USAID
FROM THE AMERICAN PEOPLE

AN ASSESSMENT OF THE QUALITY OF DATA ON HEALTH AND NUTRITION IN THE DHS SURVEYS, 1993-2003

DHS METHODOLOGICAL REPORTS 6

DECEMBER 2008

This publication was produced for review by the United States Agency for International Development. It was prepared by Thomas W. Pullum of The University of Texas at Austin.

GPO

MEASURE DHS assists countries worldwide in the collection and use of data to monitor and evaluate population, health, and nutrition programs. Additional information about the MEASURE DHS project can be obtained by contacting Macro International Inc., Demographic and Health Research Division, 11785 Beltsville Drive, Suite 300, Calverton, MD 20705 (telephone: 301-572-0200; fax: 301-572-0999; e-mail: reports@macrointernational.com; internet: www.measuredhs.com).

The main objectives of the MEASURE DHS project are:

- to provide decisionmakers in survey countries with information useful for informed policy choices;
- to expand the international population and health database;
- to advance survey methodology; and
- to develop in participating countries the skills and resources necessary to conduct high-quality demographic and health surveys.

DHS Methodological Reports No. 6

An Assessment of the Quality of Data on Health and Nutrition in the DHS Surveys, 1993-2003

Thomas W. Pullum

Macro International Inc.
Calverton, Maryland, USA

December 2008

Address for correspondence: Thomas W. Pullum, Department of Sociology, University of Texas at Austin, 1 University Station - A 1700, Austin, TX 78712, USA. Phone: 512-232-6314. Email: tom.pullum@mail.utexas.edu.

Editor: Sidney Moore
Document Production: Kaye Mitchell and Dana Thomson

This study was carried out with support provided by the United States Agency for International Development (USAID) through the MEASURE DHS project (#GPO-C-00-03-00002-00). The views expressed are those of the authors and do not necessarily reflect the views of USAID or the United States Government.

Recommended citation:

Pullum, Thomas W. 2008. *An Assessment of the Quality of Data on Health and Nutrition in the DHS Surveys, 1993-2003.* Methodological Reports No. 6. Calverton, Maryland, USA: Macro International Inc.

Contents

Tables

Figures

Preface

The Demographic and Health Surveys (DHS) program has become one of the principal sources of international data on fertility, family planning, maternal and child health, nutrition, mortality, and HIV/AIDS. The quality of these data is of utmost importance to researchers worldwide.

Because survey methodology has a major impact on data quality, one of the objectives of the MEASURE DHS project is to advance the methodology and procedures used to carry out national-level surveys. This will improve the accuracy and depth of information relied on by policymakers and program managers in developing countries.

The topics in the *DHS Methodological Reports* series are selected by MEASURE DHS staff in consultation with the U.S. Agency for International Development. While data quality is a main topic of the reports, they also examine issues of sampling, questionnaire comparability, survey procedures, and methodological approaches. Some reports are updates of previously published reports.

This report assesses the quality of health and nutrition data collected in the DHS surveys. It is part of a routine monitoring of data quality in the DHS surveys. The assessment is based on 81 surveys in 47 countries, conducted between 1993 and 2003.

It is hoped that the *DHS Methodological Reports* series will be useful to researchers and survey specialists, particularly those engaged in work in developing countries.

Ann A. Way
Project Director

Acknowledgments

Several DHS staff members provided valuable assistance during the preparation of this report. Special thanks go to Fred Arnold, Yuan Gu, Bridgette James, Vinod Mishra, Kaye Mitchell, Sidney Moore, Shea Rutstein, and Jerry Sullivan. Khatuna Doliashvili helped during the early stages of data preparation.

Executive Summary

The eleven years from 1993 through 2003 spanned two important phases of the Demographic and Health Surveys (DHS) project known as DHS III and MEASURE DHS+. During this interval, 81 surveys of women and their children were conducted in 47 different countries. A majority of the surveys (43) and countries (27) were in sub-Saharan Africa. The remaining surveys were in Latin America and the Caribbean, and in South Asia and Southeast Asia, with a few in Central Asia, Western Asia, and North Africa.

One of the most important objectives of the DHS project is to provide information about the health of women of reproductive age and young children. Many indicators of maternal and child health are provided by the responses to extensive questions about the most recent births, including antenatal care, delivery, nutrition, immunizations, symptoms of illness, and the treatment of illness. These questions are asked about surviving children born in a recent window of observation, usually the past five years, with some additional questions about the most recent birth, about children who did not survive, and about the woman herself. Using internationally accepted guidelines and standards, these indicators appear in country reports and comparative analyses and are used to assess progress and needs within the participating countries. It is crucially important for this information to be of the highest possible quality. The sampling design, structure of the questionnaire, training and supervision of interviewers, data processing, and analytical procedures are all directed toward this goal.

The purpose of this methodological report is to provide an overview of the quality of the maternal and child health data collected during DHS III and MEASURE DHS+. It is important to note that this assessment is not a response to any concerns, either general or specific, raised by users of the data. It is probably fair to say that within the general community of users, there is a general sense that the data are of very high quality. Rather, it is part of a routine monitoring process. It is desirable to conduct occasional checks and comparisons on the quality of data across surveys. There have been no comprehensive assessments of these data since DHS I and DHS II.

In this investigation it is not possible to compare responses with “true” values, or even to compare them with responses in a second interview of the same respondent. When a response is outside of a plausible range, then it can be inferred that the response is incorrect, but in general the quality of data can only be assessed at an aggregate level using indicators that are likely to be associated with incorrect responses.

Three principal indicators of data quality will be used. The first is the percentage of cases that are “missing” on a response—that is, the percentage of cases for which a numerical or pre-coded response would have been anticipated, based on the characteristics of the woman or child and the structure of the questions, but for which such a response was not obtained. We will often combine “don’t know” with “not stated” if either represents a complete loss of the case for analytic purposes, but will generally comment on the difference between the two.

The second indicator, which can only be measured for scaled or interval-level variables such as weight, height, or durations of time, is the degree of heaping or digit preference. Heaping, like “missing,” usually has little direct effect on substantive conclusions but suggests a lower standard of fieldwork and, if excessive, produces less confidence in the overall results.

The third indicator of data quality is the estimated bias in a mean or distortion in the distribution of a variable. If the level of missing data is substantial and if the probability of missing is systematically related to other characteristics, then the mean and distribution of the non-missing cases may differ from what they would be if the missing cases were imputed and added back in.

In general, it cannot be assumed that data are missing at random. When the level of missing is relatively large, the assessment uses logit regression to find whether the probability of missing is related to type of place of residence, mother's age, mother's education, the child's age, the number of surviving children in the window of observation, and the number of interviewer visits. In some instances, potential effects of interviewers, supervisors, etc., are checked. With these logit regressions it is often found that the incidence of missing is related to the covariates, and usually in an interpretable way. However, the bias is always found to be negligible. Bias requires the combination of a high level of missing as well as a highly systematic pattern to the missing responses, and the combination is never found. For that reason, this report does not actually have any findings of bias.

The assessment divides the indicators into three main sets that are considered in Chapters 2, 3, and 4. The topics of these chapters, respectively, are indicators of maternal health and maternity care; child health; and anthropometric measurements. The content and findings of these three chapters will now be briefly summarized.

Chapter 2, on maternal health and maternity care, first looks at antenatal care in terms of the date of the first visit, the number of visits, who provided such care, and whether tetanus toxoid injections were received. The overall levels of missing on these four indicators were 0.9 percent, 2.3 percent, 0.2 percent, and 1.3 percent, respectively. Information about the date of first antenatal care and who provided it was most likely to be missing for mothers in rural areas, or with no education, or for children who were relatively older (that is, for whom the care was more distant in time).

The highest level of missing was on the number of antenatal visits, exceeding 10 percent in three surveys. In those surveys, the better-educated women, who tend to have *more* such visits, were more likely to be missing on the variable. It appears that if the number is large, and therefore more uncertain, this question is more likely to go unanswered. It is recommended that respondents be advised to give a rough estimate or a range if they hesitate to give an exact number.

Two questions about recent births are considered: the type of person who assisted at the birth and where the birth occurred. The level of missing on these variables is extremely low, only about 0.2 percent across all surveys. The missing cases tend to occur for women who are rural and less educated. They are also more likely for women who had several births in the window or are older and have had many children, factors that could make it harder to recall the circumstances of a specific birth.

The durations of postpartum amenorrhea and sexual abstinence are included in Chapter 2, partly because of very high levels of heaping on multiples of six months. Myers' Index, the minimum percentage of cases that would have to be shifted to produce a smooth distribution, averages 25 percent for postpartum amenorrhea and 35 percent for abstinence. Guinea 1999 and Burkina Faso 1998/99, both in West Africa, are the two surveys with the highest levels of heaping on these durations.

With outcomes such as these, analysts often choose to estimate the median duration of amenorrhea, for example, with current status data rather than with the reported durations. "Current status" refers to whether or not a woman with a recent birth is still amenorrheic. The median duration would then be the estimated length of the open interval at which exactly half of the women are still amenorrheic. By contrast, a hazard model or failure time approach would use the reported durations, corrected for censoring for women who are still amenorrheic, to estimate the median. If computed for the same women, the two estimates would ideally agree very closely, but heaping on durations that are multiples of six months could be expected to produce a difference between the two estimates. The advantage of the current status estimate is that it is not affected by heaping, but the disadvantage is that it involves far fewer women, so the standard errors of estimates will be larger.

The current status and hazard model approaches are applied to the data on amenorrhea in the Burkina Faso 1998/99 survey, with a comparison between the two smoothed survival functions and the medians. The correspondence is surprisingly close, with an estimated median duration of 16.9 months from the hazard model versus 17.7 months from the current status model. The difference is only about 5 percent, or less than a month. These kinds of estimates appear to be robust with respect to heaping, and it may make little difference whether the hazard model or current status approach is used.

Chapter 3, on child health, begins with the responses to questions about diarrhea, fever, and cough in the past two weeks. The level of missing values is assessed for each of these symptoms as well as for whether a child with diarrhea received treatment and whether a child with fever or cough received treatment. The overall levels of missing are very low. There were six surveys in which 4 percent or more of responses were missing on all three symptoms. All of these were in sub-Saharan Africa: Côte d'Ivoire 1998/99, Gabon 2000, Namibia 2000, Tanzania 1999, Uganda 2000/01, and Zimbabwe 1999. The responses are more likely to be missing in urban areas, for younger mothers, and for older children. Most of the missing values can be traced to "don't know" responses when the mother and the child did not reside in the same household at the time of the survey.

Questions about treatment of symptoms were only asked if the symptoms had occurred during the past two weeks. The highest levels of missing on these questions, in excess of 10 percent, occurred in Bangladesh 1999/2000, Nigeria 1999, Philippines 1993, South Africa 1998, and Zimbabwe 1999. The level of missing was not significantly related to any of the covariates.

The assessment looks into the degree of seasonality, or variation across months of the year, in the reported prevalence of diarrhea, fever, and cough. Seasonality is not directly related to data quality, but is important analytically, particularly if there is substantial seasonality within a country and successive surveys in that country are conducted at different times of the year. In such situations, apparent declines or increases in the prevalence of symptoms may simply be seasonal variations.

In most DHS surveys, the fieldwork extends over an interval of three to six months, and there is usually a statistically significant level of variation, across the months of fieldwork, in the prevalence of the three symptoms. The greatest month-to-month variation, by far, is for cough. Using an arbitrary standard by which 17 surveys had very high variation across months in the prevalence of cough, only five surveys met the same standard for diarrhea and three met that standard for fever. In these surveys, the prevalence of a symptom is typically at least twice as high in the highest month as it is in the lowest month. The interval of data collection sometimes appears to include either the annual low or the annual high, but it is also common for prevalence to decline monotonically or to increase monotonically during the data collection, in which case it is impossible to be sure whether the data collection included the annual low or the annual high. The results suggest that a single survey may be able to identify differentials across sub-populations, but often cannot give a good picture of the overall seriousness of the symptoms, and consecutive surveys cannot give a good picture of trends unless they are conducted at the same time of year.

All DHS surveys collect information about duration of breastfeeding. In addition, 19 of the 81 surveys asked about liquids given to the child during the first three days after birth, and 51 asked about types of liquids and foods given to the child during the seven days before the interview.

For the duration of breastfeeding, three kinds of unusable responses are grouped together as "missing": no response; "don't know;" and "inconsistent," assigned during data processing if the stated duration was longer than the interval since the last birth. For all surveys, only 0.9 percent of cases were missing, and only eight surveys had more than 2 percent missing. All of these surveys were in sub-Saharan Africa. For them, the probability of a missing response was significantly greater for children age 2-4 than for children age 0-1, and the probability increased steadily for women with more children in the window. This pattern is plausible because breastfeeding has ended for most children age 2-4 and a woman with more children is asked to recall more durations and could confuse them.

In some surveys there is extreme heaping on durations that are multiples of six months. The greatest heaping was observed in the Bangladesh 1999/2000 survey, for which more than two-thirds of the responses were at multiples of six months, and fully one-quarter were at 24 months. The next worst heaping was in the Bangladesh 1996/97 survey. Nevertheless, as described above for postpartum amenorrhea and abstinence, there is little difference between the smoothed distributions using current status and failure time approaches.

In the 19 surveys that asked about liquids given to the child during the first three days after birth, only 0.3 percent of responses are missing. The highest level was in the Peru 2000 survey, in which it reached 2.3 percent. In that survey, "missing" was more likely in urban areas, for better educated women, and for women with only one child in the window. These are women who are most likely to have had the birth in a hospital, with a higher prevalence of the "don't know" response because the mother was unsure about liquids given by hospital staff.

Questions about liquids and foods in the last seven days have higher levels of missing responses than any other child health questions. Overall, 6.5 percent of children (in the 51 surveys that included the questions) were missing, and the levels were extremely high in Nicaragua 1997/98 (33.6 percent) and Mozambique 1997 (30.0 percent). The level exceeded 10 percent in another seven surveys. The probability of missing was not systematically related to any of the covariates. Most of the missing values, as with recent symptoms of illness, can be traced to "don't know" responses from mothers who do not live in the same household as their child. After further investigation, DHS could consider skipping the questions about current illness or about current liquids and foods if the mother and child live in different households.

DHS surveys routinely obtain information about nine childhood vaccinations: one vaccination against tuberculosis (referred to as BCG); four oral polio vaccinations (Polio 0, 1, 2, 3); three vaccinations against diphtheria, pertussis, and tetanus (DPT 1, 2, 3); and one vaccination against measles. These are supposed to be given at birth or within three months after birth, except for the measles vaccination, which is usually recommended at nine months. Ideally, vaccinations and their dates are recorded on a health card for each child.

Only a handful of surveys have a level of missing—defined to include "don't know"—that exceeds 1 percent on the BCG or Polio vaccinations, or on whether the child has a health card. The incidence of missing is somewhat higher for the DPT vaccinations, for which five surveys have at least 2 percent missing. It is highest for the measles vaccination, for which 23 surveys have at least 2 percent missing. The level of missing on the measles vaccination exceeds 4 percent in Burkina Faso 1998, Haiti 1994/95, Kazakhstan 1995, and Turkey 1999.

A selective examination of the surveys with the highest incidence of missing cases indicates that the mother's education, number of children in the window, and the elapsed time since the normal age at vaccination are sometimes related to the incidence of missing, but the impact of missing cases on the overall distribution of responses is negligible. Most of the missing responses come from the "don't know" code and from the cases for which the respondent claims to have a health card but is unable to show it to the interviewer.

Chapter 4, on anthropometric measurements, focuses on the heights and weights of mothers and children. A few surveys included other kinds of measurements (for example, of hemoglobin levels) that are not discussed in this assessment. The database consists of the same cases that are used in the other chapters. In recent DHS surveys (about one-third of the 81 surveys included in this report), heights and weights are measured for all children as part of the household survey. That procedure includes children whose mother is not living or is not in the household. The analysis omits those children, as well as women age 15-49 who do not have a child in the window. In terms of data quality, we would expect similar findings for the omitted children and women.

Heights and weights are measured in the metric system, with the measurements recorded in tenths of a centimeter (that is, in millimeters) for height and in tenths of a kilogram for weight. During data processing, the raw measurements are subjected to checks on whether they are in a plausible range. The range is the same for all women, but varies by age for children. Height and weight are also checked for compatibility with each other. These range checks are very liberal, and unlikely to lead to the deletion of cases that are extreme but correctly coded.

Using international standards, the raw measurements for women are converted to percentiles and z–scores for height, and for weight given height. Similar indices are calculated for children, but they also take the child's age into account. The body mass index (BMI) is also calculated. Using standard thresholds, the main indicators used by DHS are the percentages of children or women in a survey who are stunted, underweight, or malnourished.

If a recorded measurement of height or weight is outside the acceptable range, it is "flagged" with a special code for the constructed variables. The standard recode files distributed by DHS include the raw measurements of height and weight as well as the constructed variables, making it possible to infer the boundaries of the acceptable range.

Three potential problems are considered: incorrect measurement of height; incorrect measurement of weight; or, for children, incorrect measurement of age, which is important for determining whether a child is stunted or underweight. This chapter of the assessment goes into considerable detail about these possible factors, including the levels of missing data, heaping or digit preference, flagging, possible adjustments to the criteria for flagging, and consistency between reports of the child's age.

One finding is a much higher level of heaping on the final digit of height than on the final digit of weight, simply because a millimeter is a much finer distinction in terms of the range of heights than a tenth of a kilogram is in terms of the range of weights. Such heaping is acceptable, and there is no reason to believe that it induces any bias. Measurements of height to the nearest fifth or even half of a centimeter could be adequate.

The boundaries for flagging may be excessively liberal, bringing into the analysis a few cases at both ends of the height and weight distributions that are likely to be recording errors. If such cases survived validation in the field by a supervisor, they should of course be included.

The level of dropped data on height and weight, taking the form of "not applicable," "missing," and "flagged" values, is generally higher than for other variables in DHS surveys. Most of the "not applicable" codes, however, can be traced to the absence of children from their mother's household, or to subsampling, and do not reflect on the quality of the data. Highly significant variation across interviewers is found in most surveys, implying that the incidence of dropped cases could be substantially reduced by better training and tighter supervision of interviewers.

Chapter 5 explores two features of the standard survey design that could perhaps be modified. The first of these is a possible reduction in data collection, such that the child health information would be collected for a maximum of two children in the window.

The database for this study included 485,715 surviving children, of whom 472,670, or 97.31 percent, were either the most recent or the next most recent birth. The data include 12,236 children who are the third child in the window, 769 who are the fourth child, 39 who are the fifth child, and 1 who is the sixth child in the window. Relatively few children are in the third, fourth, fifth, or sixth positions. For various reasons, including the demand on interviewers, some evidence of poorer data, and data processing, a design that would omit such children could be considered.

Two scenarios are considered in which several vaccination and illness indicators are recalculated with partial data and compared with their values for the full data set. In one simulation, just the two youngest children are used. In the second scenario, the two youngest children are again used, but the next-to-youngest child is given a slightly inflated sampling weight to make it a better substitute for the omitted children. Under both scenarios, but particularly the second, the effect on indicators is very small. The index of dissimilarity, comparing the distribution with the full data set with the distribution under the first alternative scenario, is always less than 1 percent, reaching a maximum of 0.75 percent. With the second alternative scenario, the maximum is 0.26 percent.

A second possible modification of the survey design would be a slight increase, rather than reduction, in the data collected. This modification would be to collect vaccination data on children who have died. Apart from the measles vaccination, most vaccinations occur in the first six months. Many children who were born in the window but died before the survey date were alive at the ages when most vaccinations were received. The health card may not have been saved for a child who died, but this might not be a serious difficulty because in many surveys it is almost as common for the mother to say that the card is missing or not to show it as it is for the mother to be able to show the card. The value and possible costs of collecting such data would have to be taken into account, but it is suggested that the option be considered.

The overall conclusion of this assessment is that, through a combination of careful questionnaire design, training and supervision of interviewers, and checks and forced consistencies as part of data processing, DHS data on maternal and child health are generally of very high quality. Relatively speaking, missing or otherwise questionable data are only an issue for height and weight, particularly of children. This summary has identified a few surveys, mainly in sub-Saharan Africa, with relatively high levels of missing data or other signs of data problems. The assessment itself goes into much more detail and specificity than is possible here.

Perhaps the best evidence on the general quality of the DHS health data is given by the inclusion in the analysis of the survey conducted in Nigeria in 1999, a survey that is strikingly atypical of the general results. For example, in this survey, height and weight were missing for 15 percent of children and were flagged for another 46 percent of children. This level of flagging was more than 10 times the average level. Many other indicators were dramatically higher than for any other survey.

It happens that the Nigeria 1999 survey went into the field without DHS approval and without the normal level of training and supervision. The survey provides a textbook example of what can go wrong with a demographic survey. DHS does not generally distribute the data from the survey.

In 2003, Nigeria conducted another survey using the standard level of training and supervision. Its indicators of data quality are well within the range of all the sub-Saharan surveys. For example, in the 2003 survey, height and weight were missing for 7 percent of children and were flagged for 9 percent. A comparison of the 1999 and 2003 surveys of Nigeria is virtually a controlled experiment, illustrating the effect of DHS training and supervision with essentially the same setting and the same questionnaire.

This assessment includes some recommendations for relatively minor changes and further investigations, but the overall conclusion is that, to the extent that can be ascertained without reinterviews or factual verification of specific individual-level responses, the DHS data on maternal and child health are excellent.

1 Introduction

1.1 Objectives

One of the major functions of the Demographic and Health Surveys (DHS) program is to provide population-based estimates of the current health of women and children in developing countries. The surveys are a basis for many indicators of health, both published and unpublished, that are in turn used to evaluate the effectiveness of programs and interventions and to guide future changes in health programs. For many users of DHS data, the information about maternal and child health and nutrition has greater interest than the information about fertility and family planning.

