interviews and for births, the midpoint is 90 days or three months, which is the value of the difference in the century-month codes of the dates. The midpoint value for the group of the difference of two months and three months together is therefore 2.5.

4. The first group is special. The value of the previous group is assumed to be equal to the percentage of children born 0 to 36 months before the survey who were ever breastfed. Moreover, births that occur in the month of interview can only come before the date of interview, rather than on any day of the month of interview. Based on simulations, 0.75 is chosen as the midpoint value for the group < 2 months (0–1). The midpoint of the previous group is 0 (day of interview).

   An example of the median calculation is given below:

   Example:

Group	P_i	Midpoint	Width
< 2	0.90	0.75	0.75
2–3	0.70	2.50	1.75
4–5	0.55	4.50	2.00
6–7	0.43	6.50	2.00

   Median duration = 4.5 + (.55 – .5)/(.55 – .43)*(6.5 – 4.5) = 5.33 months

C. Mean

The mean duration is the accumulation over all groups of the proportions breastfeeding, exclusively breastfeeding, or predominantly breastfeeding multiplied by the width of the time-since-birth group. For the first group, the proportion breastfeeding is taken as the proportion ever breastfed for children born 0 to 35 months before the survey.

$$\textbf{mean duration} = \Sigma\, p_i\, w_i$$

Handling of Missing Values

Birth dates and dates of interview are not allowed to have missing data.

Respondents with missing reports of breastfeeding are considered to be not breastfeeding.

Respondents with missing reports for liquids and foods or who do not know whether the child received the liquids or foods are treated as the liquid or food not given.

Notes and Considerations

Assumptions

Medians and means are based on current status of mothers of the births. The distributions of the proportions of births, by month of birth of the child, are analogous to the l_x column of the synthetic life table. The basic assumption is that there has been no change in the proportions breastfeeding, exclusively breastfeeding, and predominantly breastfeeding over time so that births that proportions that are calculated from births, with small times since birth, are the same as those of births that occurred earlier, when those earlier births, had those durations. For the short 36-month period, this assumption is very likely to be approximately correct. The proportions are assumed to be the proportion ever breastfed at the time of birth and to decrease monotonically with time since birth. The time at which the proportions decline to 0.5 is taken as the median. The mean is taken similarly from the l_x column of the life table, which in this case is the series of proportions by time since birth. Because the sum of proportions equals one, there is no need to divide the sum of the proportions times the width of the interval.

Because of the relatively small number of births in each month before the interview, the months are first grouped and then smoothed to reduce random fluctuations due to sampling variance.

Mean

Truncated Mean: Because of the limitation to births that occurred within the three years preceding the survey, the mean is truncated if there are mothers who are breastfeeding longer than three years after their last birth. It is very unlikely that the proportions exclusively or predominantly breastfeeding are more than negligible after 35 months since birth. Therefore, it is felt that the truncated mean is very close to the full mean for both of these statistics. On the other hand, any breastfeeding can continue for considerably longer than 35 months so that the mean may be biased somewhat downwards. To get an idea of the possible extent of this bias, the proportions of children who are breastfeeding, by age of child in months, should be examined.

The measures are based on all births that occurred within the three years preceding the interview, including last and other births, surviving or not.

Decisions on Alternatives

Current status medians and means are used instead of measures based on the recall of durations because of the severe heaping (digit preference) on multiples of 3 and 6 months in the recall data.

In the current calculation of median and mean durations, a woman can contribute more than once if she had more than one birth in the three years preceding the survey. An alternative calculation would be to base the medians and means on women rather than births. In this alternative, each woman is represented only once, which is equivalent to durations based on only the last birth. Estimates of durations based on last births (open interval durations) are thought to overestimate average durations of breastfeeding, since they almost always exceed those based on recall after other births (closed interval durations).

INFANT FEEDING

FOODS

Statistics:

A. Percentage of breastfeeding children consuming specific foods.

B. Percentage of non-breastfeeding children consuming specific food.

Definition

A. Coverage:

1. Population base: Last living child born in the three years preceding the survey who is living with the mother.

2. Time period: Breastfeeding and food consumption in the 24 hours preceding the interview

3. Specific foods:

   a) Infant formula

   b) Other milk/cheese/yogurt

   c) Other liquids (does not include plain water)

   d) Food made from grains

   e) Fruits/vegetables (includes fruits and vegetables rich in vitamin A)

   f) Food made from roots/tubers

   g) Food made from legumes

   h) Meat/fish/shellfish/poultry/eggs

   i) Food made with oil/fat/butter

   j) Fruits and vegetables rich in vitamin A (include pumpkin, red or yellow yams or squash, carrots, red sweet potatoes, green leafy vegetables, mangos, papayas, and other locally grown fruits and vegetables that are rich in vitamin A)

   k) Any solid or semisolid food.

B. Numerators:

1. Numbers of breastfeeding last children who consumed specific foods.

2. Numbers of non-breastfeeding last children who consumed specific foods.

C. Denominators:

1. Number of breastfeeding last children born in the three years preceding the survey to interviewed women.

2. Number of non-breastfeeding last children born in the three years preceding the survey to interviewed women.

Calculation

Numerators divided by denominator and quotient multiplied by 100 to get percentages.

Handling of Missing Values

Breastfeeding: Missing values and “don’t know” assumed not breastfeeding.

Specific foods: Missing values and “don’t know” excluded from numerators in all percentages, but included in denominators.

Notes and Considerations

Country-specific foods may be added to each category. Percentages are based only on children living with the mother. Children living elsewhere are excluded from both the numerators and the denominators.

INFANT FEEDING

Statistics:

A. Mean number of times specific foods were consumed in the last 24 hours by breastfeeding children.

B. Mean number of times specific foods were consumed in the last 24 hours by non-breastfeeding children.

Definition

A. Coverage:

1. Population base: Last living child born in the three years preceding the survey who is living with the mother.

2. Time period: Breastfeeding and food consumption in the 24 hours preceding the interview

3. Specific foods:

   a) Infant formula

   b) Other milk/cheese/yogurt

   c) Other liquids (does not include plain water)

   d) Food made from grains

   e) Fruits/vegetables (include fruits and vegetables rich in vitamin A)

   f) Food made from roots/tubers

   g) Food made from legumes

   h) Meat/fish/shellfish/poultry/eggs

   i) Food made with oil/fat/butter

   j) Fruits and vegetables rich in vitamin A (includes pumpkin, red or yellow yams or squash, carrots, red sweet potatoes, green leafy vegetables, mangos, papayas, and other locally grown fruits and vegetables that are rich in vitamin A)

   k) Any solid or semisolid food.