Despite the widespread use of the health and nutrition data, there has not been a comprehensive assessment of its quality since the early 1990s. The absence of such an assessment is attributable to the widespread sense, with which we concur, that these data are generally very good, with low levels of missing values, good internal consistency, and good consistency with other sources. This report is intended to fill a gap and to provide some recommendations for possible marginal improvements. We will focus on the following objectives:

1) Identify surveys with relatively high levels of non-response, heaping, or other deficiencies that reflect on the quality of the data.

2) Identify possible biases in indicators of health and nutrition that arise from a non-random pattern of non-response. Non-response or incompleteness may be systematically related to such factors as the interviewer or keyer; the number of interviewer visits; seasonality of some kinds of illnesses; number of children in the interval of time before the survey (the window); and level of education, age, or other characteristics of the respondent or her children.

3) Investigate the potential impact of minor modifications to the survey design or questionnaire, specifically by obtaining information on, at most, two children in the or by obtaining vaccination information about children who have died.

1.2 Scope

The DHS survey program can be broken into a series of time intervals or phases that correspond roughly with funding cycles. This analysis will include 81 surveys, consisting of all 44 surveys conducted as part of DHS III, during the time interval from 1993 through 1998, and all 37 surveys conducted as part of MEASURE DHS+, from 1999 through 2003. DHS III and MEASURE DHS+ will sometimes be referred to as DHS phases 3 and 4, respectively.

The analysis will include most variables on maternal and child health and nutrition that appeared in these surveys with standard recode labels, including some variables that were only included in a few of the surveys. A complete list of variables is provided in Appendix B, and the next section of this introduction will give an overview. Some of the categories of these variables were country-specific. We do not include data related to HIV/AIDS or country-specific variables (which generally have the prefix “cs”). Over time, some country-specific categories of standard variables, and some country-specific variables themselves, may be used more broadly. They could be assessed using the same methods that are used here.

The child health variables are generally obtained for all surviving children born in an interval of time before the survey, usually five years. This interval will be referred to as the “window” of observation and the children will sometimes be referred to as “index” children. The health variables included in this assessment were not obtained for children born outside the window, for children who were born in the window but died before the interview, or for women who did not have a birth in the window. Maternal health variables relate to the circumstances of the most recent pregnancy that resulted in a live birth in the window. Those variables are included even if the child later died.

Table 1.1 gives the number of children born in the window, the number of surviving children for whom the detailed health information was obtained, and the number of women who had at least one child (regardless of survival status) in the window, for each of the 44 surveys conducted as part of phase 3. There were a total of 250,656 children born in the window, of whom 231,809 (92.5 percent) survived to the date of the survey. A total of 185,741 women had at least one child in the window. Table 1.2 gives the corresponding frequencies for the 37 surveys conducted as part of phase 4. These surveys included a total of 276,397 children born in the window, of whom 253,906 (91.9 percent) survived to the date of interview. The surveys obtained information on 199,845 mothers.

In the full set of 81 surveys conducted from 1993 through 2003, a total of 527,053 children were born in the window, of whom 485,715 survived to the date of the interview, with a total of 385,586 mothers.

Many countries appear in both Table 1.1 and Table 1.2, and some countries had two surveys within the same DHS phase. Of the 47 different countries that conducted a DHS survey from 1993 through 2003, 20 countries had only one survey, 20 had two surveys, and 7 had three surveys. The countries that had three surveys are Bangladesh, Bolivia, Dominican Republic, Ghana, Indonesia, Kenya, and Philippines. Of the 47 different countries, 27 are in sub-Saharan Africa, eight are in the Caribbean or Central or South America, six are in South or Southeast Asia, and the remaining six are in North Africa and in Western and Central Asia.

The consolidated data files are limited to children whose ages at the date of interview were in the range for which the child health questions were asked. This was usually ages 0-4 years (0-59 months). The age range was 0-3 years (0-47 months) for Uganda 1995. The eligible age range for children was 0-2 years (0-35 months) for the following 23 surveys: Bangladesh 1993, Benin 1996, Bolivia 1994, Cameroon 1998, Central African Republic 1994/95, Comoros 1996, Cote d'Ivoire 1994, Ghana 1993, India 1998/99, Kazakhstan 1995, Kenya 1998, Kyrgyz Republic 1997, Madagascar 1997, Mali 1995/96, Mozambique 1997, Nepal 1996, Niger 1998, Nigeria 1999, Togo 1998, Uzbekistan 1996, Vietnam 1997, Vietnam 2002, Zimbabwe 1994.

Blocks of variables will be summarized with key indicators of missing, bias, heaping, and other possible problems (see below for definitions of these terms). Except where indicated, an indicator of data quality will not include cases for which the relevant variable is completely missing, i.e., is coded as a blank in the raw data or as "." in Stata. This code signifies that a variable was not applicable to the specific case because of skip patterns, or because it was omitted from the survey. DHS is careful to distinguish that type of omission from other types of missing data that occur during the data collection process. For our purposes, "not stated"[1] generally means that the question or item was applicable to the specific case but was given a special code, such as 9, to indicate that none of the valid options or values was selected. This kind of non-response can be attributed to a refusal from the respondent, interviewer error (skipping a question, not probing sufficiently, etc.), an invalid code at some phase of data entry, etc. Another type of non-response that will generally be combined with the "not stated" code is "don't know." For example, for the question about whether a child had a fever in the past two weeks, code 8 is used to indicate that the respondent doesn't know whether the child did or did not have a fever. For most purposes, this kind of a response probably has the same sources and implications as "not stated." Rather than treat "don't know" as a valid response or analyze the patterns of "don't know" and "not stated" separately, which would have

[1] The label usually used by DHS for these non-responses is "missing." In this assessment, we will sometimes, as here, substitute the label "not stated."

required nearly doubling the numbers of tables and figures, we will usually group them together under the label "missing."

The level of missing data is important in itself as an indicator of data quality. A high proportion of "not stated" responses can imply that the interviews are too demanding, the questions are poorly worded, the interviewers have not been trained adequately, and so on. At the very least, "not stated" responses indicate wasted fieldwork and statistical inefficiency, because they limit the analysis to a smaller number of valid responses.

"Bias" refers to the possible effect of a non-random pattern of missing cases. Suppose, for example, that the percentage of children who had a fever in the past two weeks is calculated using only the valid codes 0 for "no" and 1 for "yes," and ignoring the codes 8 for "don't know" and 9 for "not stated." If codes 8 and 9 occur at random, then percentages calculated just from codes 0 and 1 will not be biased. However, suppose that codes 8 and 9 *do not* occur at random, and are actually more likely to occur when the true response would be 1 rather than 0. In this case, children who *did* have a fever in the past two weeks are more likely to be omitted from the calculation of the percentage than children who *did not* have a fever. This pattern would cause the percentage to be biased downwards; the true percentage who had a fever would be higher than what is estimated just from codes 0 and 1. The estimated percentage minus the true percentage is the bias. In this hypothetical example the bias would be negative. The actual bias can never be known, but we will apply a procedure to estimate its direction and magnitude.

There may well be other kinds of bias due to systematic measurement error that are even more important, but that cannot be assessed here. For example, with the question about whether the child had a fever in the past two weeks, there is probably some subjectivity and cultural variation in the definition of a fever. The reference period of two weeks is subject to distortion.

This report will also consider possible biases that result from age misstatement or transfers. Weights and heights of children are compared with norms based on their ages. If age is misreported then the norms that are looked up in reference tables will be incorrect, possibly leading to under- or over-statements of nutritional deficiencies.

"Heaping" refers to the tendency for reports or measurements of interval-level variables to be concentrated at certain values. This can be important for the calculation of some rates, means, or percentages if the heaped values occur at the boundary of an interval. Otherwise, heaping, like "not stated," may simply cast some doubt on the quality of the overall data collection process.

Finally, this report will include some simulations of alternative study designs that could affect the cost and efficiency of DHS surveys. These include reducing the number of children selected within the window and adding more questions about children who have died.

Table 1.1 Number of children born in the window, number of surviving children, and number of mothers in the DHS surveys, phase 3 (1993-1998)

	Children born in window		
Survey and date	All	Surviving	Mothers
Bangladesh 1993/94	3,865	3,545	3,569
Bangladesh 1996/97	6,167	5,579	4,585
Benin 1996	3,011	2,741	2,658
Bolivia 1994	3,654	3,368	3,075
Bolivia 1998	7,304	6,766	4,788
Brazil 1996	5,045	4,818	3,761
Burkina Faso 1998/99	5,953	5,076	3,960
Cameroon 1998	2,317	2,123	2,013
Central African Republic 1994/95	2,816	2,561	2,423
Chad 1996/97	7,408	6,361	4,552
Colombia 1995	5,141	4,976	3,824
Comoros 1996	1,145	1,056	934
Cote d'Ivoire 1994	3,998	3,660	3,559
Cote d'Ivoire 1998/99	1,992	1,732	1,439
Dominican Republic 1996	4,643	4,413	3,155
Egypt 1995	12,135	11,274	8,027
Ghana 1993	2,204	2,056	1,980
Guatemala 1995	9,952	9,360	6,065
Haiti 1994/95	3,564	3,208	2,273
Indonesia 1994	17,738	16,653	13,841
Indonesia 1997	17,083	16,257	13,731
Kazakhstan 1995	846	811	732
Kenya 1993	6,115	642	3,904
Kenya 1998	3,531	3,275	3,058
Kyrgyz Republic 1997	1,127	1,068	984
Madagascar 1997	3,681	3,344	3,128
Mali 1995/96	6,031	5,238	5,163
Mozambique 1997	4,122	3,770	3,732
Nepal 1996	4,417	4,097	3,845
Nicaragua 1997/98	8,454	8,084	5,824
Niger 1998	4,798	4,247	4,085
Peru 1996	17,549	16,600	12,403
Philippines 1993	9,195	8,823	5,795
Philippines 1998	8,083	7,751	5,240
Senegal 1997	7,372	6,628	4,786
South Africa 1998	4,978	4,726	4,148
Tanzania 1996	6,789	6,080	4,540
Togo 1998	4,168	3,873	3,757
Turkey 1993	3,724	3,527	2,774
Uganda 1995	5,756	5,188	4,013
Uzbekistan 1996	1,324	1,261	1,151
Vietnam 1997	1,775	1,724	1,633
Zambia 1996	7,248	6,177	4,616
Zimbabwe 1994	2,438	2,292	2,218
Total	**250,656**	**231,809**	**185,741**

Table 1.2 Number of children born in the window, number of surviving children, and number of mothers in the DHS surveys, phase 4 (1999-2003)

Survey and date	Children born in window: All	Children born in window: Surviving	Mothers
Armenia 2000	1,726	1,659	1,291
Bangladesh 1999/2000	6,813	6,309	5,176
Benin 2001	5,349	4,740	3,553
Bolivia 2003/04	10,448	9,802	7,325
Burkina Faso 2003	10,645	9,365	7,367
Colombia 2000	4,670	4,561	3,618
Dominican Republic 1999	597	583	431
Dominican Republic 2002	11,362	11,008	8,059
Egypt 2000	11,467	10,951	8,001
Ethiopia 2000	10,873	9,560	7,245
Gabon 2000	4,186	3,915	2,957
Ghana 1998	3,298	3,026	2,376
Ghana 2003	3,844	3,530	2,777
Guatemala 1998/99	4,943	4,687	3,030
Guinea 1999	5,834	5,040	4,035
Haiti 2000	6,685	6,077	4,348
India 1998/99	33,026	30,984	28,978
Indonesia 2002	16,010	15,350	13,349
Kazakhstan 1999	1,345	1,266	1,068
Kenya 2003	5,949	5,447	3,972
Malawi 2000	11,926	10,367	7,941
Mali 2001	13,097	11,109	8,277
Mozambique 2003	10,326	9,129	7,007
Namibia 2000	3,989	3,784	3,008
Nepal 2001	6,931	6,416	4,731
Nicaragua 2001	6,986	6,727	5,088
Nigeria 1999	3,549	3,208	3,106
Nigeria 2003	6,029	5,186	3,775
Peru 2000	13,697	13,130	10,499
Philippines 2003	7,145	6,892	4,920
Rwanda 2000	7,922	6,857	4,964
Tanzania 1999	3,215	2,839	2,118
Turkey 1998	3,565	3,403	2,669
Uganda 2000/01	7,113	6,350	4,252
Vietnam 2002	1,317	1,302	1,221
Zambia 2001/02	6,877	5,997	4,495
Zimbabwe 1999	3,643	3,350	2,818
Total	**276,397**	**253,906**	**199,845**

1.3 Overview of the Core Health and Nutrition Information

In the standard recode file of women and children for each survey, the variable names have the following prefixes: b (for birth), h (for child health, mainly related to vaccinations, diarrhea, fever, and cough), hw (for children's height and weight), m (for prenatal care, delivery, breastfeeding, other feeding), ml (for malaria), and v4 (for various aspects of maternal health and nutrition, and indicators specifically for the most recent birth).

The "b" variables are only used here because they include key information about the birthdate, sex, and survival of the child. The birthdate and sex of the child are the main variables from this block that will be used.

This report includes some use of most of the detailed items listed in Appendix B, but most of the items on that long list are options, dates, or recodes of a much shorter list of key variables. For example, to learn whether a child received treatment for fever or cough, DHS provides a list of possible places where the child could have been taken or medications that could have been given. These are not mutually exclusive alternatives; the child could have been taken to more than one place or could have been given more than one medication. The most efficient way to record such alternatives—and the one employed by DHS—is as separate items, typically with codes for "yes," "no," "don't know," or "not stated." If such alternatives were coded within a single variable, it would not be possible to select more than one alternative. It would also be harder to manage different menus of options in different surveys.

In order to succinctly summarize these detailed items, we use a list of primary health and nutrition indicators. For each of them, two complementary measures have been constructed. One measure consists of the non-missing values or categories of the variable. The other measure focuses on the level of "missing" for these questions and is defined to be "1" if the response to a substantive question is missing or if *any* of the components of a set of questions are "missing" (including "don't know"). A measure of "missing" is coded "." if *all* of the components are ".", and is coded 0 otherwise. The names of the variables are intended to be self-explanatory.

All of the primary indicators refer to children except for the "v4" variables on weight and height, and the durations of postpartum amenorrhea and sexual abstinence, which refer to women. Three "m" variables, on antenatal care (who and where) and on the delivery, refer to the pregnancy and delivery and therefore have implications for the mother as well as the child.

Primary indicators for children:

- Health card (h1)
- Vaccinations (h0, h2, h3, h4, h5, h6, h7, h8, h9)

- Diarrhea (h11)
- Treatment of diarrhea (h13, h14, h15)
- Fever (h22)
- Cough (h31)
- Severe cough (h31b)
- Treatment of fever and cough (h32a-z)

- Weight (hw2)
- Height (hw3)
- Date of weight and height measurements (hw17, hw18, hw19)

- Date of first antenatal visit (m13)
- Who provided antenatal care (m2a-m)
- Place antenatal care was provided (m57a-x)
- Place of delivery (m3a-m)

- Duration of breastfeeding (m4)
- Liquids in first three days (m55a-z)
- Foods and liquids in last seven days (m40a-xz)

Primary indicators for women:

- Weight (v437)
- Height (v438)
- Duration of postpartum amenorrhea (m6)
- Duration of postpartum sexual abstinence (m8)

This analysis omits a number of variables that were included in some surveys and that would be of particular interest to some analyses:

- Components of antenatal care other than tetanus toxoid injections
- Delivery characteristics (caesarean sections, birth weight and size)
- Postpartum care
- Diarrhea treatment, knowledge of oral rehydration salts, feeding during diarrhea
- Hand washing materials
- Disposal of children's stools
- Access to health care
- Smoking
- Frequency of breastfeeding and meals, use of bottles
- Salt iodization
- Micronutrients, vitamin A, and iron supplements
- Nightblindness
- Anemia in children and women

The omitted variables were only included in a few surveys, sometimes with country-specific variable names and categories. They would have added only marginally to our findings about the quality of specific surveys and different types of questions. If desired, the methods used here could be extended to those variables.

1.4 Methods

The analysis is based on consolidated computer files that were constructed from the unrestricted files of women from all the phase 3 and phase 4 surveys that are in general distribution by DHS and are listed in Tables 1.1 and 1.2. A consolidated file of all surviving children born in the specified window in all 81 surveys was constructed, along with a separate file of the mothers (including women who had at least one live birth in the window, even if they had no surviving births). The child file was constructed directly from the surveys of women, rather than from the child files that DHS distributes. Children's records include all relevant variables from the mother's data, specifically variables that begin with v0, v1, and v4. Mothers' records include all information about the youngest child born in the window.

Some of the analysis uses an entire consolidated data file without weights–that is, treating all cases from all surveys equally. Other parts of the analysis use all surveys in an entire file but with weights that are proportional to the usual weight variable, v005, re-scaled so that each survey in the file has equal weight. Re-scaled weights prevent the results from being excessively influenced by the largest surveys. Finally, some of the analysis is repeated for specific surveys or groups of surveys within the consolidated files. All file preparation and analysis were done with the Stata statistical package, version 8.

So far as possible, the analysis is done with individual-level statistical models such as binomial logit regression and multinomial logit regression. Statistical models are essential for the multivariate modeling of the missing cases, but are also helpful even when calculating simple proportions because they produce standard errors that take account of the study design.

In these models, sample weights and clustering are taken into account. In general, sample weights affect estimated coefficients, correcting for bias that would result from ignoring the variation in sampling fractions that are used in specific surveys. Taking account of clustering by primary sampling units (PSUs) does not affect estimated coefficients but generally increases the estimated standard errors somewhat. Because of the wide variation across surveys in the use of stratification, models will not be adjusted for that effect. In general, stratification does not affect estimated coefficients but slightly decreases the estimated standard errors.

Data quality is investigated with respect to the following general kinds of indicators for specific outcomes:

- The levels of "not stated" or "don't know" responses
- The extent to which these kinds of responses are systematically related to specific covariates
- The estimated bias in a distribution, inferred by comparing the distribution when "not stated" or "don't know" responses are ignored with the distribution when they are allocated to the valid responses under some rule
- Indexes of heaping or digit preference, and other internal evidence of displacement for interval-level variables
- Estimates of bias in age-related characteristics that may be due to misstatement of age.

Sometimes the standards for data quality are based on the level of the indicator. For example, if only 1 percent of the cases are "missing" on some indicator, virtually anyone would infer that "missing" responses were not a problem and that there is no point in checking whether they are random. Generally, the assessment will be in relative terms and will focus upon the surveys with the highest levels of the indicator.

The procedure for identifying covariates of missing data is described in detail in Appendix A but will be briefly summarized here. For variables of interest with a relatively high level of missing data, a binary variable is constructed that is coded 0 if the variable is not missing and 1 if it is missing. This binary variable is regressed on each of six covariates, in turn. The six covariates are type of place of residence, age of the child's mother or of the woman herself in 5-year intervals, level of education of the child's mother or of the woman herself, the child's age in single years, number of surviving children in the window, and number of visits by the interviewer. These logit regressions are weighted and corrected for clustering by primary sampling units. Association with a covariate is considered to be statistically significant if the chi-square for the logit regression exceeds the critical value for a .01 level of significance, given the degrees of freedom for the model. It is considered to be substantively important if the pseudo–R^2 for a model is .01 or greater, i.e., if the model accounts for at least 1 percent of the variation or deviance in the outcome.

The distributions of these covariates are given in Table 1.3, first for all surviving index children and then for the mothers of any children born in the window. The distributions are weighted to give equal weight to each survey. The distributions are very similar for the children and the mothers, except for the distribution of the surviving children of the mothers and the surviving sibship of the children. This is because there are twice as many children in a sibship of size two as a sibship of size one, and so on, with the result that large sibships are more common from a child's perspective than from the mother's perspective.

In a second step, the non-missing categories of the variable are used as the dependent variable in a multinomial logit regression (weighted and corrected for clustering). The predictor variables in this regression are all six of the categorical variables listed above, included additively and regardless of whether they were significant in the preceding logit regressions. We then calculate the observed distribution across the legal categories of the variable for the non-missing cases and, more importantly, the distribution across these categories for the missing cases, fitted by the multinomial logit regression. The difference between these two distributions is measured by the index of dissimilarity, which is the sum of the absolute differences between the respective percentages in the observed and fitted (for missing cases) distributions. The index of dissimilarity can be interpreted as the percentage of cases in one distribution that would have to be shifted in order to match the other distribution. Then the non-missing and missing cases are pooled, and the index of dissimilarity is again calculated to measure the difference between the observed distribution and the pooled distribution—that is, the distribution adjusted for missing cases using the association with the six covariates. The value of this index is perhaps the single best quantitative measure of any bias in the observed distribution.

Table 1.3 Percentage distributions of the six covariates used to assess possible bias in missing data, given separately for children and mothers, with equal weight for each survey

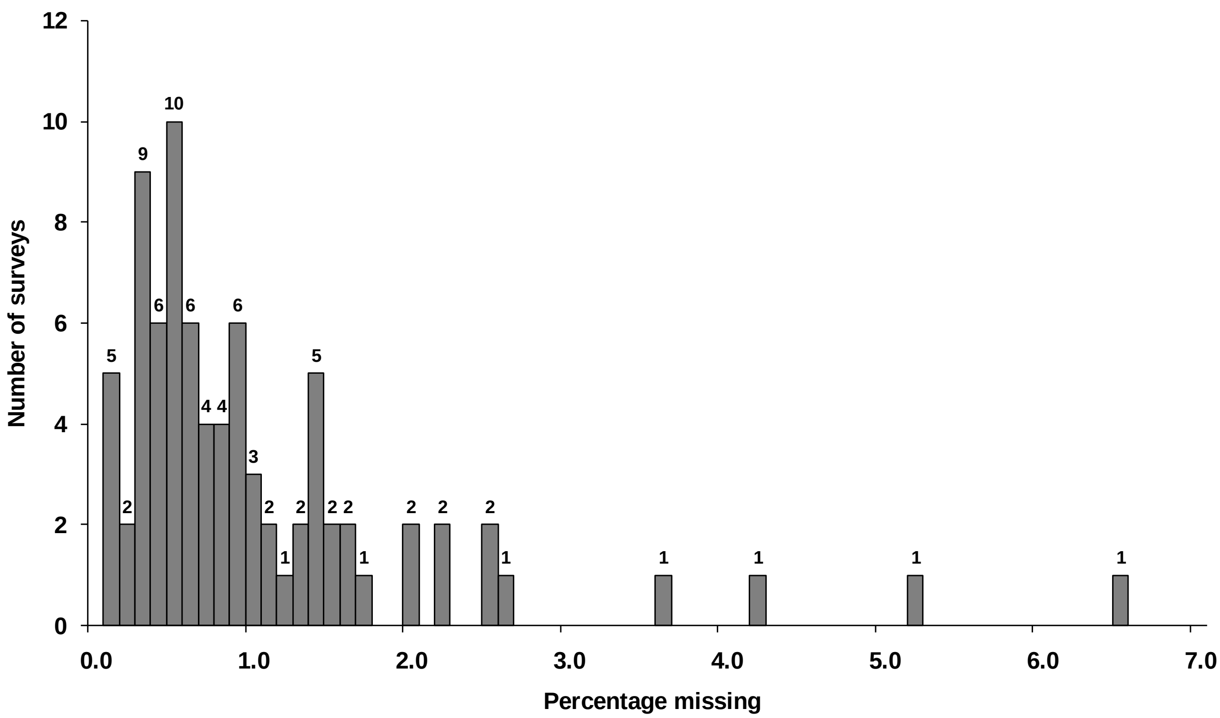
Tabulated for surviving children		Tabulated for mothers
Residence		**Residence**
Urban	34.2	34.7
Rural	65.8	65.3
Mother's age interval		**Woman's age interval**
15-19	7.3	8.6
20-24	26.4	26.0
25-29	27.6	26.4
30-34	19.6	19.2
35-39	12.4	12.4
40-49	6.8	7.5
Mother's level of education		**Woman's level of education**
None	32.6	32.6
Primary	35.9	34.9
Secondary	26.8	27.6
Higher	4.7	4.9
Age of child		**Not applicable**
0	25.0	-
1	24.1	-
2	23.2	-
3	14.2	-
4	13.6	-
Surviving sibship in window		**Surviving children in window**
1	54.8	70.8
2	37.9	25.7
3 or more	7.3	3.6
Number of interviewer visits		**Number of interviewer visits**
1	91.6	91.6
2	6.6	6.6
3 or more	1.8	1.8

2 Maternal Health and Maternity Care

2.1 Antenatal Care, Including Tetanus Toxoid Immunization before Birth

DHS surveys include many questions about antenatal care for the pregnancy preceding each index child. We will examine just four such questions: when was the first antenatal visit (m13), how many visits were there (m14), who provided the care (m2), and did the mother receive any tetanus toxoid injections during the pregnancy (m1).[2] Overall (giving equal weight to each survey), responses were missing for only 0.9 percent of responses on when the first antenatal visit occurred, 2.3 percent on how many visits there were, and 0.2 percent on who provided the care. Only 1.3 percent of responses were missing on whether the mother received a tetanus toxoid injection. The distributions of the percentages of missing cases on these four variables, across all phase 3 and 4 surveys, are given in Figures 2.1 through 2.4. (Note that these Figures have different horizontal scales.)

Figure 2.1 Distribution of the percentages of cases that are missing the date of the first antenatal visit (m13), all DHS surveys 1993-2003



[2] A question concerning where the treatment took place (m57) was included in only three surveys: Ghana 1998, Kenya 2003, and Nigeria 2003. Its level of missing was only 0.1 percent and it will not be discussed.

Figure 2.2 Distribution of the percentages of cases that are missing the number of antenatal visits (m14), all DHS surveys 1993-2003

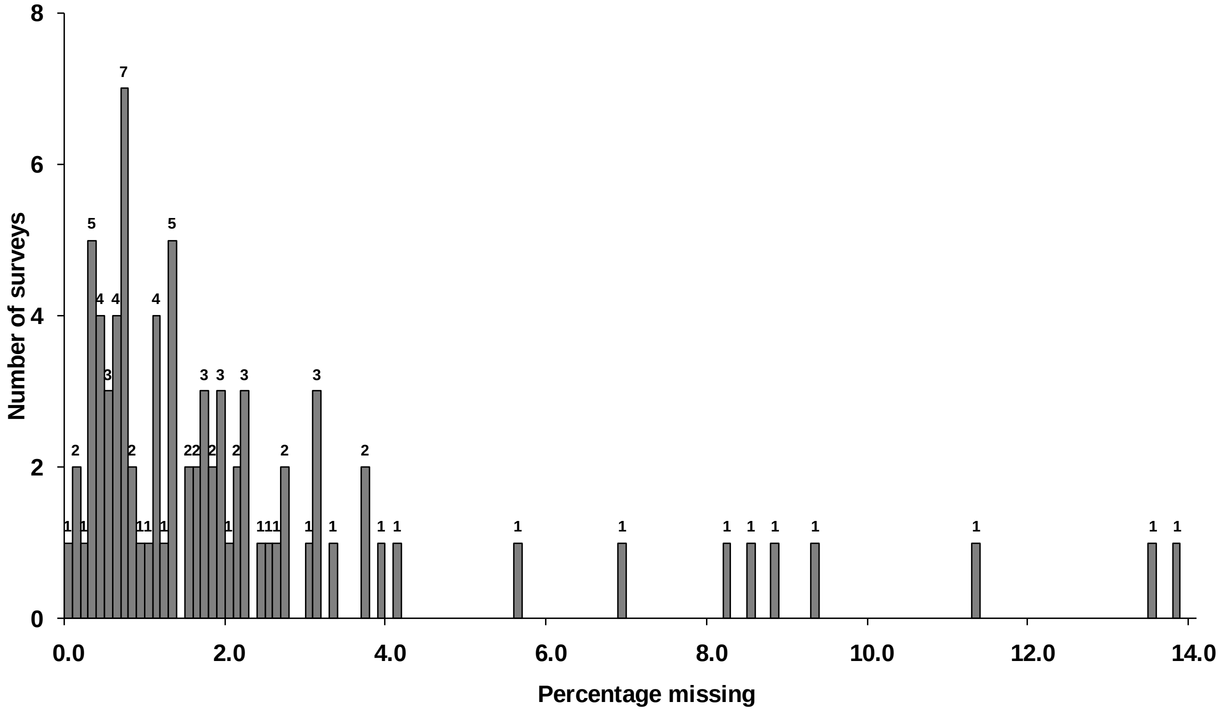


Figure 2.3 Distribution of the percentages of cases that are missing who provided antenatal care (m2), all DHS surveys 1993-2003

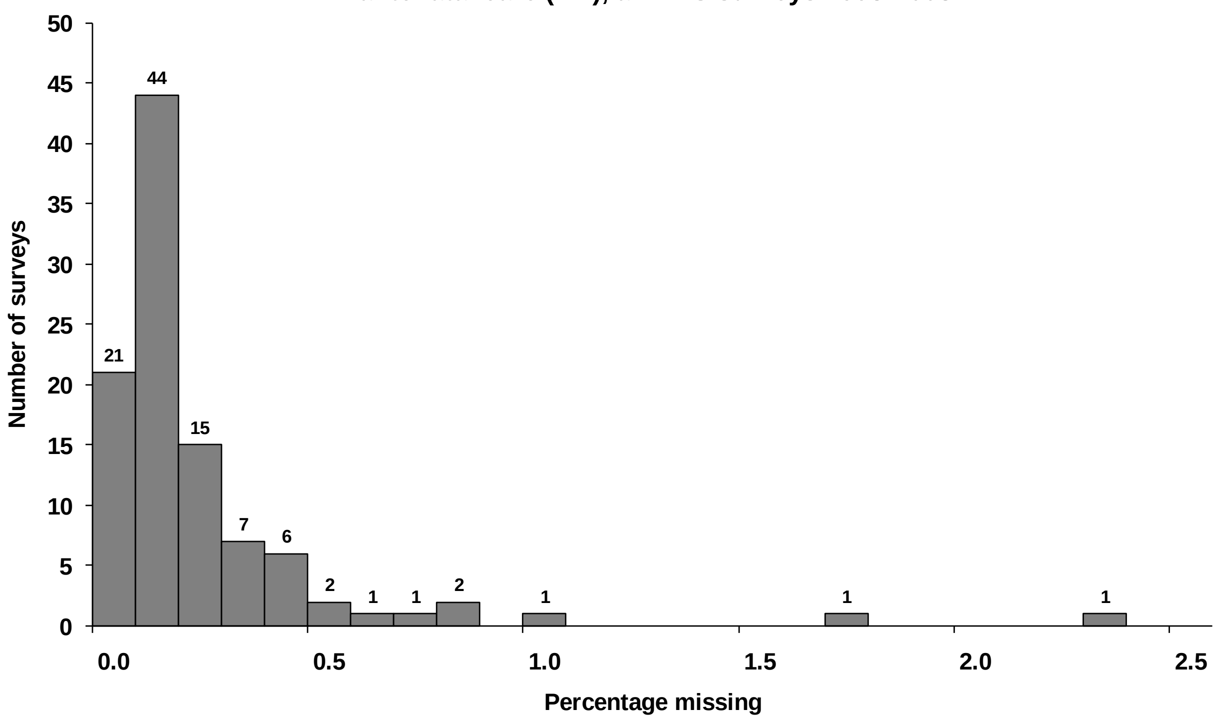


Figure 2.4 Distribution of the percentages of cases that are missing whether the mother received any antenatal tetanus toxoid injections (m1), all DHS surveys 1993 2003 other than Armenia 2000, Kazakhstan 1995 and 1999, Kyrgyz Republic 1997, Namibia 2000, and Uzbekistan 1996, which omitted this question

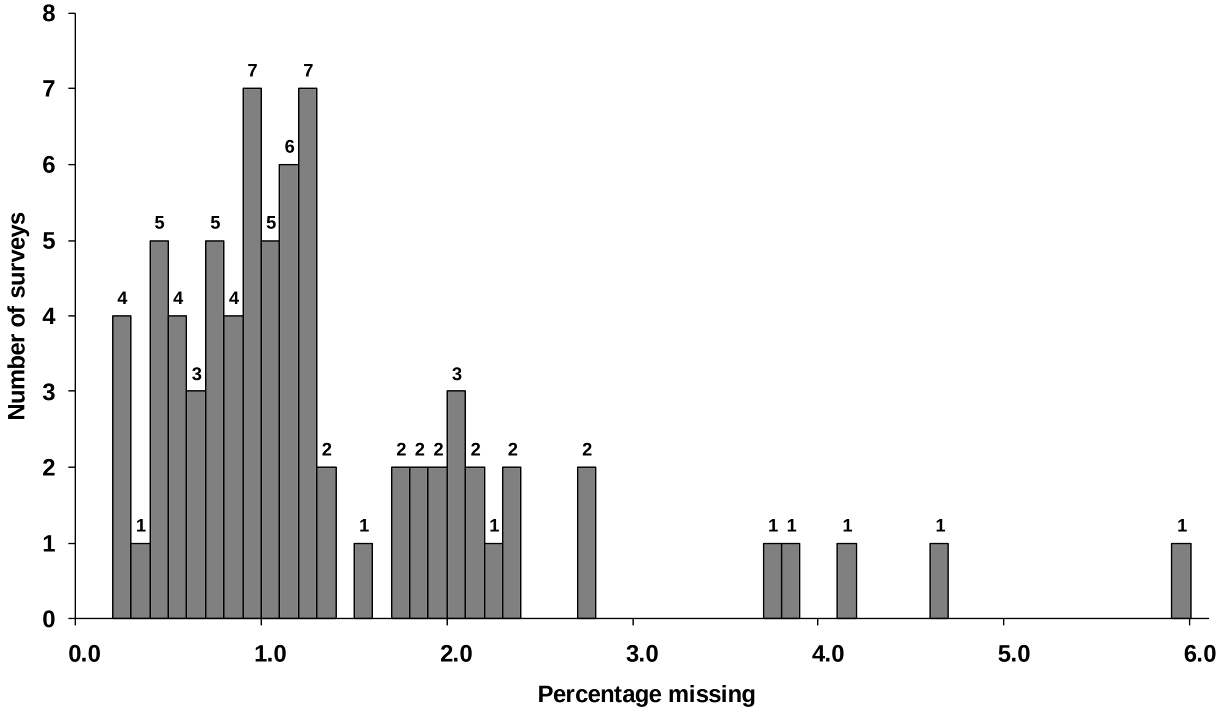


Table 2.1 gives the percentages of responses that were missing (including "don't know") that were 2 percent or greater on any of these four indicators. Percentages below 2 percent are omitted so the table is easier to read. Surveys for which none of these percentages reached 2 percent are omitted from the table. Table 2.1 shows a very low level of missing on all of these items in most surveys. Three of the highest percentages are found for the Nigeria 1999 survey, which will show up repeatedly in this assessment for reporting problems, and which was conducted largely independently of DHS.

In an effort to find a pattern to the missing responses, we have done a series of logit regressions of "missing" versus "not missing" on each of the six covariates listed in Section 1.4. The first set of regressions consists of a pooling of the surveys that were missing at least 2 percent of responses to the question about the date of the first antenatal visit—the nine surveys identified in the first column of Table 2.1. The nine surveys are weighted such that each survey counts equally (but the within-survey weights are retained). In this series of six regressions, we only consider a potential covariate to be a useful indicator of the missing response if it produces a pseudo-R^2 of at least .01, that is, if at least 1 percent of the deviance in the binary response variable is explained by the covariate. Such a pseudo-R^2 always has an extremely high level of statistical significance because the number of cases is large. Much lower values of the pseudo-R^2 are also statistically significant at the $\alpha = .01$ level, say, but a model that explains less than 1 percent of the deviance will be deemed not to be *substantively* significant.

Table 2.1 Surveys in which 2 percent or more of weighted responses to questions about antenatal care or tetanus toxoid injection are missing, all DHS surveys 1993-2003

Survey	Antenatal visits When	Antenatal visits How many	Antenatal visits Who	Tetanus toxoid vaccinations
Armenia 2000		3.6		–
Benin 2001		2.2		2.2
Bolivia 1998				2.2
Bolivia 2003/04				2.7
Brazil 1996		2.1		4.1
Burkina Faso 1998/99	4.2			
Burkina Faso 2003	2.1			
Cameroon 1998		2.1		
Central Africa Republic 1994/95		2.6		
Comoros 1996		3.0		
Egypt 2000		3.9		
Ghana 2003		3.3		2.0
Guinea 1999		3.7		
Kazakhstan 1995		3.1		–
Kazakhstan 1999		13.4		–
Kenya 1993		3.0		
Kenya 1998		2.1		
Kenya 2003		2.5		
Kyrgyz Republic 1997		9.3		–
Mali 1995/96	3.4	3.0		
Mali 2001	2.6	5.5		
Mozambique 1997	2.5	8.5		
Mozambique 2003				2.6
Namibia 2000	2.2	8.8		–
Nepal 1996	6.5	2.7	2.3	
Nicaragua 1997/98				2.1
Nicaragua 2001				3.7
Nigeria 1999	5.2	11.2		4.5
Senegal 1997	2.6			
South Africa 1998		8.2		5.9
Tanzania 1996		4.0		
Turkey 1998				2.3
Uzbekistan 1996		6.9		–
Zambia 1996		2.2		
Zambia 2001/02		2.4		
Zimbabwe 1994		2.5		
Zimbabwe 1999		13.7		3.7

Note: A dashed line indicates that the question was not included in this survey.

In the first series of logit regressions, only the woman's type of place of residence (pseudo-R^2 = .010) and level of education (pseudo-R^2 = .017) are substantively significant predictors of missing data. Figures 2.5 and 2.6 show that rural women and women with no education are much more likely to be missing the date of the first antenatal visit. Above that level, however, there is little variation by education.

Figure 2.5 Percentage of cases missing date of first antenatal visit, by woman's type of place of residence, for the nine surveys with at least 2 percent missing on this variable (pseudo-R^2 = .010)

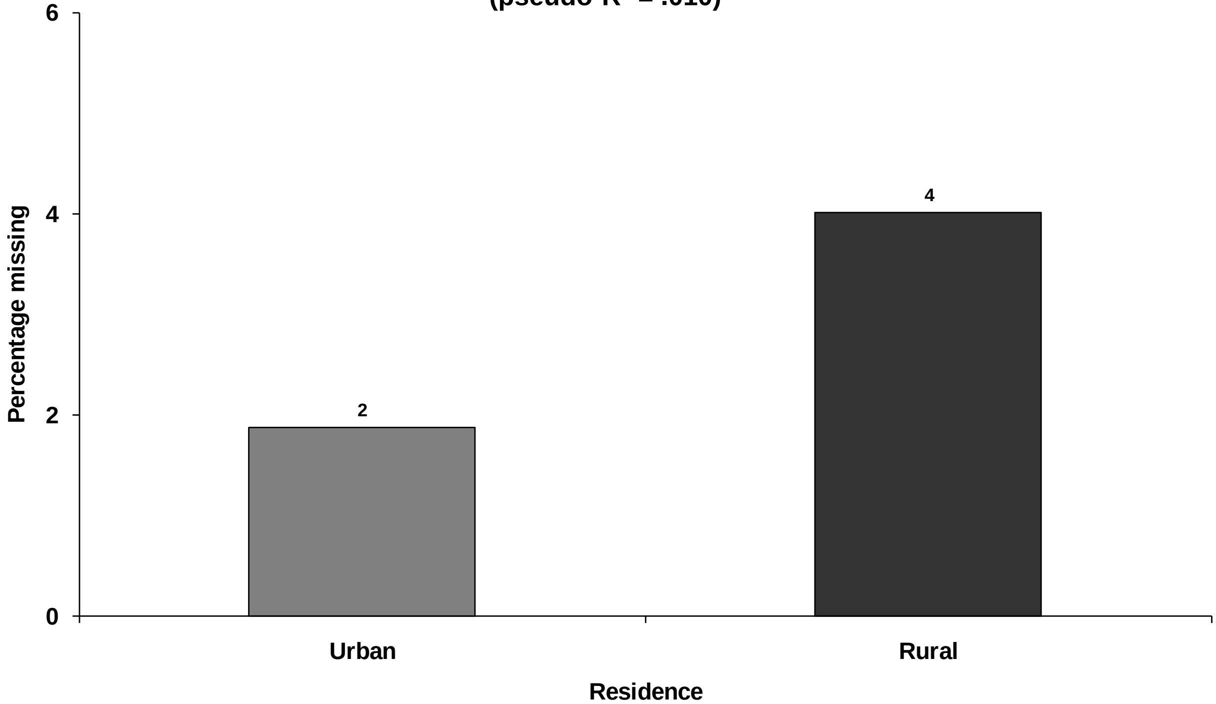
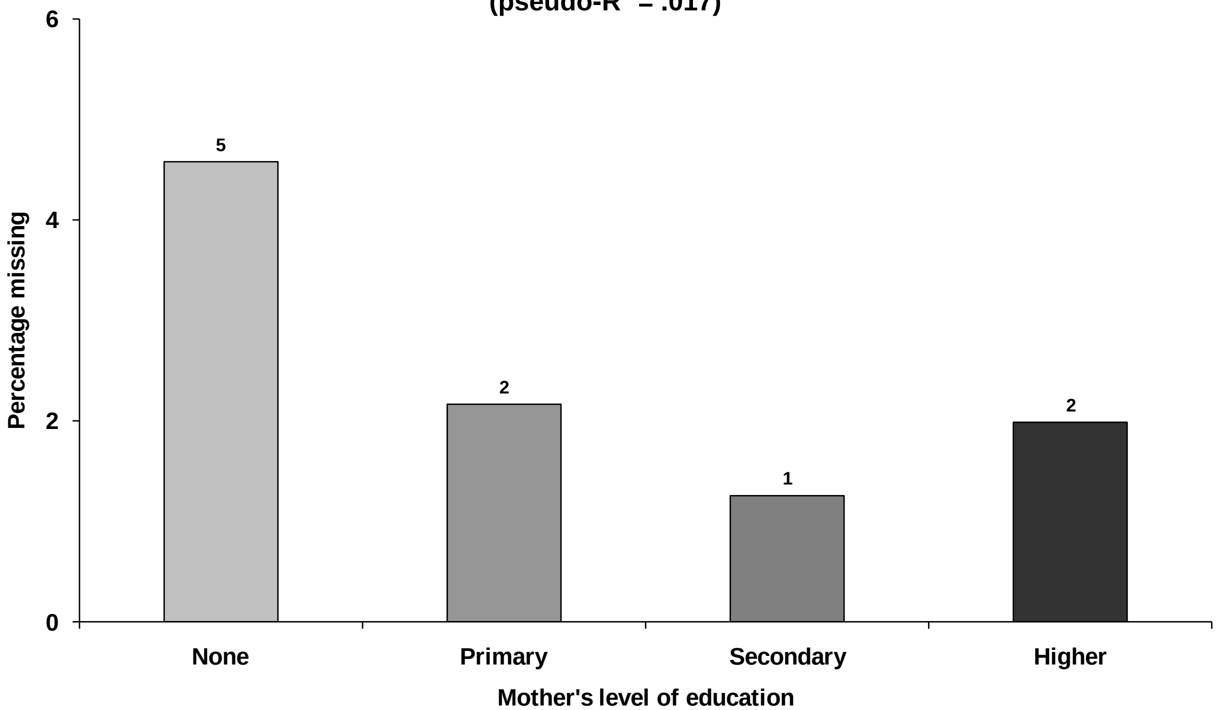
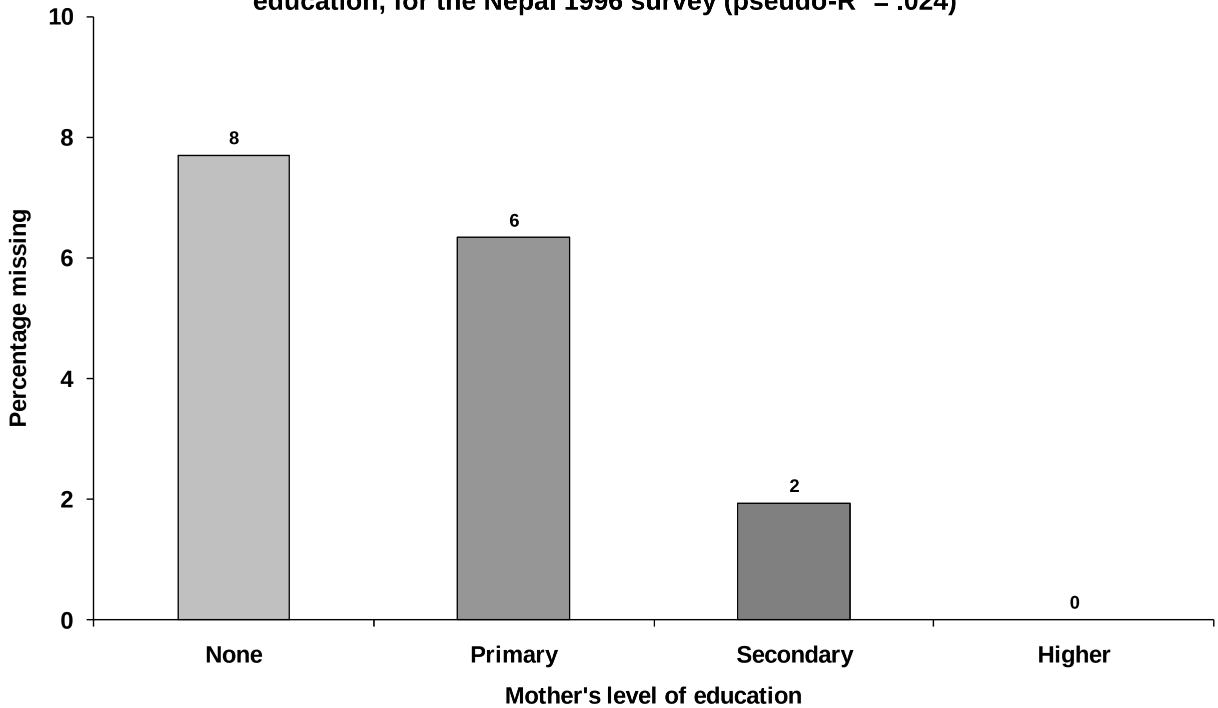