B. Numerators:

1. Number of times breastfeeding last children consumed specific foods.

2. Numbers of times non-breastfeeding last children consumed specific foods.

C. Denominators:

1. Number of breastfeeding last children born in the three years preceding the survey to interviewed women.

2. Number of non-breastfeeding last children born in the three years preceding the survey to interviewed women.

Calculation

Numerators divided by denominator to get mean number of times.

Handling of Missing Values

Breastfeeding: Missing values and “don’t know” assumed not breastfeeding.

Specific foods: Missing values and “don’t know” excluded from numerators in all percentages, but included in denominators (assumes 0 time).

Notes and Considerations

Country-specific foods may be added to each category. Percentages are based only on children living with the mother. Children living elsewhere are excluded from both the numerators and the denominators. Children who did not get specific food are included in mean as 0 time.

INFANT FEEDING

Statistics:

A. Mean number of days specific foods were consumed in the last 24 hours by breastfeeding children.

B. Mean number of days specific foods were consumed in the last 24 hours by non-breastfeeding children.

Definition

A. Coverage:

1. Population base: Last living child born in the three years preceding the survey who is living with the mother.

2. Time period: Breastfeeding in the 24 hours preceding the interview and food consumption in the 7 days preceding the interview.

3. Specific foods:

   a) Plain water

   b) Infant formula

   c) Other milk

   d) Fruit juice

   e) Other liquids (does not include plain water)

   f) Food made from grains

   g) Food made from roots/tubers

   h) Fruits and vegetables not rich in vitamin A

   i) Food made from legumes

   j) Cheese/yogurt

   k) Meat/fish/shellfish/poultry/eggs

   l) Food made with oil/fat/butter

   m) Pumpkin/red or yellow yams or squash/carrots/red sweet potatoes

   n) Green leafy vegetables

   o) Mangos/papayas/and other locally grown fruits rich in vitamin A.

B. Numerators:

1. Number of days breastfeeding last children consumed specific foods.

2. Numbers of days non-breastfeeding last children consumed specific foods.

C. Denominators:

1. Number of breastfeeding last children born in the three years preceding the survey to interviewed women.

2. Number of non-breastfeeding last children born in the three years preceding the survey to interviewed women.

Calculation

Numerators divided by denominator to get mean number of days.

Handling of Missing Values

Breastfeeding: Missing values and “don’t know” assumed not breastfeeding.

Specific foods: Missing values and “don’t know” excluded from numerators in all percentages, but included in denominators (assumes 0 day).

Notes and Considerations

Country-specific foods may be added to each category. Percentages are based only on children living with the mother. Children living elsewhere are excluded from both the numerators and the denominators. Children who did not get specific food are included in mean as 0 day. Categories cannot be added to form grouped categories.

MICRONUTRIENTS

IODIZATION OF HOUSEHOLD SALT

Statistics

A. Percent distribution of households, by iodine content of salt.

B. Percentage of households tested for iodized salt.

C. Percentage of households with no salt.

Definition

A. Coverage:

1. Population base: Interviewed households.
2. Time period: Time of interview.

B. Numerators:

1. Number of tested households distributed according to iodine content of household salt.
2. Number of households where salt was tested.
3. Number of households with no salt.

C. Denominators:

1. Number of households where salt was tested for iodine content.
2. Number of households interviewed.
3. Number of households interviewed.

Calculation

Numerators divided by denominator and quotient multiplied by 100 to get percent distribution and percentages.

Handling of Missing Values

Missing values and “don’t know” on iodine content or whether salt tested excluded from both numerators and denominators.

Missing values and “don’t know” excluded from numerator (assumes not tested) but included in denominator.

Missing values and “don’t know” excluded from numerator (assumes has salt) but included in denominator.

MICRONUTRIENT INTAKE AMONG CHILDREN

Statistics:

A. Percent distribution of children under 3 years who consumed fruits and vegetables rich in vitamin A.

B. Percentage of children 6–59 months who received vitamin A supplements

C. Percentage of children under five years who live in household with adequately iodized salt.

Definition

A. Coverage:

1. Population base:

   a) Youngest children under three years of age living with an interviewed mother.

   b) Living children under 6–59 months of age.

   c) Living children under five years of age.

2. Time period:

   a) Seven days preceding interview.

   b) Six months preceding interview.

   c) Time of interview.

B. Numerators:

1. Number of youngest living children under three years who consumed fruits and vegetables rich in vitamin A at any time in the seven days preceding the interview.

2. Number of living children 6–59 months who received vitamin A supplements in the six months preceding the interview.

3. Number of children under five years living in households with adequately iodized salt.

C. Denominators:

1. Number of youngest children under three years of age living with an interviewed mother.

2. Number of living children 6–59 months of age.

3. Number of living children under five years of age.

Calculation

Numerators divided by denominators and quotients multiplied by 100 to get percentages.

Handling of Missing Values

Missing values and “don’t know” excluded from numerator (assumes did not consume).

Missing values and “don’t know” excluded from numerator (assumes did not receive).

Children living in households that were not tested or where information is missing on salt content are excluded from numerator and denominator.

Notes and Considerations

The category “Fruits and vegetables rich in vitamin A” includes pumpkin, red or yellow yams or squash, carrots, red sweet potatoes, green leafy vegetables, mango, papaya, and other locally grown fruits and vegetables that are rich in vitamin A.

Meats and other foods high in vitamin A may also have been consumed by children.

Adequately iodized salt contains 15 parts per million or more of iodine.

MICRONUTRIENT INTAKE AMONG WOMEN

Statistics:

A. Percentage of women with a birth in last five years who received a vitamin A dose postpartum.

B. Percentage of women with a birth in last five years who suffered night blindness during pregnancy, reported and adjusted.

C. Percent distribution of women with a birth in last five years, by number of days they took iron tablets or syrup during pregnancy.

D. Percentage of women with a birth in last five years who live in a household with adequately iodized salt.

Definition

A. Coverage:

1. Population base: Women with a birth in the five years preceding the survey.

2. Time period: Last birth in the five years preceding the survey.

B. Numerators:

Number of women who received a vitamin A supplement dose during the first two months after the birth of their most recent child.