Figure 2.6 Percentage of cases missing date of first antenatal visit, by woman's level of education, for the nine surveys with at least 2 percent missing on this variable (pseudo-R^2 = .017)



In the second series of logit regressions, there are no substantively significant predictors of missing data on the number of antenatal visits. The third series of logit regressions focuses on Nepal 1996. This is the only survey with more than 2 percent missing on who provided antenatal care. This outcome is not strongly related to any of the covariates. Nepal 1996 also stands out for the highest percent missing the date when antenatal care was first provided, 6.5 percent. This type of missing is strongly related to both place of residence and level of education. Figure 2.7 shows that the level of missing decreases monotonically as the mother's level of education increases.

Figure 2.7 Percentage of cases missing date of first antenatal visit, by woman's level of education, for the Nepal 1996 survey (pseudo-R^2 = .024)



We will look in more detail at the three highest levels of missing in Table 2.1—Kazakhstan 1999, Nigeria 1999, and Zimbabwe 1999—on the question of how many antenatal visits occurred. The full distribution of the number of visits (m14, unweighted) in these three surveys is given in Table 2.2.

It is clear from Table 2.2 that most of the missing responses are due to "don't know," code 98, which we are combining with "not stated," code 99. There are other somewhat irregular features in these three distributions, such as a long tail on the distribution for Kazakhstan 1999, heaping at numbers with final digits 8 and 0 in Nigeria 1999, and heaping at 12 in all three surveys.

In the Kazakhstan 1999 survey, none of the six covariates used in this analysis is significantly related to the probability of missing on this variable. The pattern of missing appears to be random, at least with respect to these covariates, and the observed distribution of number of visits would not change under any reasonable imputation procedure.

In the Nigeria 1999 and Zimbabwe 1999 surveys, there is a systematic relationship between level of education and the probability of being missing, but it is the opposite of what one might at first expect: the probability of missing *increases* with level of education. At the same time, level of education is strongly and positively related to the number of antenatal visits (for those women reporting a number).

Table 2.2 Distribution of the number of antenatal visits for the three DHS surveys with the highest levels of missing responses

Number of antenatal visits for pregnancy	Kazakhstan 1999	Nigeria 1999	Zimbabwe 1999
0	77	945	120
1	48	58	28
2	56	123	100
3	39	163	317
4	35	226	393
5	47	201	382
6	72	238	446
7	50	136	180
8	43	205	129
9	40	61	68
10	77	183	74
11	11	24	9
12	68	98	35
13	9	11	10
14	23	22	5
15	38	38	7
16	32	39	9
17	7	7	1
18	34	10	5
19	2	4	0
20	41	58	8
21	10	0	0
22	10	0	0
23	2	0	0
24	14	0	0
25	1	0	0
26	5	0	0
27	1	0	0
28	2	0	0
30	7	0	0
32	3	0	0
36	2	0	0
39	2	0	0
40	2	0	0
44	1	0	0
49	1	0	0
56	1	0	0
Don’t know	111	279	322
Not stated	4	79	4
Total	**1,028**	**3,206**	**2,652**

Table 2.3 shows how both the number of antenatal visits and the percent missing tend to increase strongly by level of education in Nigeria 1999 and Zimbabwe 1999. It cannot be proven with these data, but we speculate that the reason for the increase in missing, by education, which is mostly due to the "don't know" response, is that women who made many visits were unable to give an exact count. This uncertainty, when the number of visits is large, is probably also behind the increasing heaping of responses at large values.

Table 2.3 Weighted number of antenatal visits and percentage of mothers who are missing on the number of antenatal visits, by level of education, in Nigeria 1999 and Zimbabwe 1999

	Nigeria 1999		Zimbabwe 1999	
Level of education	Mean visits	Percent missing	Mean visits	Percent missing
None	2.32	9.0	4.61	7.1
Primary	6.25	12.5	4.82	12.5
Secondary	7.99	12.9	5.66	15.1
Higher	9.13	19.8	7.94	27.4
Total	**4.85**	**11.2**	**5.27**	**13.7**

Interviewers should make it clear to the respondent that an approximate or estimated number of visits is better than "don't know" or "not stated." The difficulty of giving an exact number when that number is large was presumably an issue in other surveys and was probably handled during the training of the interviewers, but less effectively in these two surveys. We note that the Nigeria 2003 survey did not encounter this problem.

2.2 Birth Information

DHS obtains detailed information about who was present and assisted at the birth of an index child (m3) and where the birth occurred (m15). This section describes the cases that are missing on these two items.

Figures 2.8 and 2.9 give the distributions of the percentages missing on who assisted at the birth and where the birth took place. The distributions are very similar in shape and indicate extremely low levels of missing responses. Overall, giving equal weight to each survey, only 0.2 percent of responses were missing on either question.

Figure 2.8 Distribution of the percentages of cases that are missing on who assisted at the child's birth, all DHS surveys 1993-2003

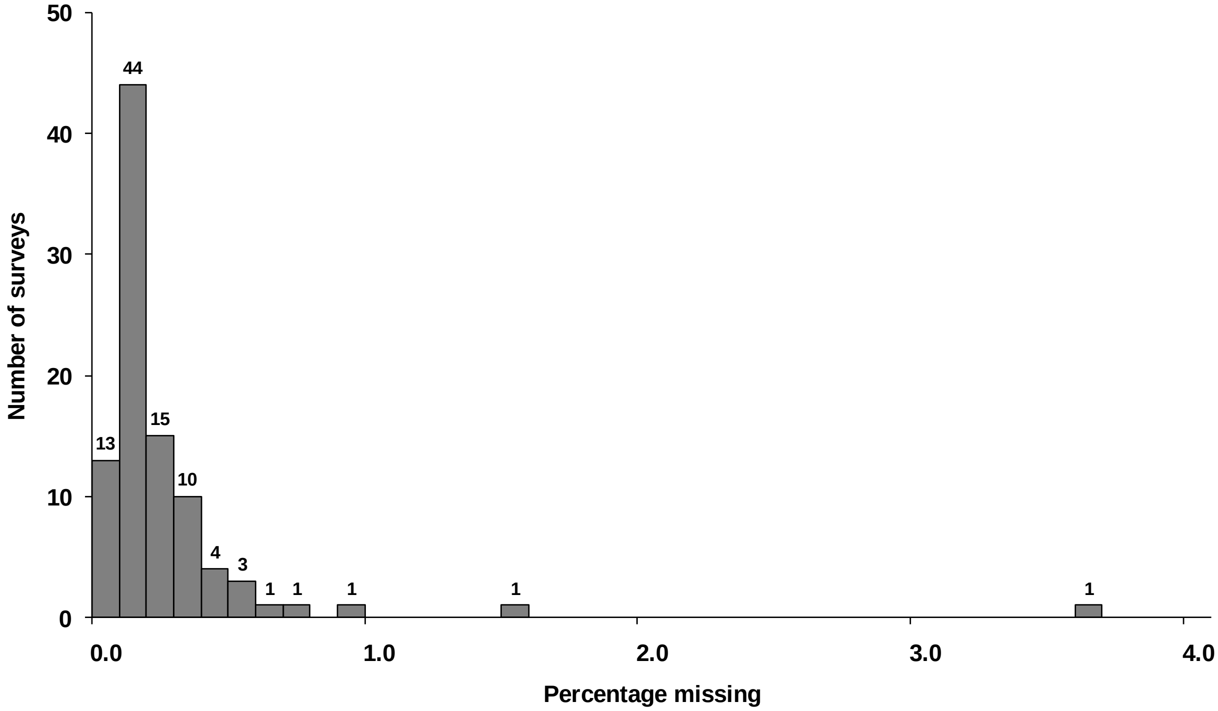
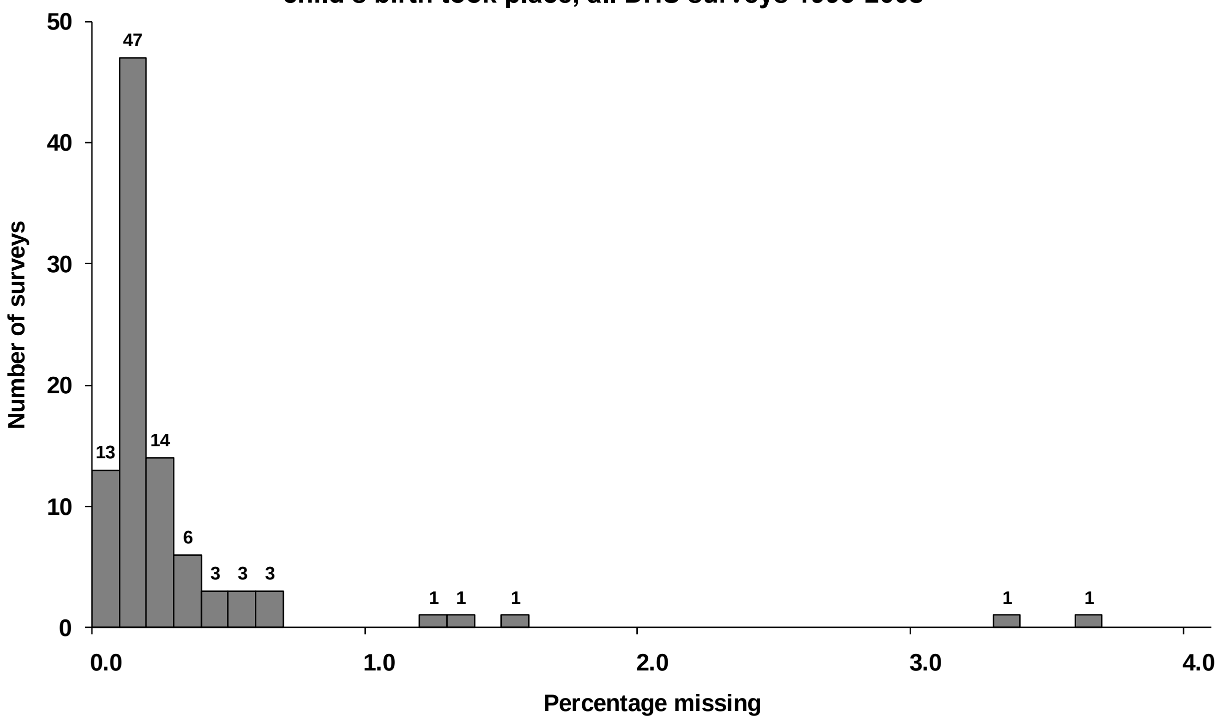


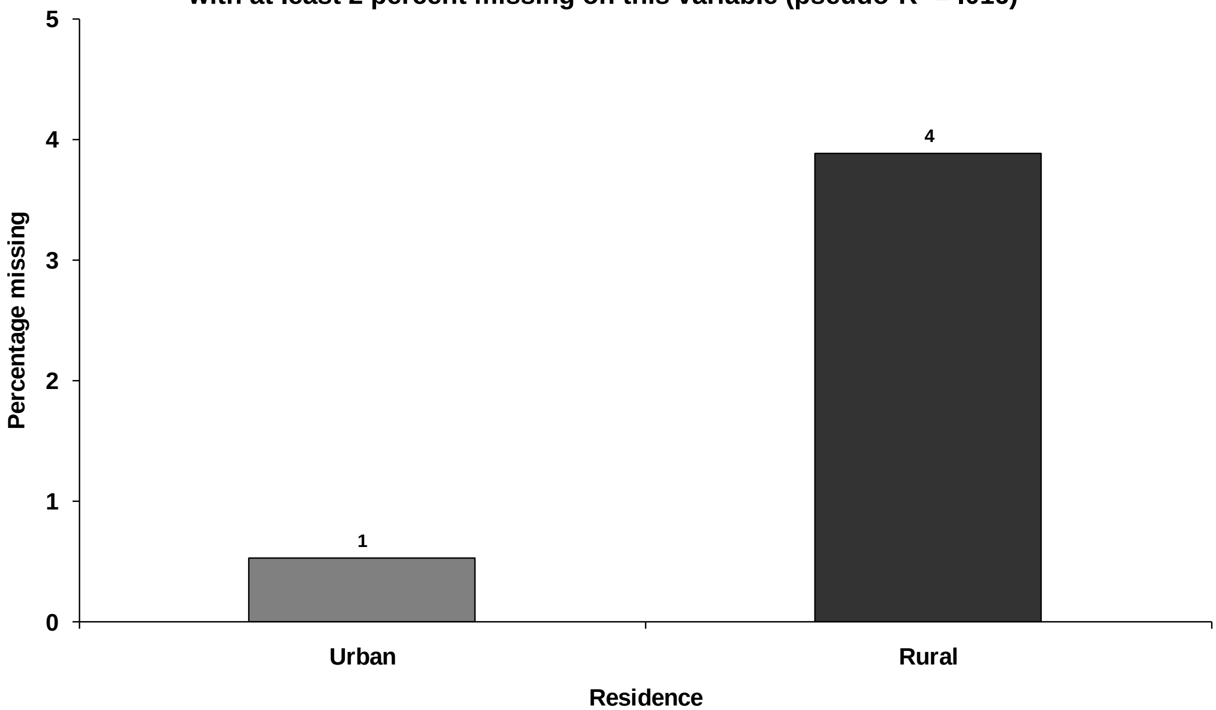
Figure 2.9 Distribution of the percentages of cases that are missing on where the child's birth took place, all DHS surveys 1993-2003



Only three surveys exceeded the arbitrary threshold of 2 percent missing. The Uganda 2000/01 survey had 3.5 percent missing[3] on who attended at the delivery, and the Tanzania 1996 and Zimbabwe 1999 surveys had 3.3 percent and 3.6 percent, respectively, missing on where the delivery took place.

In the Uganda 2000/01 survey, the probability of being missing on the type of person who assisted at the delivery is strongly associated with three of the covariates, and for each covariate the relationship is in the direction we would expect. The patterns are shown in Figures 2.10 through 2.12. The risk of missing is higher in rural areas, for an older child, and for women with more index children. It is weakly (pseudo–R^2 > .01) but not significantly (with a .01 standard) associated with education, also in the expected direction. It is *not* significantly associated with the age of the woman or the number of interviewer visits. The strongest association, by far, is with the age of the child. For the 1,389 children in this survey who had not yet reached their first birthday, not a single one was missing on this variable. For the 2,372 children who were age 3 or 4, the missing rate was about 7 percent (virtually the same for age 3 and age 4). Similarly, women with more than one index child were much more likely to be missing. This pattern clearly implies recall error, but it is interesting that Uganda 2000/01 had a very low level of missing (only 0.5 percent) on the *place* of delivery, which was apparently easier to recall.

Figure 2.10 Percentage of cases missing who assisted at the birth, by type of place of residence, for the only survey (Uganda 2000/01) with at least 2 percent missing on this variable (pseudo-R^2 = .016)



[3] A variable such as who attended at the delivery (m3) is coded by DHS as a set of binary variables (m3a, m3b, etc.) for each possible option (doctor, nurse/midwife, etc.). As a rule, if the question is not answered but should have been, all of the options will be coded as missing (code 9). In the Uganda 2000/01 survey this rule was violated; of the eight possible options, two have 24 cases with code 9 and the other six have 251 cases with code 9 (unweighted numbers of cases). The main report on this survey, using the 24 cases, reported a level of 0.4 percent missing on this variable. Here we use the 251 cases, giving 3.5 percent missing. The fact that 24 cases and 251 cases are separately stated to be the number of missing cases for this survey and variable may indicate a failure of edit checks during data processing.

Figure 2.11 Percentage of cases missing who assisted at the birth, by child's age, for the only survey (Uganda 2000/01) with at least 2 percent missing on this variable (pseudo-R^2 = .118)

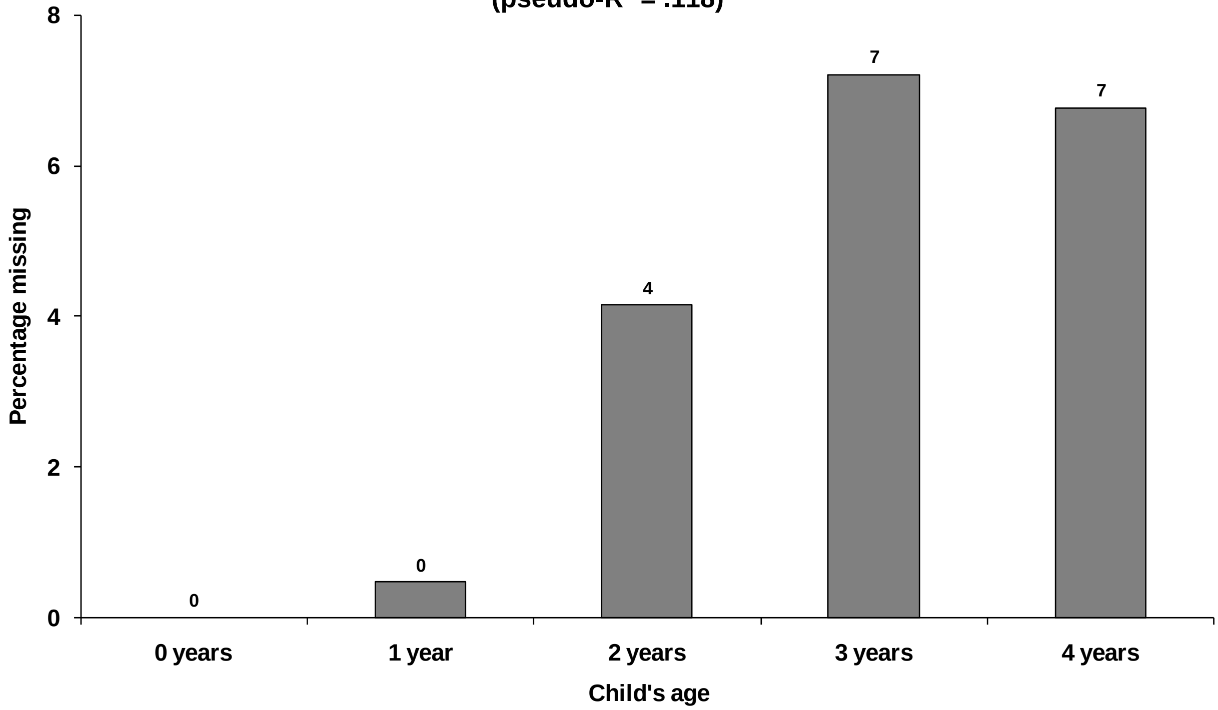
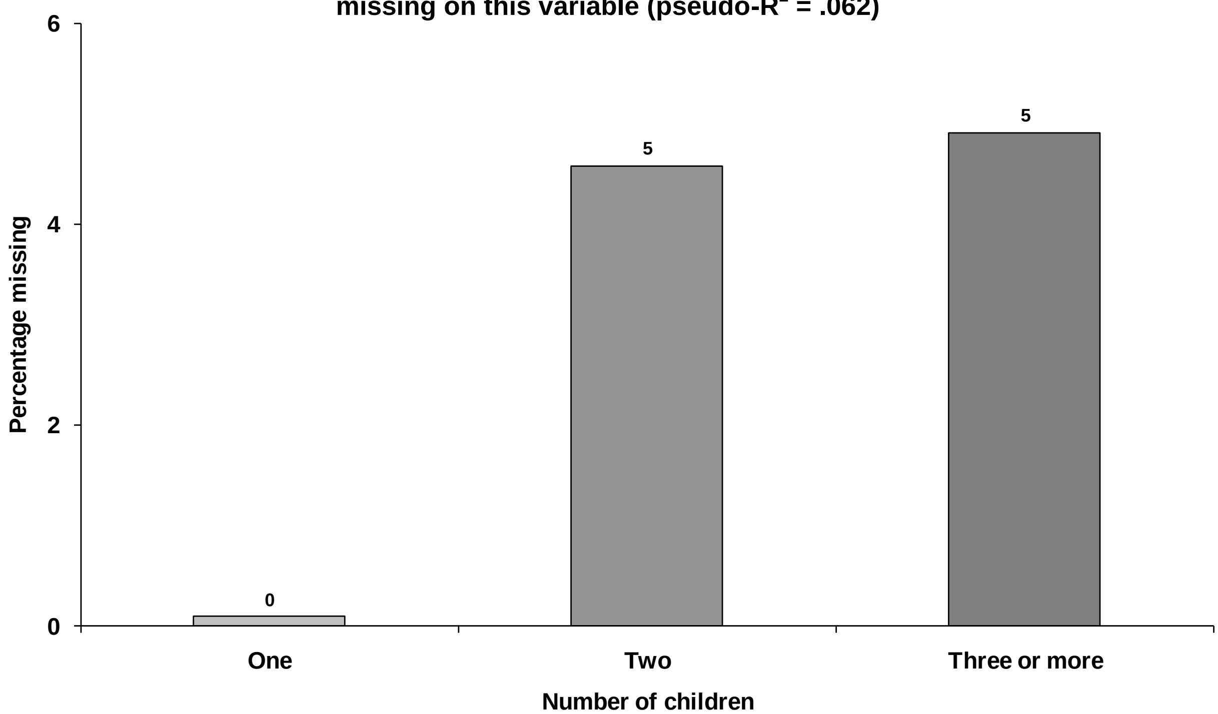


Figure 2.12 Percentage of cases missing who assisted at the birth, by number of children in the window, for the only survey (Uganda 2000/01) with at least 2 percent missing on this variable (pseudo-R^2 = .062)



Potential covariates of the probability of missing on place of birth were examined for a pooling of the two surveys that exceeded the 2 percent threshold, Tanzania 1996 and Zimbabwe 1999, giving equal weight to each. The missing response is again significantly related to four covariates, but with some differences from the pattern just discussed for Uganda 2000/01. It is higher in rural areas, higher for older women, higher for women with low education, and higher for women with more index children. The highest levels of missing were for women with no education and for women age 40-49 years. These patterns are shown in Figures 2.13 through 2.16. We have not estimated the bias in the observed distribution of provider and place that would be induced by this pattern of non-random missing, but on the basis of estimates for higher levels of missing, given later in this report, we believe any bias would be negligible.

Figure 2.13 Percentage of cases missing the place of birth, by type of place of residence, for the two surveys (Tanzania 1996 and Zimbabwe 1999, pooled) with at least 2 percent missing on this variable (pseudo-R^2 = .010)

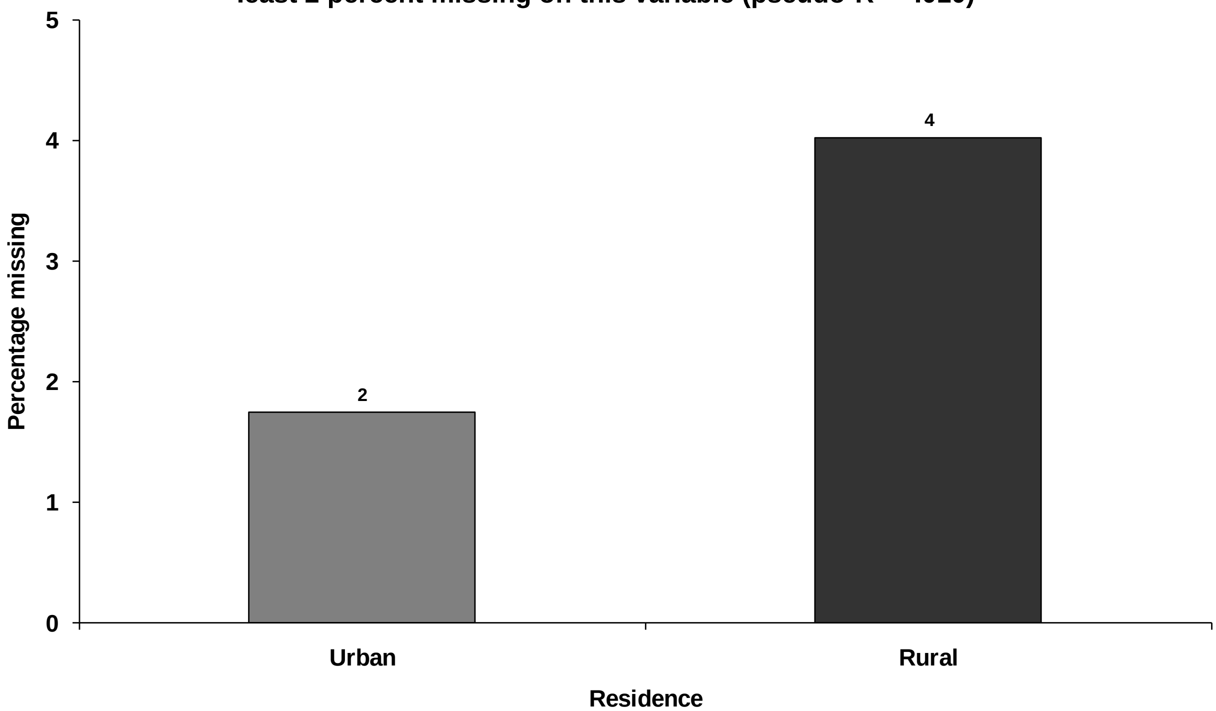


Figure 2.14 Percentage of cases missing the place of birth, by mother's age, for the two surveys (Tanzania 1996 and Zimbabwe 1999, pooled) with at least 2 percent missing on this variable (pseudo-R^2 = .025)

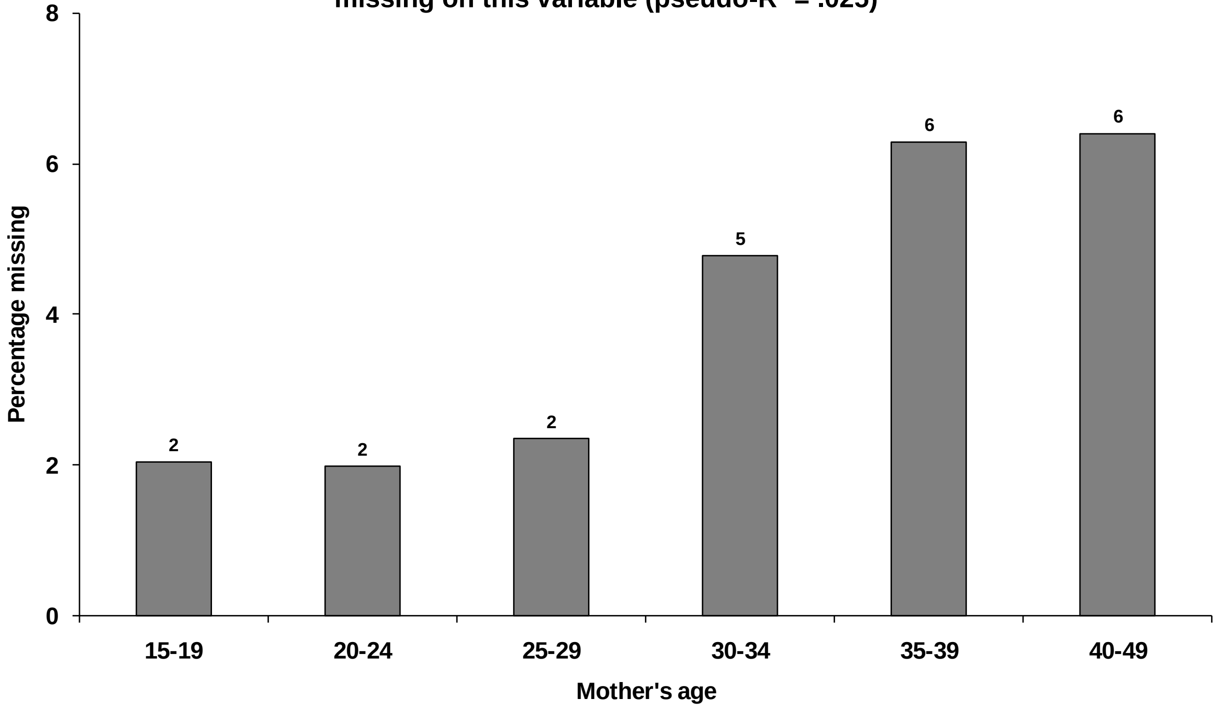


Figure 2.15 Percentage of cases missing the place of birth, by mother's level of education, for the two surveys (Tanzania 1996 and Zimbabwe 1999, pooled) with at least 2 percent missing on this variable (pseudo-R^2 = .020)

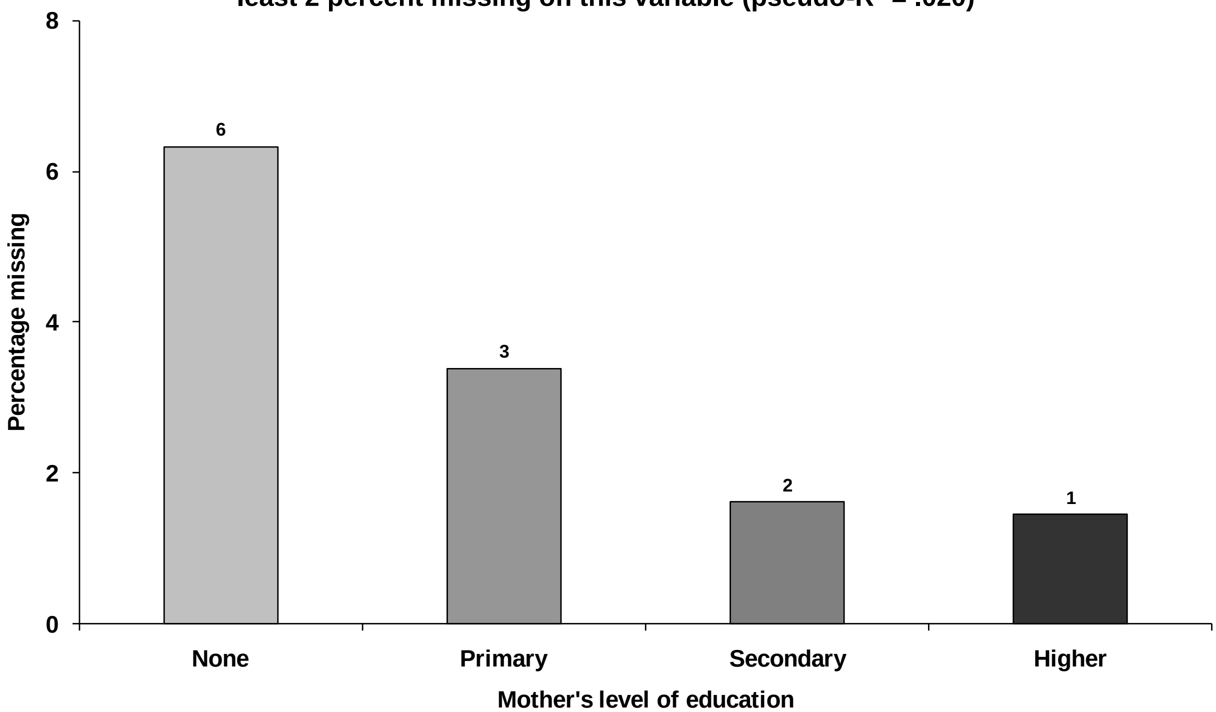
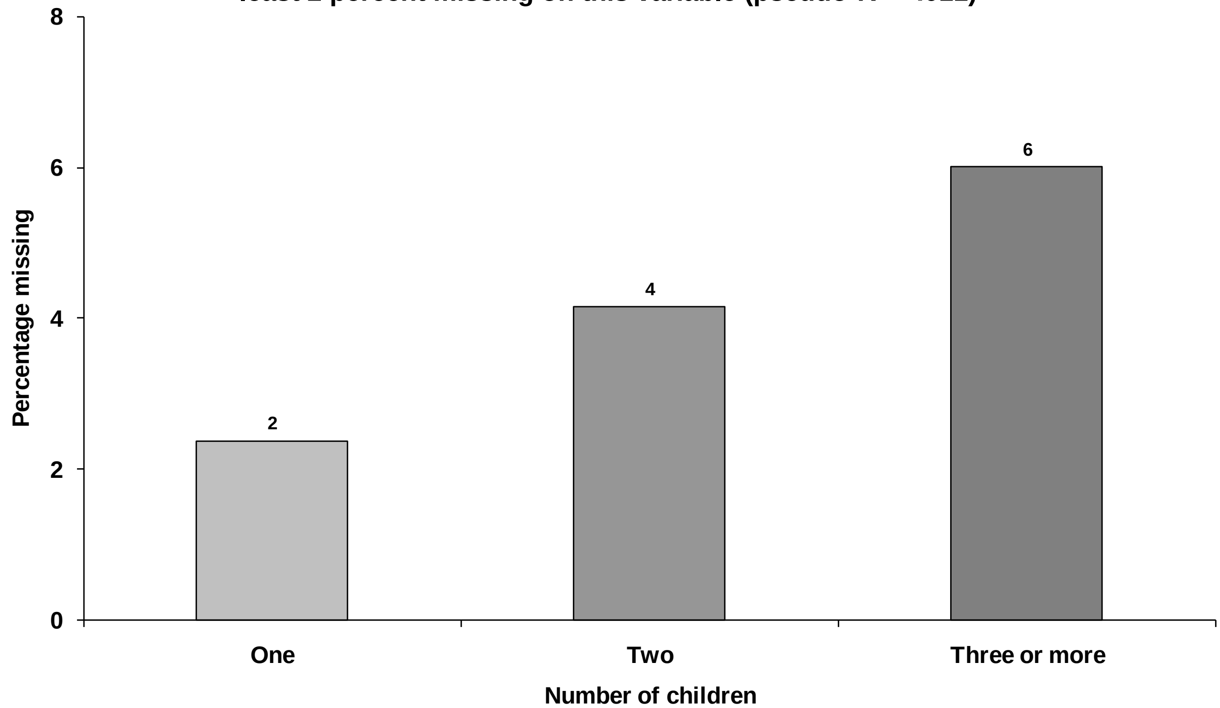


Figure 2.16 Percentage of cases missing the place of birth, by number of children in the window, for the two surveys (Tanzania 1996 and Zimbabwe 1999, pooled) with at least 2 percent missing on this variable (pseudo-R^2 = .011)



2.3 Postpartum Amenorrhea and Sexual Abstinence

This section describes the cases that are missing on questions about the length of postpartum amenorrhea and sexual abstinence. First, it is necessary to clarify the difference between two standard DHS variables that give these lengths.

For breastfeeding, amenorrhea, and abstinence, DHS includes pairs of closely related variables with the labels "duration" and "months." Thus, m4, m6, and m8 are "duration of breastfeeding," "duration of amenorrhea," and "duration of abstinence," respectively. The variables m5, m7, and m9 are "months of breastfeeding," "months of amenorrhea," and "months of abstinence," respectively. This section will examine m6 through m9, and Section 3.2 will look at m4 and m5.

All of m4 through m9 are measured in months and include codes 97, "inconsistent;" 98, "don't know;" and 99, "not stated." m6 comes directly from the questionnaire and has one additional non-numeric category: 96, "period not returned." (The equivalent code for censored observations on m8 is 96, "still abstaining." A different code is used for censored observations on m4: 95, "still breastfeeding;" m4 also includes 94, "never breastfed;" and 96, "breastfed until child died.")

We will review how m7, "months of amenorrhea," is constructed from m6, "duration of amenorrhea." Similar comments could be made about the links between m4 and m5 and between m8 and m9. m7 is equal to m6, with two principal exceptions. First, if an observation on m6 is censored, i.e., has code 96, then m7 is given the value of the open interval in months, namely v008-b3. Second, if m6 gives a stated duration of amenorrhea that is longer than the open interval, then m7 is recoded as 97, "inconsistent." (There appear to be a few exceptions to the second rule, 18 cases in all, for which an excessively long interval on m6 was replaced with a different interval on m7, instead of being recoded as 97. We suspect that these exceptions could be traced to illegible values or data entry errors that were resolved by checking against the questionnaires, and we do not question them.)

The first part of this section deals with the incidence of invalid or missing responses using m7 and m9. Codes 97, “inconsistent;” 98, “don’t know;” and 99, “not stated” are all counted as “missing.”

Overall, giving equal weight to every survey, 3.1 percent of cases were missing a valid duration of amenorrhea. Of these, again giving equal weight to each survey, 47.7 percent were inconsistent (almost always because the stated duration was longer than the length of the open interval), 34.8 percent were “don’t know,” and 17.5 percent were completely missing. For duration of abstinence, 3.2 percent were missing a valid response, which breaks down into 37.1 percent inconsistent, 45.7 percent “don’t know,” and 17.2 percent completely missing. We will first look at the distribution of missing responses across all 81 surveys, then estimate any bias arising from non-randomness in the missing, and then examine the pattern of heaping.

Figure 2.17 gives the distribution of the percentage missing a valid code (i.e., with codes 97, 98, or 99) for the months of postpartum amenorrhea (m7). One survey, Cote d’Ivoire 1994, is omitted because it was an outlier, with 58.9 percent of the durations of amenorrhea stated to be invalid. Almost all of the allegedly invalid responses in this survey are coded “inconsistent.” Because we do not have access to the number of months stated in the interview, no effort will be made to account for the apparently serious coding problems with m7 in this survey.