Reported: Number of women who reported suffering night blindness during the pregnancy of their most recent child.

Adjusted: Number of women who reported suffering night blindness during the pregnancy of their most recent child and who did not report difficulty with vision during the day.

Number of women distributed by the number of days they took iron tablets or syrup during the pregnancy of their most recent child.

Number of women living in households with adequately iodized salt.

C. Denominators:

1. Number of women with a birth in the five years preceding the survey.

2. Number of women with a birth in the five years preceding the survey who live in households where salt was tested.

Calculation

A. Numerators divided by denominators and quotients multiplied by 100 to get percentages.

B. Handling of missing values.

C. Missing values and “don’t know” excluded from numerator (assumes did not receive).

D. Missing values and “don’t know” excluded from numerators (assumes did not suffer night blindness).

E. Missing values and “don’t know” excluded from numerators and denominator of distribution.

F. Women living in households that were not tested or for which information is missing on salt content are excluded from numerator and denominator.

Notes and Considerations

Adequately iodized salt contains 15 parts per million or more of iodine.

Statistics: *Percentages of Children between Ages 6 and 59 Months, by Anemia Status*

Definition

A. Coverage:

1. Population base: Living children born 6 to 59 months before the survey.

2. Time period: Current status at time of the survey.

B. Numerators:

1. Any anemia: Number of children whose hemoglobin count is less than 11 grams per deciliter (g/dl).

2. Mild anemia: Number of children whose hemoglobin count is between 10.0 and 10.9 grams per deciliter (g/dl).

3. Moderate anemia: Number of children whose hemoglobin count is between 7.0 and 9.9 grams per deciliter (g/dl).

4. Severe anemia: Number of children whose hemoglobin count is less than 7.0 grams per deciliter (g/dl).

C. Denominator: Number of children between ages 6 and 59 months before the survey.

Calculation

A. At the time of creation of a recode file, an adjustment of the hemoglobin count is made for altitude. Rather than change the cutoff points, the effective hemoglobin count is lowered as altitude increases, since oxygen is less available. The adjustment is made with the following formulas:

$$\mathbf{adjust = -0.032*alt + 0.022*alt^2}$$

$$\mathbf{adjHb = Hb - adjust}$$

where *adjust* is the amount of the adjustment, *alt* is altitude in 1,000 feet (converted from meters by dividing by 1,000 and multiplying by 3.3), *adjHb* is the adjusted hemoglobin level, and *Hb* is the measured hemoglobin level in grams per deciliter. No adjustment is made for altitudes below 1,000 meters.

B. The percentage anemic, by category, is obtained by dividing the numerators by the denominator and multiplying the quotient by 100.

Handling of Missing Values

Children who were not tested and those children whose values were not recorded are excluded from both the denominator and the numerators.

Notes and Considerations

Children less than six months of age are not included because they have higher levels of hemoglobin at birth and just after birth and thus may distort the indication of prevalence of anemia.

People residing at higher altitudes (greater than 1,000 meters (3,300 feet)) have higher Hb and Hct levels than those residing at sea level. This variation is due to the lower oxygen partial pressure at higher altitudes, a reduction in oxygen saturation of blood, and a compensatory increase in red cell production to ensure adequate oxygen supply to the tissues. Thus, higher altitude causes a generalized upward shift of the Hb distributions. This shift may be associated with the underdiagnosis of anemia for residents of higher altitudes when sea-level cutoffs are applied (CDC, unpublished data). Therefore, the proper diagnosis of anemia for those residing at higher altitudes requires an upward adjustment of Hb cutoffs. The values for altitude-specific adjustment of Hb are derived from data collected by the CDC Pediatric Nutrition Surveillance System on children residing at various altitudes in the mountain states. Altitude affects Hb levels throughout pregnancy in a similar way (J.N. Chatfield, unpublished data). Altitude data should always be obtained to adjust where the altitude of the dwelling is more than 1,000 meters. Substitution of cluster altitude is usually accepted instead of altitude for each dwelling in the cluster.

References

Centers for Disease Control and Prevention. 1998. Recommendations to prevent and control iron deficiency in the United States. *Morbidity and Mortality Weekly Report 47* (RR-3): 1–29.

Centers for Disease Control and Prevention. 1989. Current trends CDC criteria for anemia in children and childbearing-aged women. *Morbidity and Mortality Weekly Report 38,* (22): 400–404.

Statistics: *Percentages of Women between Ages 15 and 49 Years, by Anemia Status*

Definition

A. Coverage:

Population base: Women of all marital status age 15–49 years at the time of the survey.

Time period: Current status at the time of the survey.

B. Numerators:

1. Any anemia: Number of not pregnant women whose hemoglobin count is less than 12 grams per deciliter (g/dl) plus number of pregnant women whose count is less than 11 g/dl.

2. Mild anemia: Number of not pregnant women whose hemoglobin count is between 10.0 and 11.9 grams per deciliter (g/dl) plus number of pregnant women whose count is between 10.0 and 10.9 g/dl.

3. Moderate anemia: Number of women pregnant or not whose hemoglobin count is between 7.0 and 9.9 grams per deciliter (g/dl).

4. Severe anemia: Number of women pregnant or not whose hemoglobin count is less than 7.0 grams per deciliter (g/dl).

C. Denominator: Number of women between age 15 and 49 years at time of survey.

Calculation

A. At the time of creation of a recode file, an adjustment of the hemoglobin count is made for altitude. Rather than change the cutoff points, the effective hemoglobin count is lowered as altitude increases, since oxygen is less available. The adjustment is made with the following formulas:

$$\mathbf{adjust = -0.032*alt + 0.022*alt^2}$$

$$\mathbf{adjHb = Hb - adjust}$$

where *adjust* is the amount of the adjustment, *alt* is altitude in 1,000 feet (converted from meters by dividing by 1,000 and multiplying by 3.3), *adjHb* is the adjusted hemoglobin level, and *Hb* is the measured hemoglobin level in grams per deciliter. No adjustment is made for altitudes below 1,000 meters.

B. Similarly, an adjustment is made for women who smoke (if information was collected). The adjustment is to be made in accordance with the following table:

Cigarettes Smoked	Adjust Hb (g/dl) Concentration by
Less than 10 per day	No adjustment
10–19 per day	–0.3
20–39 per day	–0.5
40 or more per day	–0.7
Unknown quantity or non-cigarettes smoking	–0.3

C. The percentage anemic, by category, is obtained by dividing the numerators by the denominator and multiplying the quotient by 100.

Handling of Missing Values

Women who were not tested and women whose values were not recorded are excluded from both the denominator and the numerators.

Notes and Considerations

Pregnancy causes an increase in blood fluid diluting somewhat the concentration of hemoglobin. Therefore the cutoff for mild anemia is lowered for pregnant women. No adjustment is made for the cutoff values for either moderate or severe anemia. Because the availability of oxygen is less at higher altitudes (lower partial pressures) and because smoking interferes with the blood's capacity to transfer oxygen, adjustments need to be made for both altitude and smoking.

People residing at higher altitudes (greater than 1,000 meters (3,300 feet)) have higher Hb and Hct levels than those residing at sea level. This variation is due to the lower oxygen partial pressure at higher altitudes, a reduction in oxygen saturation of blood, and a compensatory increase in red cell production to ensure adequate oxygen supply to the tissues. Thus, higher altitude causes a generalized upward shift of the Hb distributions. This shift may be associated with the underdiagnosis of anemia for residents of higher altitudes when sea-level cutoffs are applied (CDC, unpublished data). Therefore, the proper diagnosis of anemia for those residing at higher altitudes requires an upward adjustment of Hb cutoffs. The values for altitude-specific adjustment of Hb are derived from data collected by the CDC Pediatric Nutrition Surveillance System on children residing at various altitudes in the mountain states. Altitude affects Hb levels throughout pregnancy in a similar way (Chatfield, unpublished data).

The influence of cigarette smoking is similar to that of altitude; smoking increases Hb level substantially. The higher Hb of smokers is a consequence of an increased carboxyhemoglobin from inhaling carbon monoxide while smoking. Because carboxyhemoglobin has no oxygen-carrying capacity, its presence causes a generalized upward shift of the Hb distribution curves (CDC, unpublished data). Therefore, a smoking-specific adjustment to the anemia cutoff is necessary for the proper diagnosis of anemia in smokers.

The adjustment for altitude can be substantial, but the adjustment for smoking is less substantial. Thus, where no smoking information is available, no adjustment is made. However, altitude data should always be obtained to adjust where the altitude of the dwelling is more than 1,000 meters. Substitution of cluster altitude is usually accepted instead of altitude for each dwelling in the cluster.

References

Centers for Disease Control and Prevention. 1998. Recommendations to prevent and control iron deficiency in the United States. *Morbidity and Mortality Weekly Report 47* (RR-3): 1–29.

Centers for Disease Control and Prevention. 1989. Current trends CDC criteria for anemia in children and childbearing-aged women. *Morbidity and Mortality Weekly Report 38,* (22): 400–404.

CHILDREN'S AND WOMEN'S NUTRITIONAL STATUS

CHILDREN'S NUTRITIONAL STATUS

Statistics: *Percentages of Children Under 5 Years of Age, by Stunted, Wasted, and Underweight*

Definition

A. Coverage:

1. Population base: Living children born 0 to 59 months before the survey.

2. Time period: Current status at the time of the survey.

B. Numerators:

1. Severely stunted: Number of children whose height for age z-score is less than -3.0 standard deviations (SD) below the mean on the NCHS/CDC/WHO international references standard.

2. Moderately stunted: Number of children whose height for age z-score is between –2.0 and -2.99 standard deviations (SD) below the mean on the NCHS/CDC/WHO international references standard.

3. Severely wasted: Number of children whose weight for height z-score is less than –3.0 standard deviations (SD) below the mean on the NCHS/CDC/WHO international references standard.

4. Moderately wasted: Number of children whose weight for height z-score is between –2.0 and –2.99 standard deviations (SD) below the mean on the NCHS/CDC/WHO international references standard.

5. Severely underweight: Number of children whose weight for age z-score is less than –2.0 standard deviations (SD) below the mean on the NCHS/CDC/WHO international references standard.

6. Moderately underweight: Number of children whose weight for age z-score is between –2.0 and –2.99 standard deviations (SD) below the mean on the NCHS/CDC/WHO international references standard.

C. Denominator: Number of living children between ages 0 and 59 months before the survey.

Calculation

A. The assignment of anthropmetric z-scores based on the NCHS/CDC/WHO International Reference Standard is done through a complicated interpolation function that takes into account sex, age (measured by difference in date of birth and date of interview, both precise to day of month), height in centimeters, and weight in kilograms (precise to 100 grams). As part of a creation of a recode file variables with the z-scores are calculated and included in that file. In the process of assigning the z-scores, checks are made on their plausibility. Z-scores are assigned missing to children with incomplete date of birth (month or year missing or "don't know"). The reason is because z-scores are very sensitive to changes in age. Children with height for age z-scores below –6 SD or above +6 SD, with weight for age z-scores below –6 SD or above +6 SD, or with weight for height z-scores below –4 SD or above +6 SD are flagged as having invalid data. Also invalid are combinations of z-scores where height for age is less than –3.09 SD and weight for age is more than +3.09 SD, or where height for age is more than +3.09 SD and weight for age is less than –3.09 SD.

B. The percentages of children stunted, wasted, and underweight are equal to the specific numerators divided by the denominators and multiplied by 100.

Handling of Missing Values

Children who were not weighed and measured and children whose values for weight and height were not recorded are excluded from both the denominator and the numerators. Children whose day of month of birth is missing or unknown are assigned day 15. Children who are flagged for out-of-range z-scores or invalid z-scores are excluded from both the denominator and the numerators. Children with missing data in the z-scores (date of birth incomplete or missing) are excluded from both the denominator and the numerators.

Notes and Considerations

Stunting, based on a child's height and age, is a measure of chronic nutritional deficiency. Wasting, based on a child's weight and height, is a measure of acute nutritional deficiency. Underweight, based on weight and age, is a composite measure of both acute and chronic statuses. The NCHS/CDC/WHO standard is based on well-nourished children in the United States. Due to natural variations in a well nourished population, 2.2 percent of children will be between –2.0 and –2.99 SD below the mean, and 0.1 percent will be –3.0 or more SD below the mean. The extent of malnutrition in a population should be taken by the extent the proportions moderate and severe exceed these percentages that occur in a well-fed population of children.