Figure 2.17 shows more dispersion than was seen for earlier measures. The percentage missing exceeds 8 percent (barely) for one survey—South Africa 1998; exceeds 6 percent for an additional three surveys—

Figure 2.17 Distribution of the percentages of cases that are missing a valid duration of postpartum amenorrhea, all DHS surveys 1993-2003, except Cote d’Ivoire 1994

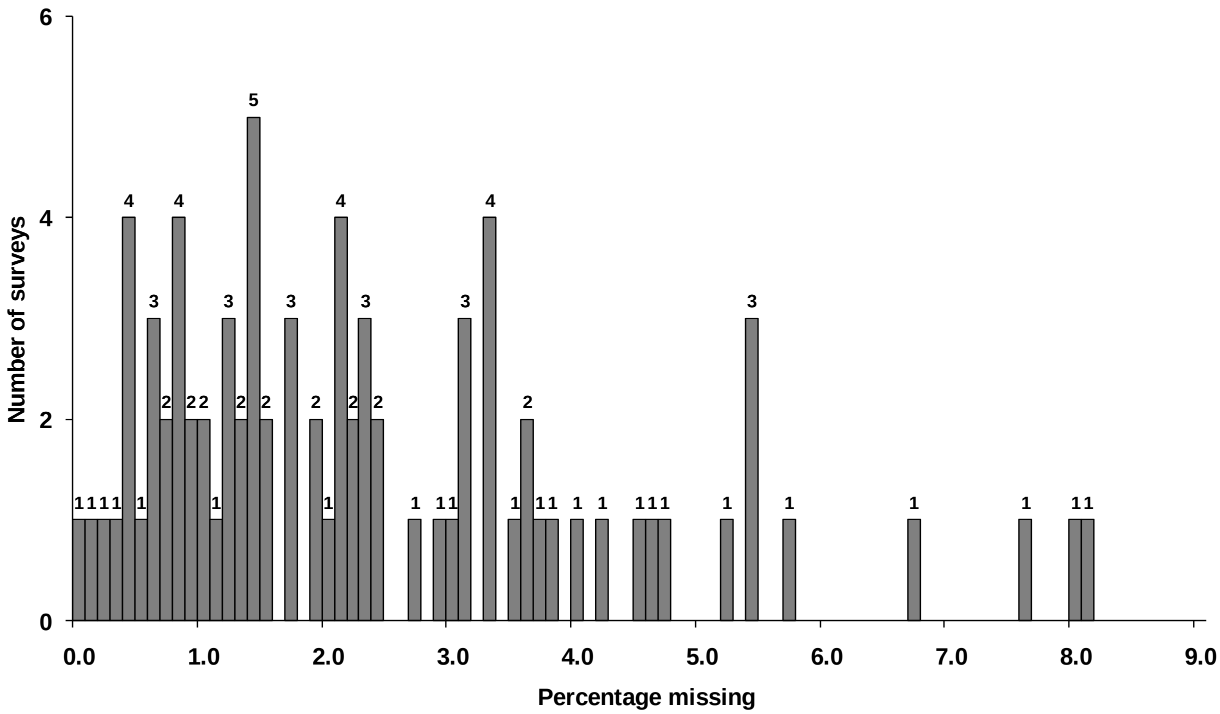


Table 2.4 Surveys in which 2 percent or more of cases are missing a valid duration of postpartum amenorrhea, after removing cases with "period not returned," all DHS surveys 1993-2003 that included these questions, except Cote d'Ivoire 1994

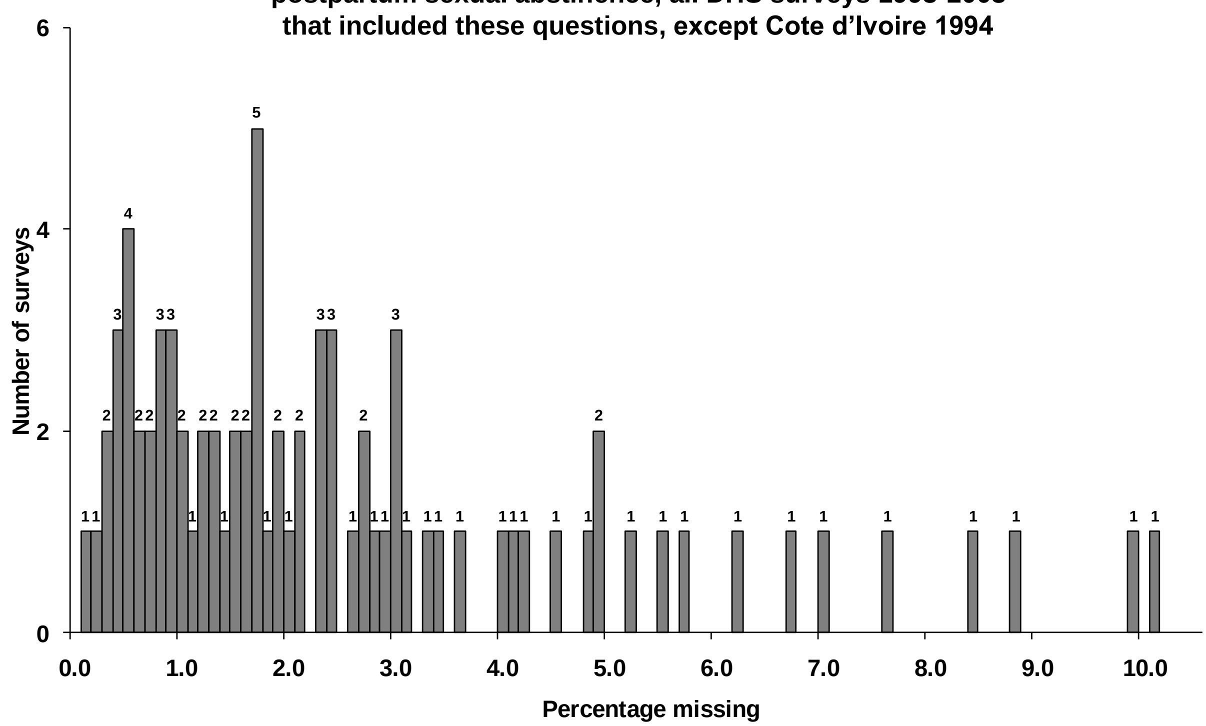
Survey	Percent missing	Estimated difference[1]	Estimated bias[2]
Benin 2001	2.65	1.61	0.0
Bolivia 1998	3.72	1.31	0.0
Bolivia 2003/04	3.04	1.07	0.0
Brazil 1996	3.68	0.09	0.0
Burkina Faso 1998/99	5.34	1.80	0.1
Burkina Faso 2003	3.28	1.82	0.0
Cameroon 1998	2.12	0.97	0.0
Chad 1996/97	3.29	1.01	0.0
Comoros 1996	2.94	0.35	0.0
Cote d'Ivoire 1998/99	2.10	0.86	0.0
Dominican Republic 1999	2.34	0.51	0.0
Ethiopia 2000	3.99	1.27	0.0
Gabon 2000	2.30	0.45	0.0
Ghana 2003	2.03	1.19	0.0
Guinea 1999	5.40	0.67	0.0
Haiti 1994/95	3.40	0.25	0.0
Haiti 2000	2.26	1.03	0.0
Kenya 1993	4.43	0.61	0.0
Kenya 2003	2.15	1.31	0.0
Madagascar 1997	2.09	0.82	0.0
Mali 1995/96	4.64	2.79	0.1
Mali 2001	7.54	1.21	0.0
Mozambique 1997	5.40	2.35	0.1
Mozambique 2003	3.51	0.90	0.0
Namibia 2000	5.69	0.61	0.0
Nicaragua 1997/98	3.06	0.99	0.0
Nicaragua 2001	2.32	0.15	0.0
Niger 1998	2.04	2.58	0.0
Nigeria 1999	7.94	0.58	0.0
Nigeria 2003	6.61	1.45	0.1
Rwanda 2000	3.20	1.75	0.0
Senegal 1997	5.13	1.17	0.0
South Africa 1998	8.03	0.04	0.0
Tanzania 1996	4.52	0.99	0.0
Tanzania 1999	4.17	1.08	0.0
Uganda 1995	3.56	0.92	0.0
Uganda 2000/01	3.25	1.39	0.0
Zambia 1996	2.82	0.74	0.0
Zambia 2001/02	2.30	1.84	0.0
Zimbabwe 1999	3.06	0.65	0.0

[1] The "estimated difference" is the estimated mean duration for the missing cases minus the observed mean duration.
[2] The "estimated bias" is the adjusted mean duration minus the observed mean duration.

Mali 2001 and Nigeria 1998 and 2003; and exceeds 4 percent for an additional nine surveys—Burkina Faso 1998/99, Guinea 1999, Kenya 1993, Mali 1995/96, Mozambique 1997, Namibia 2000, Senegal 1997, and Tanzania 1996 and 1999. Table 2.4 lists the 40 surveys with at least 2 percent missing. This table also gives two measures of the amount of bias attributable to a non-random pattern of missing cases.[4] The second column of the table gives an estimate of the amount by which the mean duration for the missing cases exceeds the mean duration for the non-missing cases. The differences range from 0.04 months for South Africa 1998 to 2.79 months for Mali 1995/96. The last column of the table estimates the amount by which duration for the combined missing and non-missing cases would exceed the mean for the non-missing cases. This difference is described as the "estimated bias." Because the percent missing and the estimated difference between the missing and non-missing means are generally small, the estimated bias is negligible for all surveys.

Turning to the duration of postpartum sexual abstinence, Figure 2.18 and Table 2.5 give the distribution of the percentage missing and estimated bias in this duration. The Cote d'Ivoire 1994 survey, which was omitted as an outlier from the preceding figure and table, is again omitted because the level of missing is 60.0 percent. The distribution is essentially similar to that for amenorrhea (the correlation between the two levels of missing, across all 81 surveys, is .56). One survey exceeds 10 percent missing—Guinea 1999; three more exceed 8 percent—Burkina Faso 1998/99, Mozambique 1997, and Nigeria 1999; four more exceed 6 percent—Bolivia 1998, Mali 2001, Namibia 2000, and Senegal 1997; and nine more exceed 4 percent. There are some differences, but the same surveys tend to appear high on both lists, and these are primarily the surveys from sub-Saharan Africa. The largest bias appears for Burkina Faso 1998/99 and Guinea 1999, but the net effect of non-randomness in the missing or invalid responses is again negligible.

Figure 2.18 Distribution of the percentages of cases that are missing the duration of postpartum sexual abstinence, all DHS surveys 1993-2003 that included these questions, except Cote d'Ivoire 1994



[4] These means are not adjusted for censoring, an issue that will be discussed explicitly at the end of this section.

Table 2.5 Surveys in which 2 percent or more of cases are missing a valid duration of postpartum abstinence, after removing cases that are still abstaining, all DHS surveys 1993-2003 that included these questions, except Cote d'Ivoire 1994

Survey	Percent missing	Estimated difference[1]	Estimated bias[2]
Benin 2001	4.83	1.19	0.06
Bolivia 1994	3.04	0.27	0.01
Bolivia 1998	6.64	-0.07	0.00
Bolivia 2003/04	4.84	-0.08	0.00
Brazil 1996	3.34	0.12	0.00
Burkina Faso 1998/99	8.36	2.08	0.17
Burkina Faso 2003	5.12	1.02	0.05
Cameroon 1998	2.96	0.42	0.01
Chad 1996/97	4.12	0.48	0.02
Comoros 1996	4.07	-0.05	0.00
Cote d'Ivoire 1994	60.01	-0.77	-0.46
Gabon 2000	4.49	0.56	0.03
Ghana 1998	2.36	1.27	0.03
Ghana 2003	3.00	1.36	0.04
Guatemala 1995	2.07	0.11	0.00
Guatemala 1998/99	2.26	-0.11	0.00
Guinea 1999	10.05	4.10	0.41
Haiti 1994/95	2.39	0.19	0.00
Haiti 2000	2.74	-0.12	0.00
Indonesia 1994	2.66	0.39	0.01
Kenya 1993	2.63	-0.42	-0.01
Madagascar 1997	3.58	0.13	0.00
Mali 1995/96	3.94	0.62	0.02
Mali 2001	6.17	0.16	0.01
Mozambique 1997	8.72	1.73	0.15
Mozambique 2003	4.74	0.95	0.04
Namibia 2000	7.54	-0.55	-0.04
Nicaragua 1997/98	2.53	-0.19	0.00
Nigeria 1999	9.85	0.48	0.05
Nigeria 2003	5.41	0.19	0.01
Philippines 2003	2.04	-0.22	0.00
Senegal 1997	6.96	-0.24	-0.02
South Africa 1998	5.65	-0.54	-0.03
Tanzania 1996	2.86	0.06	0.00
Tanzania 1999	3.26	0.00	0.00
Togo 1998	2.95	1.85	0.05
Zambia 1996	2.25	0.04	0.00
Zambia 2001/02	2.25	0.37	0.01
Zimbabwe 1999	2.35	-0.15	0.00

[1] The "estimated difference" is the estimated mean duration for the missing cases minus the observed mean duration.
[2] The "estimated bias" is the adjusted mean duration minus the observed mean duration

Myers' Index will now be used to assess heaping, that is, the tendency for stated durations of amenorrhea and abstinence to be multiples of six months and especially 12 months. The index is calculated for m6 and m8, rather than m7 and m9, because for m7 and m9, censored cases give the length of the open interval rather than stated durations, and there is no reason to expect heaping for those calculated intervals. The usual Myers' Index applied to years of age, which is characteristically heaped at multiples of five years and especially ten years, has been modified for this purpose.

Heaping for months is typically greatest for multiples of six months and especially 12 months. The modified Myers' Index indicates a very high level of heaping for durations of both amenorrhea and abstinence. The value for amenorrhea ranges from 13.3 to 38.5, with an average of 25.1, implying that an average 25.1 percent of reported values would have to be shifted to another value in order to obtain a distribution with no heaping. The distribution of the index for amenorrhea, across all 81 surveys, is given in Figure 2.19. It is relatively symmetric and is even similar to a normal distribution. Seven surveys, from five countries, have an index of 35 or more: Burkina Faso, 1998/99 and 2003; Bangladesh 1999/2000 and 1996/97; Ethiopia 2000; Guinea 1999; and Tanzania 1996.

For abstinence, Myers' Index ranges from 15.5 to 72.7, with a mean of 35.0. The generally higher level of Myers' Index for abstinence than for amenorrhea can be partially traced to the fact that extended postpartum abstinence is culturally, rather than biologically, determined. In many countries the typical interval is only one or two months long. This lack of dispersion in duration causes Myers' Index to be misleading. Therefore, Figure 2.20 is limited to the 56 surveys in which mean duration is at least three months. For those surveys, the range in the index is from 15.5 to 44.4, with a mean of 29.9. Fourteen surveys from nine countries have an index of 35 or more: Burkina Faso 1998/99, Colombia 1995 and 2000, Guinea 1999, Haiti 1994/95 and 2000, Indonesia 1997, Nicaragua 1997/98 and 2001, Peru 2000, Philippines 1998 and 2003, and Vietnam 1997 and 2000. These levels of heaping are high but not unexpected. Durations of amenorrhea and abstinence are simply estimated by the respondent.

The two surveys with the most serious heaping for both of these durations combined are Burkina Faso 1998/99 and Guinea 1999. Their unweighted distributions of reported durations of amenorrhea and abstinence are given in Table 2.6 There is obvious and increasing heaping at durations 6, 12, 18, 24, 30, and 36 months.

It may be helpful to investigate the potential impact of heaping on inferences about mean or median durations. Actual analyses of durations of amenorrhea or abstinence, as well as breastfeeding, would require the inclusion of the censored cases, i.e., those women who are still amenorrheic, abstaining, or breastfeeding. There are two typical ways to do this kind of analysis. The first uses hazard or failure time modeling; the other can be described as current status modeling. We shall illustrate these methods with the Burkina Faso 1998/99 data on duration of amenorrhea, m6. Data out to 27 months are used because of the pronounced heaping at 24 months, which is probably partially due to cases pulled down from months above 24. Cases with code 97, "inconsistent," on m7, are dropped.

In failure time modeling, the stated durations are corrected for censoring. The conditional probabilities of exiting the original status are calculated for every month, and these are combined into a survival function. The irregular line in Figure 2.21 shows the survival function for this example, produced by the "stcox" routine in Stata. The irregularities, most pronounced for months 12 and 24, are induced by the heaping of stated durations. The smooth line in this figure is the best-fitting logistic function. The logistic is not the optimal function for fitting the survival function but is the easiest one to use; several alternatives are possible, including splines. This fit uses weights proportional to the number of women at each stated duration, resulting in a worse fit at the highest durations because few respondents are found at these durations.

Figure 2.19 Myers' Index for duration of postpartum amenorrhea (m6), all DHS surveys 1993-2003

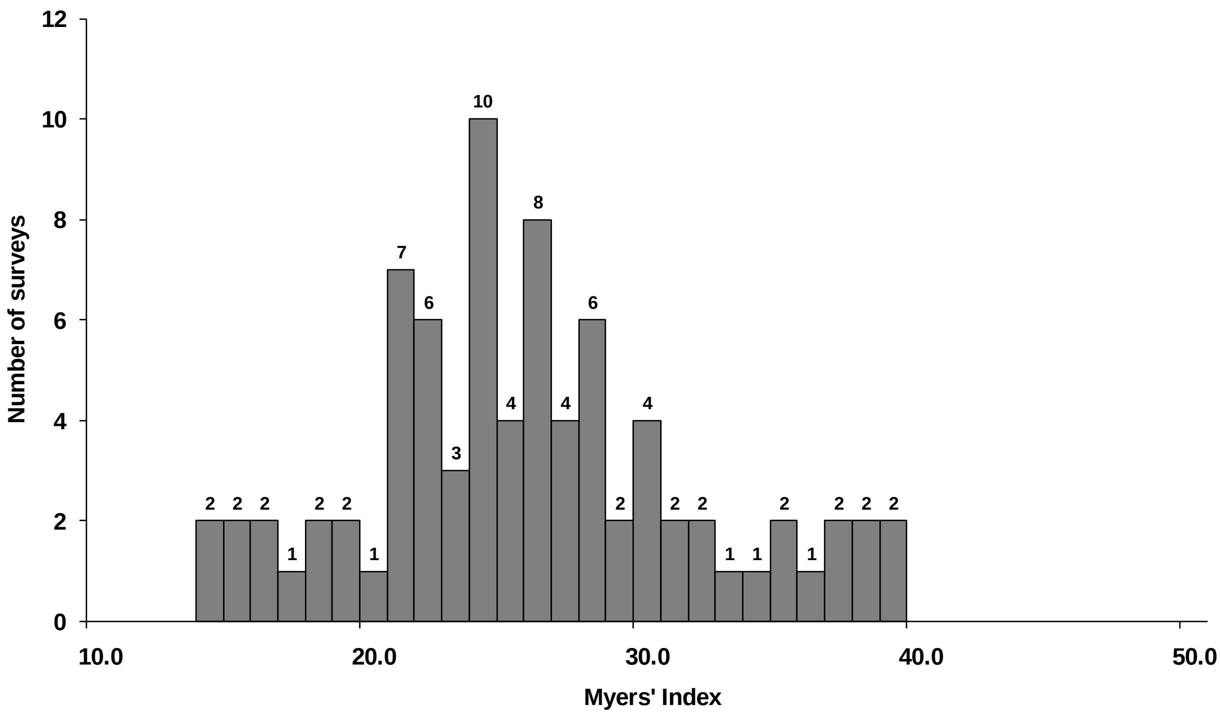


Figure 2.20 Myers' Index for duration of postpartum sexual abstinence (m8), all DHS surveys 1993-2003 in which the mean duration is at least three months

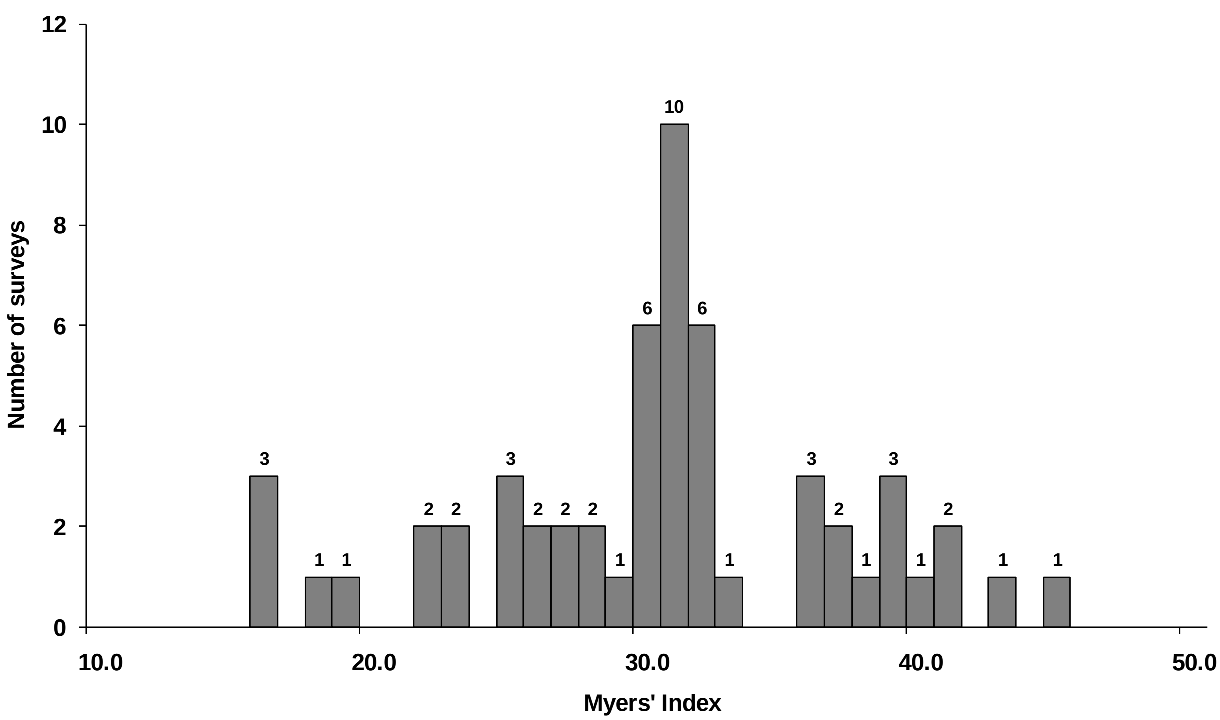
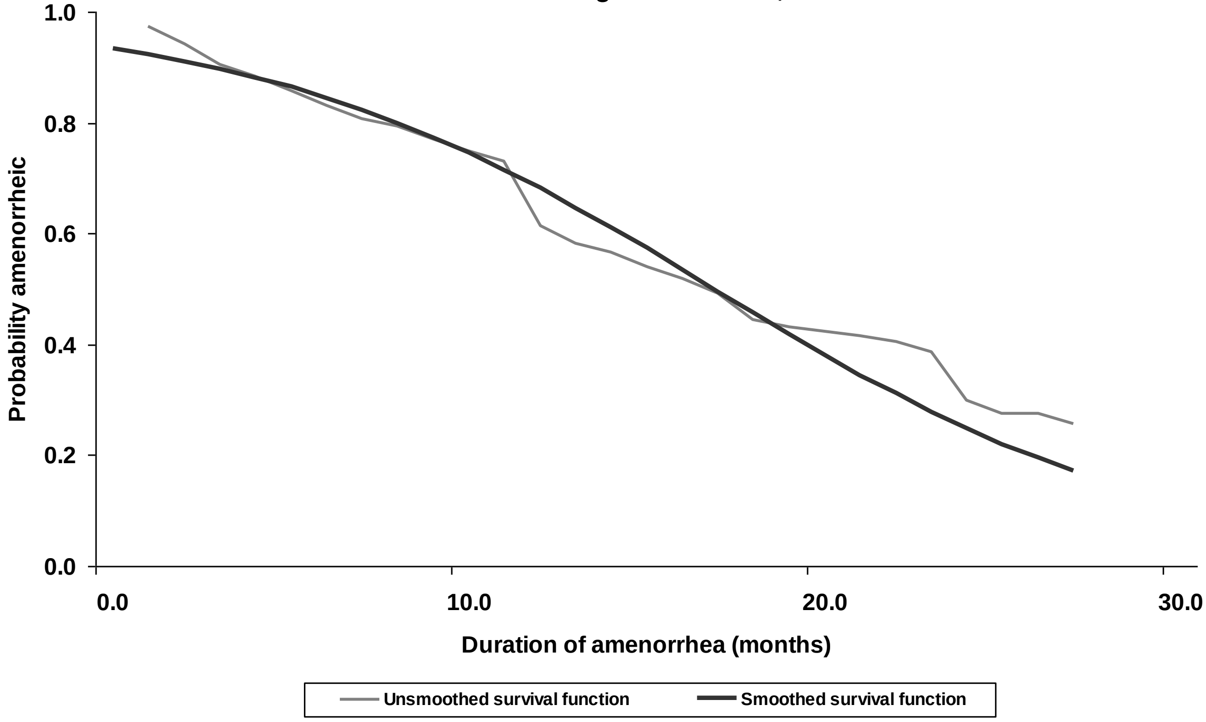


Table 2.6 Frequency distributions (unweighted) of the reported durations of postpartum amenorrhea (m6) and sexual abstinence (m8) in Burkina Faso 1998/99 and Guinea 1999, the two surveys with the highest levels of heaping on these variables

	Burkina Faso 1998/99		Guinea 1999	
Duration in months	Amenorrhea	Abstinence	Amenorrhea	Abstinence
0	7	24	8	2
1	143	83	274	34
2	170	194	266	40
3	199	217	251	39
4	136	164	153	36
5	123	175	119	12
6	142	172	178	45
7	98	96	132	26
8	73	93	123	36
9	99	72	101	25
10	104	81	115	27
11	69	36	28	7
12	602	477	703	191
13	127	68	64	19
14	83	54	76	41
15	89	67	54	33
16	43	35	37	44
17	33	17	48	42
18	149	122	111	155
19	28	25	28	53
20	27	32	56	93
21	10	11	16	24
22	20	14	13	31
23	18	11	8	29
24	467	444	428	932
25	26	17	20	80
26	15	44	20	128
27	13	19	14	75
28	13	15	15	85
29	7	8	2	47
30	17	37	16	129
31	0	1	5	19
32	2	6	4	33
33	0	0	3	8
34	1	0	2	4
35	0	1	0	1
36	63	71	26	92
37	0	0	1	3
38	0	0	0	0
39	1	2	0	0
40	0	1	1	4
41	1	0	0	1
42	0	0	0	1
43	0	0	0	0
44	0	0	0	0
45	0	0	0	0
46	0	0	0	0
47	0	0	0	0
48	0	1	1	2
Censored	1,714	1,800	1,412	2,199
Don't know	132	255	25	18
Not stated	12	14	83	94
Total	**5,076**	**5,076**	**5,040**	**5,039**

Figure 2.21 The survival function for duration of amenorrhea, using hazard modeling, unsmoothed and smoothed with a logistic function, Burkina Faso 1998/99



In Figure 2.22, the irregular line shows the survival function estimated simply from the proportion of women who are still amenorrheic for each length of the open interval, from 0 to 27 months. The irregular line in Figure 2.21 was required to decline monotonically, but in Figure 2.22 this is not required. The smooth line in this figure is again the best-fitting logistic function, estimated with weights that are proportional to the number of women at each length of the open interval. These numbers are nearly uniform from month to month, so the smooth line fits better at the upper end in Figure 2.22 than in Figure 2.21.