The NCHS/CDC/WHO international reference standard is a combination of two reference standards: one for children under 24 months and the other for children 24–216 months. The first standard, called the Fels standard, is based on children whose height was taken while prone. The other standard, based on NHANES data, is derived from children 24 months and older whose height was measured while they were standing. For children 24 months of age or over, whose height was measured while lying down, one centimeter should be subtracted from their measured height before calculating the z-scores.

Changes over Time

In phases of the DHS survey before phase IV (*DHS+*), only children of interviewed women and who were under five years old (or the cutoff for the health section of the individual questionnaire) were weighed and measured. In many surveys, only a subsample of these children were selected for anthropometry. All comparisons between surveys, either over time or between countries, should take into account the possible differences in the defined population base.

References

National Center for Health Statistics. 1977. Growth curves for children birth–18 years. *Vital Health Statistics, 11:* 1–165.

United Nations Department of Technical Co-operation for Development and Statistical Office. 1986. *How to weigh and measure children: Assessing the nutritional status of young children in household surveys*. United Nations, New York.

Centers for Disease Control and Prevention. 1999. ANTHRO: Software for calculating pediatric anthropometry, version 1.02. Software and documentation available at http://www.cdc.gov/nccdphp/dnpa/anthro.htm.

Statistics: *Percentages of Women Age 15–49 by Nutritional Status*

Definition

A. Coverage:

1. Population base: Non-pregnant and non-postpartum women age 15–49 years at the time of the survey.

2. Time period: Current status at time of the survey.

B. Numerators: Number of women with a body mass index (BMI) with the following values:

1. Normal: 18.5 to 24.9

2. Mildly thin: 17.0 to 18.4

3. Moderately thin: 16.0 to 16.9

4. Severely thin: less than 16.0

5. Overweight: 25.0 to 29.9

6. Obese: 30.0 or more.

C. Denominator: Number of women age 15–49 years at the time of the survey, excluding women who are pregnant or who gave birth in the two months preceding the date of the interview.

Calculation

A. The body mass index is calculated as weight in kilograms divided by the square of height in meters.

B. The percentages of women, by category, are equal to the category numerators divided by the denominator and multiplied by 100.

Handling of Missing Values

Women who were not weighed and measured and women whose values for weight and height were not recorded are excluded from both the denominator and the numerators. Women whose calculated BMI is below 12.0 or above 60.0 are flagged as out of range and are excluded from both the denominator and the numerators.

Notes and Considerations

BMI, also known as the Quetelet Index, is a measure of acute nutritional status. It is based on the Fogarty Metropolitan Life tables of ideal weight for height.

Changes Over Time

In phases of the DHS survey before phase IV (*DHS+*), either only interviewed women or mothers of children under five years (or the cutoff for the health section of the individual questionnaire) were weighed and measured. In some surveys, only a subsample of these women were selected for anthropometry. All comparisons between surveys, either over time or between countries, should take into account the possible differences in the defined population base. While pregnant women and women two months postpartum are weighed and measured, they are excluded from the report tabulations because of weight gain during pregnancy.

References

Nestel, P., and S. Rutstein 2002. Defining nutritional status of women in developing countries. *Public Health Nutrition, 5* (1): 17–27.

Metropolitan Life Insurance Company. 1942. Ideal weights for women. *New York Statistical Bulletin, 23* (6).

James, W. P., T. A. Ferro-Luzzi, and J. C. Waterlow, 1988. Definition of chronic energy deficiency in adults: Report of a working party of the International Dietary Energy Consultative Group. *European Journal of Clinical Nutrition, 42:* 969–981.

HIV/AIDS/STDs

HIV/AIDS PREVENTION KNOWLEDGE

Statistics: *Percent Distribution of Women, by Knowledge of Programmatically Ways to Avoid HIV/AIDS, and Percentage Who Know Two Specific Ways*

Definition

A. Coverage: All women.

B. Numerators:

1. Numbers of women distributed according to knowledge of number of ways to avoid HIV/AIDS.

2. Number of women who declared that using condoms prevent them from getting HIV/AIDS.

3. Number of women who declared that limiting the number of sexual partners prevent from getting HIV/AIDS.

C. Denominator: Number of women in the sample.

Calculation

A. The programmatically important ways to determine the knowledge of number of ways to avoid HIV/AIDS are: abstaining from sex, using condoms, and limiting the number of sexual partners. Abstinence from sex is measured from spontaneous response only. Using condoms and limiting the number of sexual partners is measured from spontaneous and probed responses.

B. Women who have not heard of HIV/AIDS or that do not know of the three programmatically ways to avoid HIV/AIDS are classified are knowing no way to avoid HIV/AIDS.

Handling of Missing Values

Women with missing values for knowledge of HIV/AIDS are included in separate categories in the distribution of number of ways. Women with missing values in the knowledge for ways to avoid AIDS are considered as not knowing any source.

Notes and Considerations

The percent distribution adds up to 100 percent.

KNOWLEDGE OF STD SYMPTOMS

Statistics: *Percentage of Women with Knowledge of Symptoms Associated with Sexually Transmitted Diseases (STDs)*

Definition

A. Coverage: All women.

B. Numerators:

1. Number of women who do not know about infections that can be transmitted through sexual contact.

2. Number of women who know no symptoms, one symptom, and two or more STDs symptoms for men.

3. Number of women who know no symptoms, one symptom, and two or more STDs symptoms for women.

C. Denominator: All women in the sample.

Calculation

To determine the number of symptoms, it is necessary to add up all responses with code 1 in the corresponding (man/woman) multiple-response variable.

Handling of Missing Values

Women with missing values on knowledge of STDs will be classified as missing in both the man and woman knowledge of STDs. Women with missing values in the specific STDs symptoms for men or women are considered as if they have not mentioned any symptoms.

Notes and Considerations

There may be some country-specific symptoms listed as STDs symptoms when, in fact, they are not. They should not be included when determining the number of symptoms. For example, the other response is not included because there is no way to know what was included in that category.

CONDOM USE AT LAST SEX

Statistics: *Percentage of Women Who Used Condom during the Last Sexual Intercourse*

Definition

A. Coverage: Women who had sexual intercourse in the past 12 months.

B. Numerators:

1. Number of women who had sexual intercourse in the last 12 months with spouse or cohabiting partner, and during that episode they used condoms.

2. Number of women who had sexual intercourse in the last 12 months with other partners than spouses or cohabiting partners, and during that episode they used condoms.

3. Number of women who had sexual intercourse in the last 12 months, and during that episode they used condoms.

C. Denominator:

1. Number of women who had sexual intercourse in the last 12 months with spouse or cohabiting partner.

2. Number of women who had sexual intercourse in the last 12 months with other partners than spouses or cohabiting partners.

3. Number of women who had sexual intercourse in the last 12 months, and during that episode they used condoms.

Calculation

Once the numerators and denominators are properly established, the proportion is reduced to the division of the numerators by the corresponding denominators and multiplying them by 100.

Handling of Missing Values

Women with missing values on time since last sexual intercourse are excluded from both numerators and denominators. Women with missing values in variables related to relationship to sexual partner are considered as having sex with other partners than husbands or cohabiting partners. Women with missing values in use of condoms are considered as not having used the condom for that sexual episode.

Notes and Considerations

The DHS core questionnaire collects information for the last two episodes of sexual intercourse in the last 12 months. When the AIDS module is used, it collects information for the last three episodes. Special care needs to be taken when matching the relationship with the sexual partner and the use of condoms. This becomes even more critical with the AIDS module where three episodes of sexual intercourse have to be analyzed. For example, the respondent could have had sex with a non-cohabiting partner and used condoms during that episode, but that sexual episode was not the last one the respondent had with a non-cohabiting partner.

Statistics: *Age-specific Adult Mortality Rates; Age 15-49 Years Mortality Rate*

Definition

D. Coverage: Women and men age 15-49 years.

E. Numerators:

1. Number of female siblings of respondents who died in the period 0-6, 7-13, and 0-13 years prior to the interview by five-year age group at time of death.

2. Number of male siblings of respondents who died in the period 0-6, 7-13, and 0-13 years prior to the interview by five-year age group at time of death.

F. Denominator:

1. Number of years of exposure of female siblings of respondents during the period 0-6, 7-13, and 0-13 years prior to the survey by five-year age group.

2. Number of years of exposure of female siblings of respondents during the period 0-6, 7-13, and 0-13 years prior to the survey by five-year age group.

C. **Numerator:** Deaths are tabulated according to period of death and the age of sibling at the time of the death.

3. **Period of death:** The period of death is calculated as the difference in months between the date of interview and the date of death of the sibling, both in century-month code format (CMC). Deaths of siblings are included in the tabulation if they occurred within the period tabulated.

4. **Age of sibling at the time of the death:** The difference in months between the date of death of the sibling and the date of birth of the sibling both in CMC. The difference is then divided by 60 and truncated to whole numbers to form the age groups. Deaths are tabulated by age group.

D. **Denominator:** Person-years of exposure are calculated as the sum of the number of months exposed in the five-year age group during the time period divided by 12. A sibling can contribute exposure to two or three five-year age groups during an 84-month period and three or four age groups during the 168 month period.

E. **Calculations:**

a. Numerator (deaths):

For each dead sibling, age at death is determined directly from the response by classification into 5 year age groups. Period of death is determined by using the sibling's date of death. Deaths age ages less than 15 years or more than 49 years are not tabulated as are deaths occurring earlier or later than the period.

b. Denominator (exposure): [Example for period 0-6 years]

For surviving siblings:

1. Highest age group: A surviving sibling's age at the end of the period determines the highest age group. The highest age group is calculated by subtracting the sibling's date of birth from the date of interview minus one (in CMC), dividing the difference by 60 and truncating to a whole number. The number of months spent in the highest age group is the difference in months between the age at the end of the period of exposure (date of interview less one month) and the lower age limit of the age group plus one month. If the number of months in the age group is less than 60 months, then the sibling contributes exposure to both the highest age group and the middle age group.

2. Middle age group: A surviving sibling's age at the end of the period less 60 months determines the middle age group. The middle age group is calculated by subtracting the sibling's date of birth from the date of interview minus sixty-one (in CMC), dividing the difference by 60 and truncating to a whole number. The number of months spent in the middle age group is the difference in months between the age at the end of the period of exposure less 60 months (date of interview minus sixty-one months) and the lower age limit of the age group plus one month. If the sum of number of months exposed in both the highest and middle age groups is less than 84, then the sibling contributes exposure to both the middle age group and the lowest age group.

3. Lowest age group: The contribution to the lowest age group is 84 less the number of months in the sum of the highest and middle age groups. If the number of months in the sum of the highest and middle age groups is greater than or equal to the duration of the time period (i.e., ≥ 84 months), then the exposure in the lowest age group is zero.

For dead siblings:

1. Highest age group: The above procedure for the highest age group is followed using the date of death (in CMC) instead of the date of interview minus one month (in CMC). If the date of death is less than the date of interview minus sixty months, no exposure is contributed to the highest age group.

2. Middle age group: The above procedure middle age group is followed using the lower of the date of death (in CMC) and the date of interview minus sixty-one months (in CMC). If the date of death is less than the date of interview minus 120 months, no exposure is contributed to the middle age group.

3. Lowest age group: The above procedure middle age group is followed using the lower of the date of death (in CMC) and the date of interview minus 121 months (in CMC). If the date of death is less than the date of interview minus 84 months, no exposure is contributed to the lowest age group.

c. **For other periods (7-13 years and 0-13 years prior to survey):** The limits of the periods and the durations of the periods are adjusted accordingly. For the 0-13 year period, there are two middle age groups.

d. **Tabulation:** Each sibling is tabulated three times, once according to her/his highest age group accumulating the exposure she/he contributes to that group, once in the middle age group accumulating middle age group exposure, and once in the lowest age group accumulating lowest age group exposure. (In CSPRO and ISSA, the same table is used, effectively summing the accumulations within each age group.)

Once the numerators and denominators are properly established, age-specific mortality rates are obtained by the division of the numerators by the corresponding denominators and multiplying by 1000. The general mortality rate for age 15-49 is obtained by multiplying the age-specific mortality rates by the proportion of respondents in the five-year age group and then summing the age-distribution-adjusted mortality rates.

Handling of Missing Values

Siblings whose survival status is unknown to the respondent are excluded from both numerators and denominators. Dates of birth are calculated from current age for living siblings and age at death and number of years ago sibling died for dead siblings. Dates of birth and death are imputed where age, age at death, or number of years ago is missing or unknown, taking into consideration of sibling's birth order and number of children (information collected for dead siblings only).