To facilitate comparison, the smoothed lines from Figures 2.21 and 2.22 are both given in Figure 2.23. The two lines are very similar, but there is a displacement, such that the median duration from the failure time model is estimated at 16.9 months, versus 17.7 months for the current status model. We hesitate to over-generalize from this example, but the two estimates are within a month of each other in a context with very prolonged duration of amenorrhea due to prolonged breastfeeding. It is surprising that the agreement is this good when the data on durations are so questionable. There seems to be little basis for choosing one method over the other.

Comparisons similar to these have been made, but are not shown here, using surveys for which the data show less heaping. In general, regardless of data quality, the two methods seem to agree but with a longer median duration estimated by the current status method than with the failure time method. Other issues that could affect this comparison, such as whether a stated duration should be interpreted as rounded months rather than completed months, have not been taken into account but could be investigated. The important point here is that both methods involve a smoothing of the data that overcomes the severe heaping at multiples of six months. Of course, if the heaping involves a bias, then the smoothing cannot overcome that bias.

Figure 2.22 The survival function for duration of amenorrhea, using current status modeling, unsmoothed and smoothed with a logistic function, Burkina Faso 1998/99

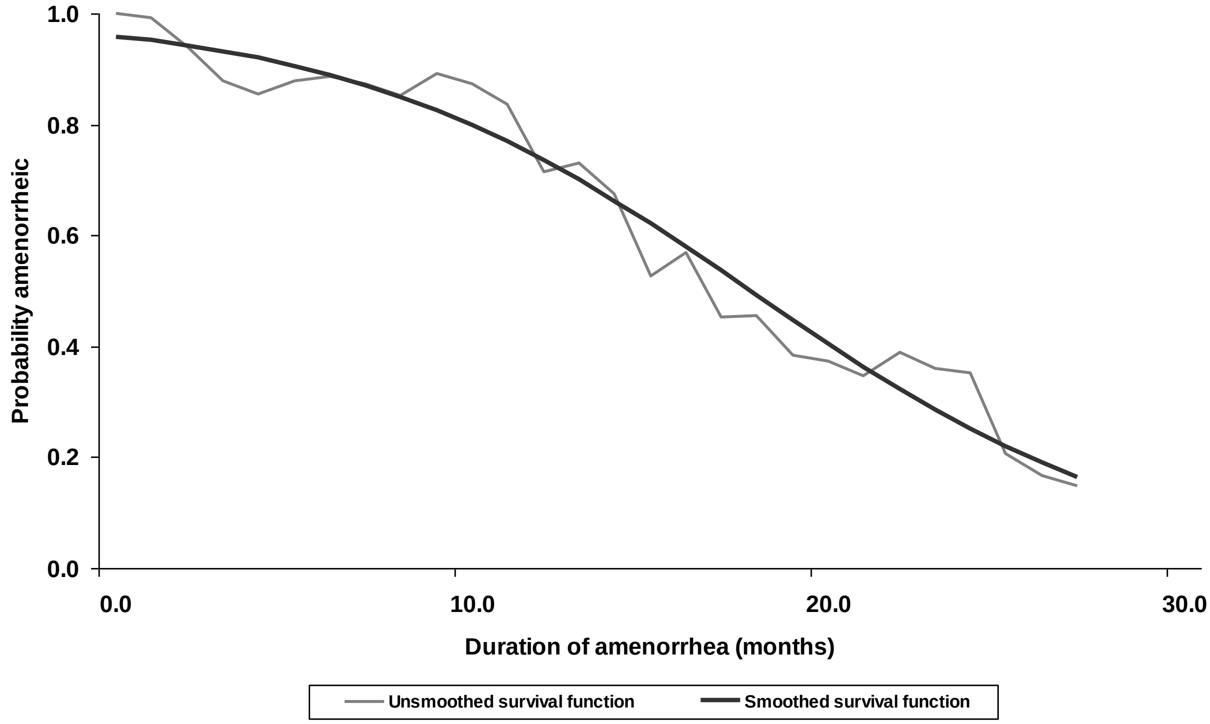
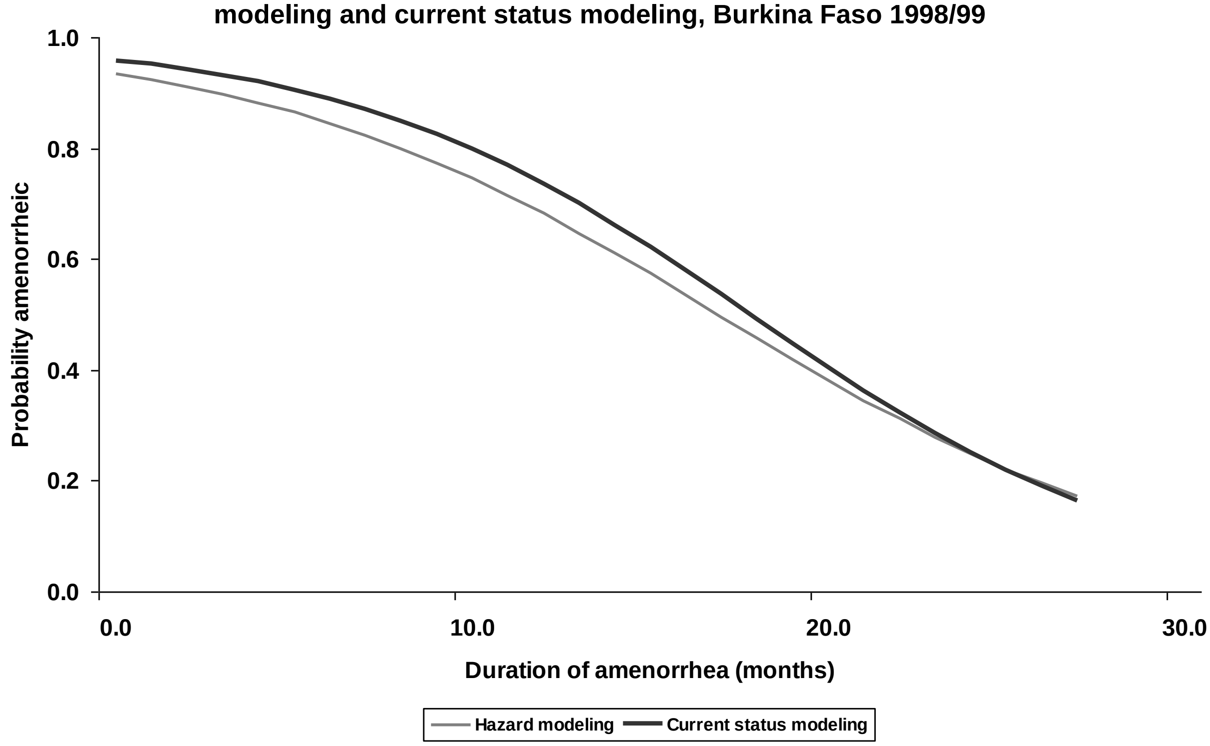


Figure 2.23 The smoothed survival functions for duration of amenorrhea, using hazard modeling and current status modeling, Burkina Faso 1998/99



3 Child Health

3.1 Morbidity and Treatment: Diarrhea, Fever, and Cough

Virtually all phase 3 and phase 4 DHS surveys included questions about recent episodes of diarrhea (h11), fever (h22), and cough (h31); the only exceptions are that Bangladesh 1993/94, Senegal 1997, South Africa 1998, and Turkey 1998 omitted questions about fever, and Senegal 1997 and Turkey 1998 omitted questions about cough. In most surveys the questions refer specifically to a reference period of two weeks preceding the date of interview.

Some of the DHS III surveys also included a reference period of 24 hours preceding the interview for the questions about diarrhea and cough. (No surveys used the shorter reference period for fever.) Brazil 1996 included the 24-hour option for diarrhea only. Haiti 1994/95 included the 24-hour option for cough only. The following 10 surveys included the 24-hour option for both diarrhea and cough: Burkina Faso 1998/99, Cote d'Ivoire 1994, Dominican Republic 1996, Ghana 1993, Indonesia 1994 and 1997, Kenya 1993, Philippines 1993, Senegal 1997, and Turkey 1993. In the surveys that included a 24-hour window, 22 percent to 45 percent of children with diarrhea during the past two weeks also had it in the past 24 hours; 52 percent to 61 percent of children with cough during the past two weeks also had it in the past 24 hours. A high percentage indicates that when the symptoms occur, they tend to be prolonged, and the same children will be identified with both a 24-hour window and a two-week window.

For diarrhea and cough, the coded responses were 0: "no;" 1: "yes, in the past 24 hours" (only included in some surveys, see above); 2: "yes, in the past two weeks;" 8: "don't know;" and 9: "not stated." For fever, the codes were the same except that code 1 was "yes, in the past two weeks" and code 2 was never used, even though 2 was the code for "yes, in the past two weeks" for diarrhea and cough. If the response to an item was "yes," a variety of possible treatments were asked about. Responses for the possible treatments were coded "no" and "yes," with the possibility of more than one treatment. The options included "don't know" or "not stated" (distinguished from "." for "not applicable").

We will consider five binary measures of non-response. The first, h_diarrhea_missing, is coded 0 if the response to h11 was "no" or "yes," 1 if the response was "don't know" or "not stated," and "." if h11 was coded ".". h_fever_missing and h_cough_missing are constructed in a similar way from h22 and h31, respectively. The fourth measure, h_diarrhea_treatment_missing, is coded 1 if any of the items about diarrhea treatment is "don't know" or "not stated," 0 if all of those items received "no" or "yes" responses, and "." if h11 was coded anything other than "yes." The fifth measure, h_fevercough_treatment_missing, was constructed similarly, combining fever and cough treatments. The level of missing on these five measures across all 81 surveys, giving equal weight to each survey, was 1.9 percent, 2.1 percent, 2.1 percent, 2.2 percent, and 0.8 percent, respectively. In most surveys, the levels of missing on the first three indicators—recent symptoms of diarrhea, fever, and cough—are very similar.

Figures 3.1 to 3.5 graph the distributions of the percentages missing on each of the five measures; the figures differ somewhat in the scales of their axes. They show that most surveys are close to the respective means and only a few are scattered out at higher levels.

Figure 3.1 Distribution of the percentages of cases that are missing whether a child recently had diarrhea, all DHS surveys 1993-2003

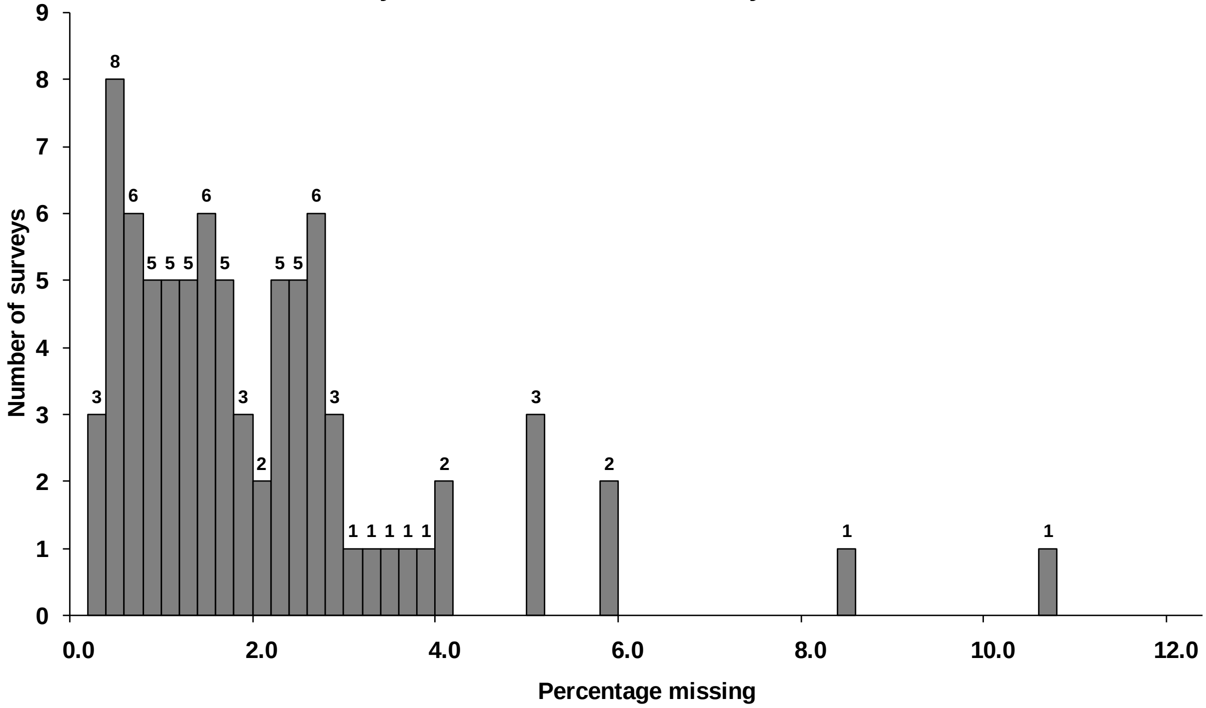


Figure 3.2 Distribution of the percentages of cases that are missing whether a child recently had fever, all DHS surveys 1993-2003 other than Bangladesh 1993/94, Senegal 1997, South Africa 1998, and Turkey 1998

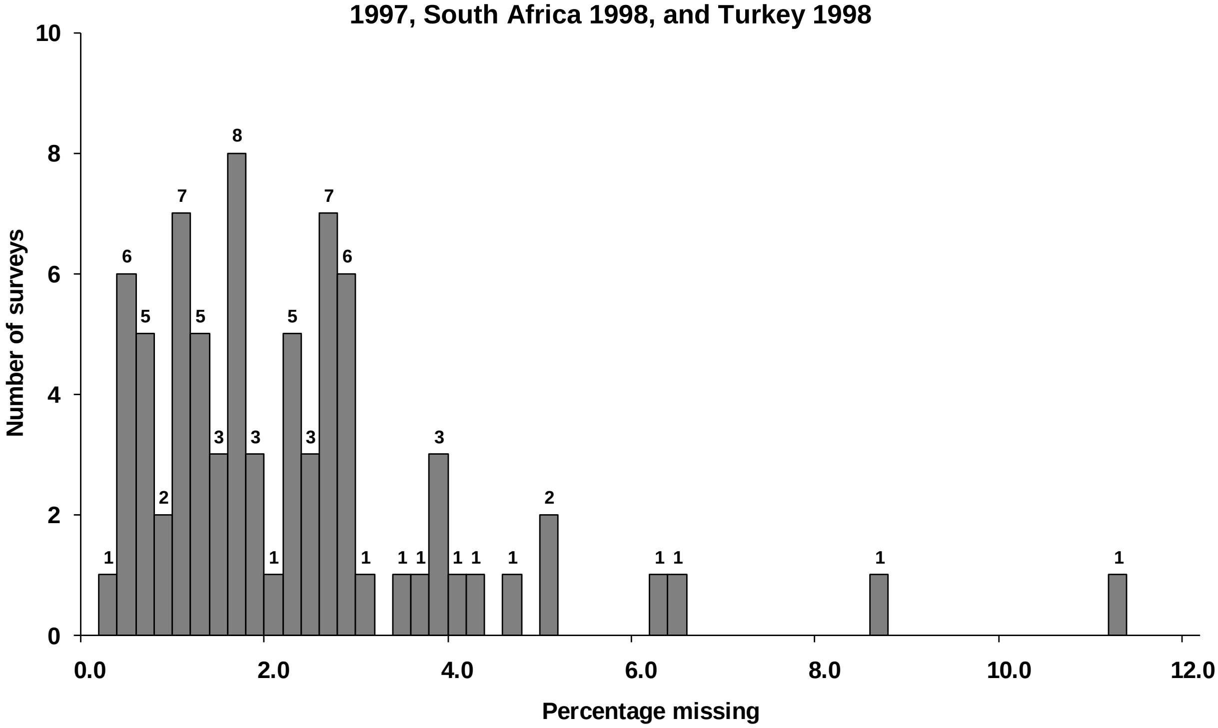


Figure 3.3 Distribution of the percentages of cases that are missing whether a child recently had a cough, all DHS surveys 1993-2003 other than Senegal 1997 and Turkey 1998

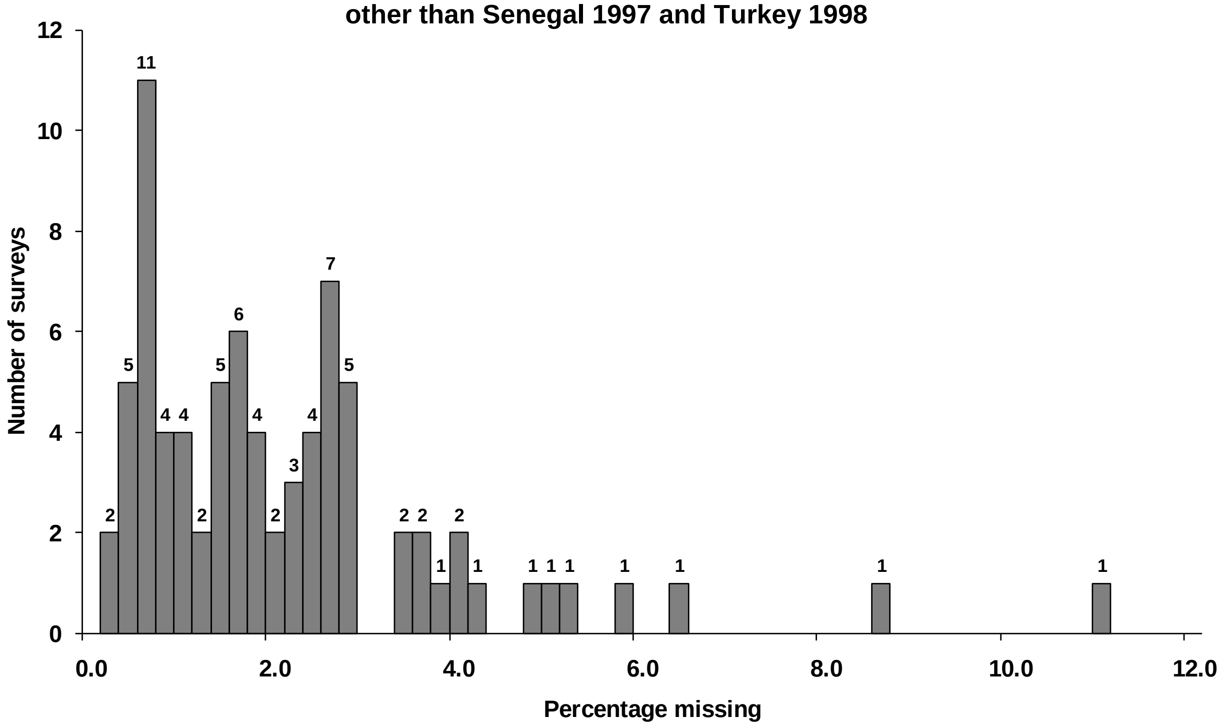


Figure 3.4 Distribution of the percentages of cases that are missing whether a child who recently had diarrhea received treatment, all DHS surveys 1993-2003