Notes and Considerations

The DHS maternal mortality module questionnaire collects information from respondents (female and male about all of their siblings born to the same mother, starting with the oldest. For living siblings, the date of birth in the recode data file is calculated by subtracting the age from the date of interview. For dead siblings, the date of birth is calculated by subtracting the sum of the responses on age at death and the number of years ago the death occurred from the date of interview. To calculate the date of death, the number of years ago the death occurred is subtracted from the date of interview. After these calculations, month of date of birth and date of death are assigned by subtracting a random number between 0 and 11 and making sure that the birth order and minimum birth intervals are maintained.

However, the age distribution of siblings is very much different than the age distribution of the population since only eligible women and men can report on their siblings. For example if a girl or boy is sixteen years old and has only siblings younger than 15 years, she or he will not be reported. The same holds true at the upper end of the eligible age range. Thus the age distribution of siblings is a curve with minimums at the ends of the eligible age range of respondents and a maximum at about the midpoint of the eligible age range (30 to 35 years). In order to properly calculate general or total rates, age-specific rates must be adjusted to a more representative age distribution. The distribution of respondents is used for this adjustment.

One might think that the calculation of mortality rates is biased because the (living) respondent is not included. Similarly, people with no siblings are not included since there is no one to report on them. However, it has been shown by German Rodriguez that these two potential biases cancel each other out, under the assumption that mortality rates are unrelated to the size of the sibship.

Another important issue is location. The DHS does not collect information on the residence of neither siblings who died nor of the residence during the exposure period of both living and dead siblings. The residence at the time of interview of respondents is not necessarily the same as that of their siblings. Therefore, DHS usually does not publish adult mortality rates by area.

ADULT MORTALITY (DIRECT METHOD)

Statistics: *Age-specific Maternal Mortality Rates; Total Maternal Mortality Rate; Maternal Mortality Ratio (Direct Method)*

Definition

A. Coverage: Women age 15-49 years.

B. Numerator:

Number of female siblings of respondents who died during pregnancy, delivery or within two months of delivery in the period 0-6, 7-13, and 0-13 years prior to the interview by five-year age group at time of death. Maternal deaths of sisters 50 years and older are added to the number of maternal deaths of sisters 45-49 years.

G. Denominator:

Number of years of exposure of female siblings of respondents during the period 0-6, 7-13, and 0-13 years prior to the survey by five-year age group.

H. **Calculation of Maternal Mortality Rates:**

e. **Numerator:** Deaths are tabulated in the same way as for adult age-specific mortality rates and then multiplied by the proportion of deaths that occurred during pregnancy, delivery or within two months of delivery.

f. **Denominator:** Women-years of exposure are calculated in the same way as for adult age-specific mortality.

Once the numerators and denominators are properly established, age-specific maternal mortality rates are obtained by the division of the numerators by the corresponding denominators and multiplying by 100,000. The total maternal mortality rate (for age 15-49) is obtained by multiplying the age-specific mortality rates by the proportion of respondents in the five-year age group and then summing the age-distribution-adjusted maternal mortality rates. The total proportion of deaths maternal is calculated by dividing the total maternal mortality rate by the general 15-49 adult mortality rate of women.

I. **Calculation of the Maternal Mortality Ratio:**

The maternal mortality ratio is calculated by dividing the total maternal mortality rate by the general fertility rate (see page 32) for the period and is expressed per 100,000 births by multiplying the product by 1000.

Handling of Missing Values

Siblings whose maternal status at the time of death are unknown to the respondent or are missing in the data set are allocated to maternal deaths in the proportion they are to deaths of any maternal status. For each age group of siblings, the number of deaths of known maternal status is divided by the number of deaths of any status to get the proportion of maternal deaths. This proportion is then multiplied by the number of deaths of unknown status to calculate the number of deaths to add to the known maternal deaths to get the total number of maternal deaths in each age group.

Notes and Considerations

The DHS maternal mortality module questionnaire collects information from respondents (female and male) about the maternal status of the death of their sisters born to the same mother by asking if the sister died while pregnant, during delivery or within two months after the end of a pregnancy or a childbirth. Younger (and male) respondents may not know that their older sister was even pregnant if the sister was several years older or died during pregnancy or from an induced abortion, thus biasing maternal mortality rates downward. Women who intend to have an induced abortion may also not disclose their pregnancy status to family members. On the other hand, deaths due to non-maternal causes, such as accidents and illnesses, will be included as maternal deaths if they occurred during pregnancy or within two months after the end of the pregnancy or childbirth. Simulation models show that up to one-third of classified as maternal may not be due to maternal causes, resulting in an upward bias. The final result of both these biases, which operate simultaneously, is unknown.

Another important issue is location. The DHS does not collect information on the residence of neither sisters who died nor of the residence during the exposure period of both living and dead sisters. The residence at the time of interview of respondents is not necessarily the same as that of their sisters. Therefore, DHS usually does not publish maternal mortality rates and maternal mortality ratios by area.

Maternal mortality rates and ratios are subject to high levels of relative sampling error due to their relatively rare occurrence. For example, a maternal mortality ratio of 500 maternal deaths per 100,000 births has the same sampling error as an infant mortality rate of 5 infant deaths per 1000 births. For a sample of about 15,000 respondents, the 95% confidence interval of the MMR would be about 406 to 594 maternal deaths per 100,000 births.

References

Rutenberg, N. and Sullivan, J.M. *Direct and indirect estimates of maternal mortality from the sisterhood method*. IRD/Macro International Inc., Washington DC, 1991.

Hanley, J.A., Hagen, C.A. and Shiferaw, T. Confidence intervals and sample size calculations for the sisterhood method of estimating maternal mortality. *Studies in Family Planning* 27(4) July/August 1996.

WHO/UNICEF. *The Sisterhood method to estimate maternal mortality*. Report of a technical meeting, 5-6 December 1996.

Stanton, C., Abderrahim, N. and Hill, K. *DHS maternal mortality indicators: An assessment of data quality and implications for data use.* Demographic and Health Surveys Analytical Report No 4, Macro International Inc. Calverton, Maryland, USA, September 1997.

ADULT MORTALITY (INDIRECT METHOD)

Statistics: *Total Maternal Mortality Rate; Maternal Mortality Ratio (Indirect Method)*

Definition

J. Coverage: Women age 15-49 years.

K. Numerator:

1. Number of female siblings of respondents who died during pregnancy, delivery or within two months of delivery by five-year age group of respondent.

L. Denominator:

1. Number of years of sister-units of exposure of female siblings of respondents by five-year age group of respondent.

M. **Calculation of Maternal Mortality Rates:**

g. **Numerator:** Deaths that occurred during pregnancy, delivery or within two months of delivery are tabulated.

h. **Denominator:**

i. For respondents age 30 to 49 the numbers of sisters who attained fifteen or more years of age are tabulated by the five-year age-group of the respondents. For respondents 15 to 29 the numbers of sisters for age groups 15-19, 20-24 and 25-29 are calculated by multiplying the number of respondents in each five-year age group by the average number of sisters who attained age 15 or more for the total of respondents age 30 to 49 years.

ii. Sister-units of exposure are calculated by applying an adjustment factor to the number of sisters who attained fifteen or more years of age tabulated by five-year age groups of respondents. The adjustment factors are 0.107, 0.206, 0.343, 0.503, 0.664, 0.882, and 0.900 for the five year age groups from 15-19 to 45-49.

Once the numerators and denominators are properly established, the life-time risks of maternal mortality by five-year age group are obtained by the division of the numerators by the corresponding denominators. The total life-time risk of maternal mortality (for age 15-49) is obtained by the quotient of the sums of the age-group specific numerators and denominators.

N. **Calculation of the Maternal Mortality Ratio:**

The maternal mortality ratio (MMR) is calculated by taking the root (equal to the total fertility rate--TFR for the period 10-14 years prior to the survey) of (one minus the total life-time risk of maternal mortality) and subtracting the result from one. This calculation gives the average per birth risk of maternal mortality. Multiplying by 100,000 gives the maternal mortality ratio. The formula is expressed as:

$$MMR = 100{,}000 \times (1 - [1 - \text{total lifetime risk}]^{(1/TFR)})$$

Handling of Missing Values

Siblings whose maternal status at the time of death are unknown to the respondent or are missing in the data set are allocated as maternal deaths in the proportion they have to all deaths (e.g., if unknown status deaths are 10% of all deaths in an age group of respondent, then the number of maternal deaths in that age group is increased by 10%).

Siblings whose maternal status at the time of death is unknown to the respondent or is missing in the data set are allocated to maternal deaths in the proportion they are to deaths of any maternal status. For each age group of respondents, the number of deaths of known maternal status is divided by the number of deaths of any status to get the proportion of maternal deaths. This proportion is then multiplied by the number of deaths of unknown status to calculate the number of deaths to add to the known maternal deaths to get the total number of maternal deaths in each respondent age group.

Notes and Considerations

The indirect method of calculating maternal mortality lifetime risk and the maternal mortality ratio is based on applying models of fertility and maternal mortality rates. The number of sisters of respondents is converted into "sister units of risk" equivalent to women-years of exposure to the risk of mortality in the reproductive age (i.e. since age 15 years) by multiplication with adjustment factors calculated on a model age-fertility distribution. The conversion from the lifetime risk of maternal mortality to the maternal mortality ratio assumes a model pattern of maternal mortality by age and by parity.

> "Because the method relies heavily on a number of assumptions about the relationships between fertility and age-specific maternal mortality, it should not be used in settings where levels of fertility are low (Total Fertility Rate below 3), or where there have been recent and marked declines in fertility, or where major migration has occurred. While the method is relatively simple and inexpensive to use, the overall results relate to a point around 10-12 years prior to the survey, a major disadvantage. However, it is possible to calculate estimates for more recent periods by limiting the upper age of the respondents to, say, adults aged below 30 years old. In this case, the overall maternal mortality estimate would relate to a period some seven years prior to the survey. The disadvantage of limiting the upper age of respondents is, however, that a large number of households need to be visited to achieve the desired sample size of adult respondents". [World Health Organisation, 1997. **The sisterhood method for estimating maternal mortality: Guidance for potential users.** WHO/RHT/97.28].

The DHS maternal mortality module questionnaire collects information from respondents (female and male) about the maternal status of the death of their sisters born to the same mother by asking if the sister died while pregnant, during delivery or within two months after the end of a pregnancy or a childbirth. Younger (and male) respondents may not know that their older sister was even pregnant if the sister was several years older or died during pregnancy or from an induced abortion, thus biasing maternal mortality rates downward. Women who intend to have an induced abortion may also not disclose their pregnancy status to family members. On the other hand, deaths due to non-maternal causes, such as accidents and illnesses, will be included as maternal deaths if they occurred during pregnancy or within two months after the end of the pregnancy or childbirth. Simulation models show that up to one-third of classified as maternal may not be due to maternal causes, resulting in an upward bias. The final result of both these biases, which operate simultaneously, is unknown.

Another important issue is location. The DHS does not collect information on the residence of neither sisters who died nor of the residence during the exposure period of both living and dead sisters. The residence at the time of interview of respondents is not necessarily the same as that of their sisters. Therefore, DHS usually does not publish maternal mortality rates and maternal mortality ratios by area.

Maternal mortality rates and ratios are subject to high levels of relative sampling error due to their relatively rare occurrence. For example, a maternal mortality ratio of 500 maternal deaths per 100,000 births has the same sampling error as an infant mortality rate of 5 infant deaths per 1000 births. For a sample of about 15,000 respondents, the 95% confidence interval of the MMR would be about 406 to 594 maternal deaths per 100,000 births.

References

Graham, W., Brass, W. and Snow, R.W. Indirect estimation of maternal mortality: the sisterhood method. *Studies in Family Planning*, 1989, 20 (3):125-135.

Rutenberg, N. and Sullivan, J.M. *Direct and indirect estimates of maternal mortality from the sisterhood method*. IRD/Macro International Inc., Washington DC, 1991.

Hanley, J.A., Hagen, C.A. and Shiferaw, T. Confidence intervals and sample size calculations for the sisterhood method of estimating maternal mortality. *Studies in Family Planning* 27(4) July/August 1996.

WHO/UNICEF. *The Sisterhood method to estimate maternal mortality*. Report of a technical meeting, 5-6 December 1996.

Stanton, C., Abderrahim, N. and Hill, K. *DHS maternal mortality indicators: An assessment of data quality and implications for data use.* Demographic and Health Surveys Analytical Report No 4, Macro International Inc. Calverton, Maryland, USA, September 1997.

Trussel J. and Rodriguez G. (1990) A Note On the Sisterhood Estimation of Maternal Mortality. Studies in *Family Planning*, 21, 6: 344-346.

DHS
TOOLKIT