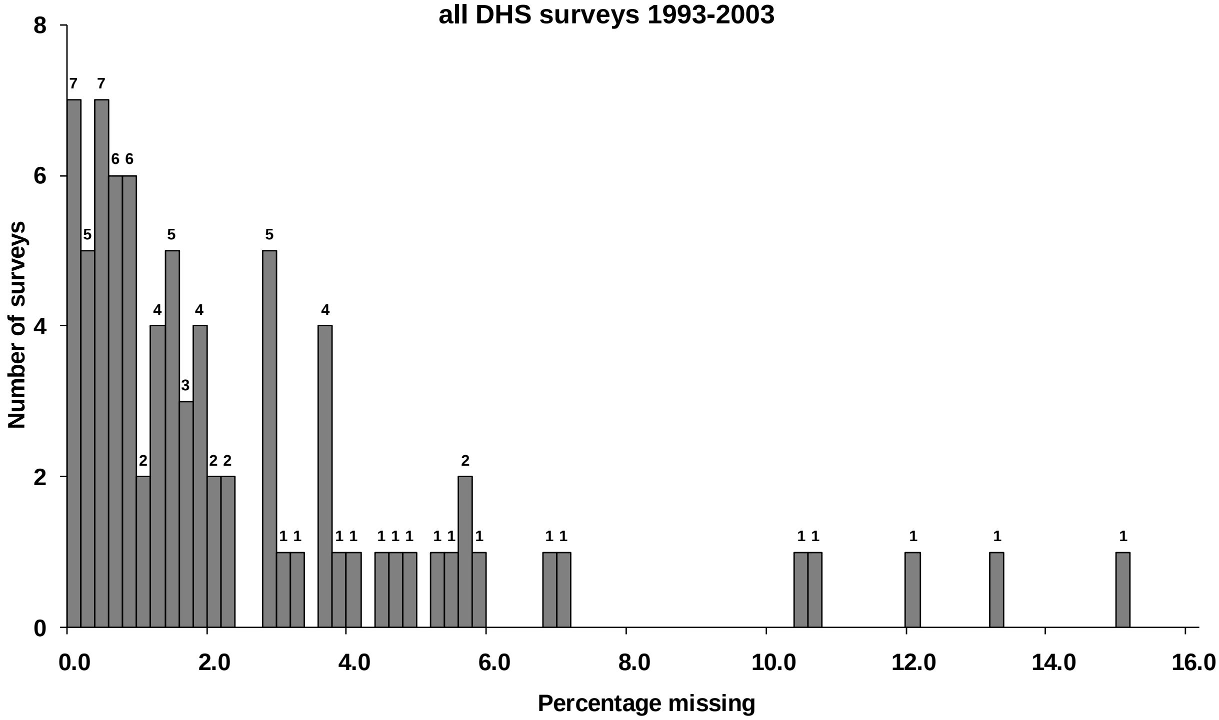
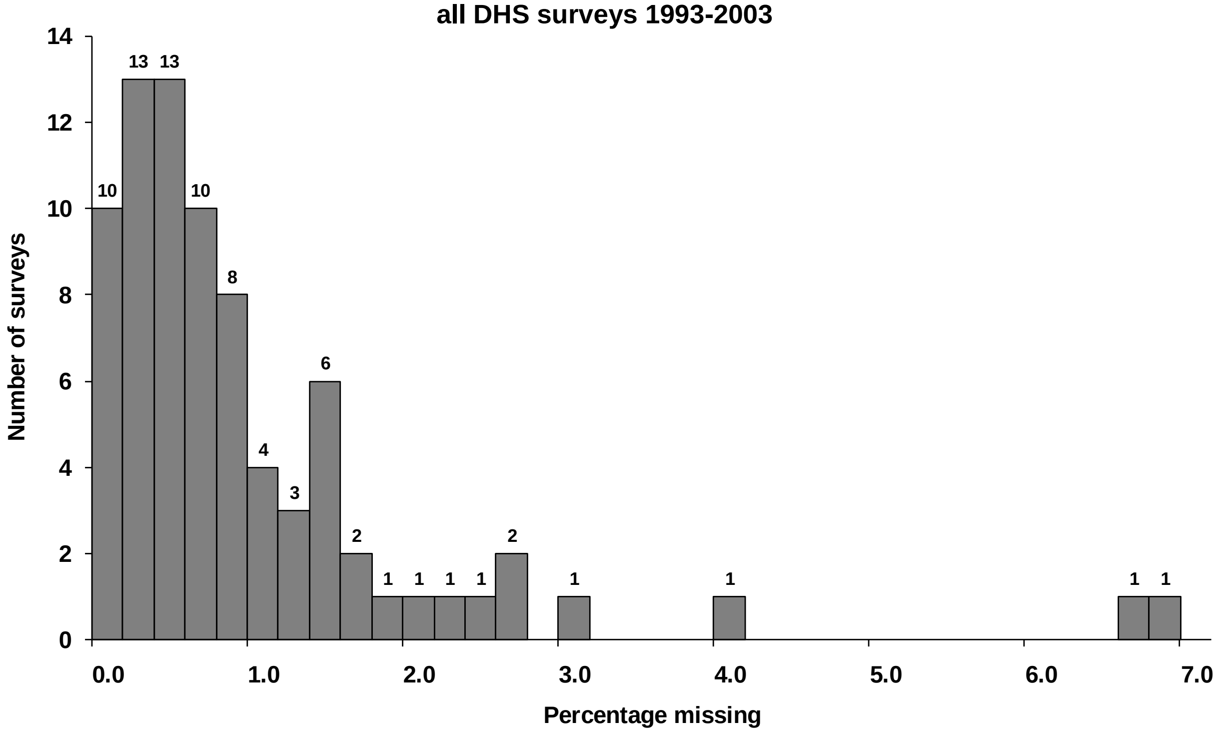


Figure 3.5 Distribution of the percentages of cases that are missing whether a child who recently had fever or cough received treatment, all DHS surveys 1993-2003



The surveys with the highest levels of missing responses on any of these five indicators are listed in Table 3.1. The threshold level for this table is arbitrarily set at 4.0 percent. Surveys are listed if any of the five percentages reached that level and any percentages below that level are blanked out. Most of the problems fall into two clear patterns, with some overlap. Nine surveys had high (i.e., ≥ 4.0 percent) levels of missing for symptoms of diarrhea, fever, and/or cough: Cote d'Ivoire 1998/99, Gabon 2000, Namibia 2000, South Africa 1998, Tanzania 1996 and 1999, Uganda 1995 and 2000/01, and Zimbabwe 1999. All of these surveys were in sub-Saharan Africa. Gabon 2000 and Namibia 2000 have the highest levels for all three indicators. A second set of 17 surveys had high levels of missing on treatment for diarrhea or fever/cough, but mainly for diarrhea. Only three surveys in this group occurred outside of sub-Saharan Africa: Bangladesh 1996/97 and Philippines 1993 and 1998. The overlap of these two patterns includes only Gabon 2000, Namibia 2000, South Africa 1998, and Zimbabwe 1999.

In an effort to find covariates of these two patterns of missing, we will pool the surveys that display the patterns and then look for any systematic relationship with the standard six covariates.

Figures 3.6 to 3.8 show the percentage of cases missing symptoms of diarrhea in the past two weeks within categories of the only three covariates that are related to this variable with a pseudo-R^2 of at least .01. These figures are limited to the seven surveys listed in Table 3.1 with levels of missing of at least 4.0 percent, pooled with equal weight given to each survey. The seven surveys are Cote d'Ivoire 1998/99, Gabon 2000, Namibia 2000, South Africa 1998, Tanzania 1999, Uganda 2000/01, and Zimbabwe 1999. They show a much higher level of missing in urban areas than in rural areas, higher levels for women under age 30 than for women 30 and over, and a steady increase in the level of missing for children who are older. The child's age is by far the most important covariate. For all covariates, "don't know" is by far the largest component of missing responses. An interpretation of the cases will be given below.

Table 3.1 Surveys in which 4.0 percent or more of weighted responses to questions about diarrhea, fever, cough, or treatment for any of these is coded "don't know" or "not stated," all DHS surveys 1993-2003 that included these questions

	Symptoms in past 2 weeks			Treatment given for:	
Survey	Diarrhea	Fever	Cough	Diarrhea	Fever/ cough
Bangladesh 1996/97				10.5	
Bangladesh 1999/2000				6.9	
Benin 2001				4.6	
Cote d'Ivoire 1998/99	5.6	6.0	5.7		
Dominican Republic 2002				5.4	
Gabon 2000	8.3	8.5	8.4	5.6	
Ghana 1993				6.8	
Guinea 1999				5.4	
Indonesia 2002				5.4	
Kazakhstan 1999					6.6
Mozambique 2003				4.6	
Namibia 2000	10.5	11.0	11.0		6.7
Nigeria 1999				14.9	
Nigeria 2003				5.2	
Philippines 1993				10.4	
Philippines 1998				4.4	
South Africa 1998	4.9		5.2	13.0	
Tanzania 1996		4.6	4.2		
Tanzania 1999	4.8	5.0	4.8		
Uganda 1995			4.2		
Uganda 2000/01	4.8	4.9	4.9		
Zimbabwe 1999	5.7	6.3	6.4	11.9	

Figure 3.6 Percentage of cases missing symptoms of diarrhea in the past two weeks, by type of place of residence, for the seven surveys with at least 4 percent missing on this variable (pseudo-R^2 = .013)

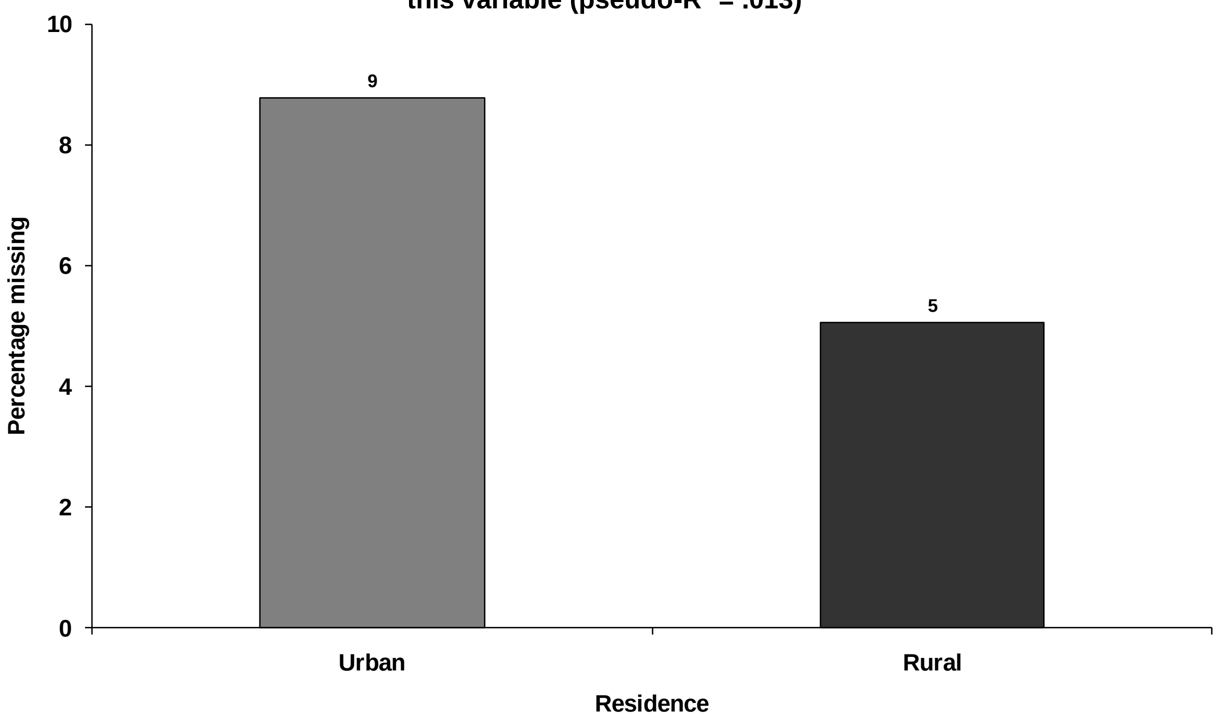


Figure 3.7 Percentage of cases missing symptoms of diarrhea in the past two weeks, by age group of mother, for the seven surveys with at least 4 percent missing on this variable (pseudo-R^2 = .012)

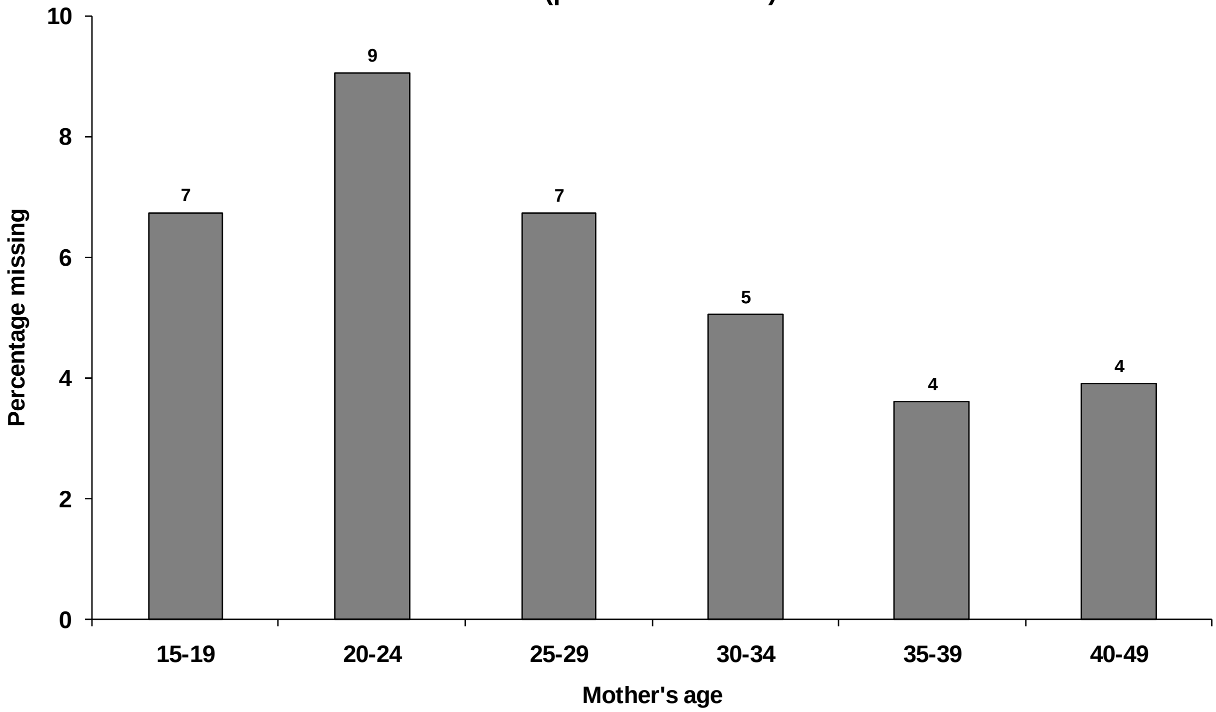


Figure 3.8 Percentage of cases missing symptoms of diarrhea in the past two weeks, by age of child, for the seven surveys with at least 4 percent missing on this variable (pseudo-R^2 = .077)

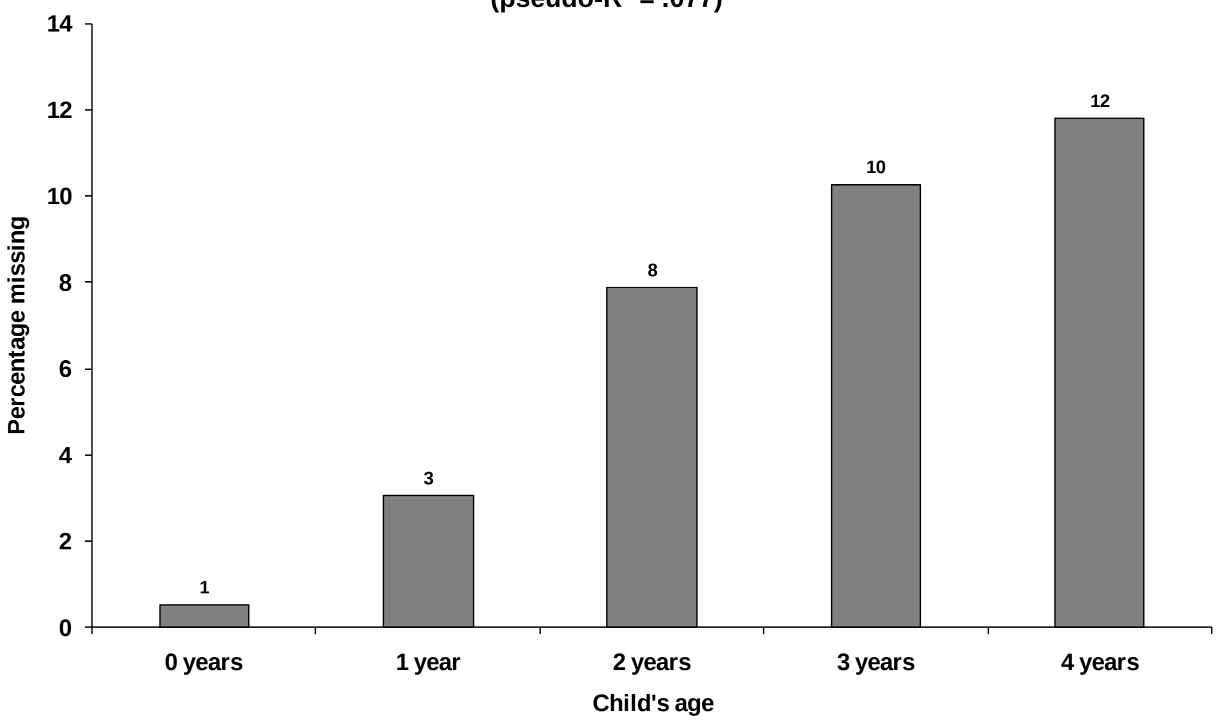


Table 3.2 Distribution of responses to the question about diarrhea symptoms, within type of place of residence, for the seven surveys with at least 4 percent missing on this variable

Had diarrhea recently	Residence		Total
	Urban	Rural	
No	77.39	78.58	78.14
Yes, past 2 weeks	13.84	16.38	15.45
Don't know	8.35	4.75	6.08
Not stated	0.41	0.30	0.34
Total	**100**	**100**	**100**

Table 3.3 Distribution of responses to the question about diarrhea symptoms, within intervals of mother's age, for the seven surveys with at least 4 percent missing on this variable

Had diarrhea recently	Woman's (mother's) age interval						Total
	15-19	20-24	25-29	30-34	35-39	40-44	
No	71.6	74.23	78.4	81.4	81.7	83.8	78.1
Yes, past 2 weeks	21.6	16..73	14.9	13.6	14.7	12.3	15.5
Don't know	6.52	8.79	6.41	4.74	3.03	3.34	6.08
Not stated	0.22	0.25	0.32	0.3	0.58	0.56	0.34
Total	**100**	**100**	**100**	**100**	**100**	**100**	**100**

Table 3.4 Distribution of responses to the question about diarrhea symptoms, by age of child, for the seven surveys with at least 4 percent missing on this variable

Had diarrhea recently	Age of child					Total
	0	1	2	3	4	
No	79.8	70.9	77.5	81.3	82.1	78.1
Yes, past 2 weeks	19.7	26.0	14.6	8.5	6.1	15.5
Don't know	0.5	2.9	7.4	9.8	11.2	6.1
Not stated	0.1	0.2	0.5	0.5	0.6	0.3
Total	**100**	**100**	**100**	**100**	**100**	**100**

Further investigation of the "don't know" responses indicates that they are largely attributable to the child not residing in the same household as the mother. When this is the case, the mother will have less frequent contact with the child, and "don't know" is an appropriate response to questions about current illness.

If the child is in a separate household, then b16, the child's line in the household survey, is assigned code 0. This variable was only included in 24 of the 81 surveys and was only coded 0 for 6,131 children, less than 2 percent of the children in the consolidated file. (b16 is now included in all surveys.) However, with a limitation to those surveys, we can examine the relationship of b16 to responses about current illness. Table 3.5 gives the unweighted frequencies of the "don't know" and the "not stated" responses (grouped together as "missing") within the two categories of co-residence for the questions about diarrhea, fever, and cough in the last two weeks. For each of the three questions, the frequencies in the "no" column add to 6,131.

Table 3.5 shows that the overwhelming majority of unusable cases are "don't know" responses given for children who do not live with the mother. Indeed, of the 3,412 + 305 = 3,717 unusable cases for recent diarrhea, 3,122, or 84.0 percent, have this source. For fever, the corresponding percentage is 82.7 percent, and for cough it is 83.9 percent.

Of the seven surveys with the highest levels of missing (including "don't know") data in Table 3.1, three surveys (Namibia 2000, Uganda 2000/01, and Zimbabwe 1999) included b16. For these surveys, the pattern is the same: about 89 percent of the unusable cases are due to children living separately from their mothers. It is very likely, although it cannot be demonstrated, that in the surveys that *do not* include b16 the pattern is also the same.

Table 3.5 Unweighted frequencies of the responses to questions about diarrhea, fever, and cough in the past two weeks, by co-residence status of the mother and child, all DHS surveys 1993-2003 that included co-residence status

	Mother and child in the same household		
Recent symptom	No	Yes	Total
Diarrhea			
No	2,640	139,678	142,318
Yes	274	27,905	28,179
Don't know	3,122	290	3,412
Not stated	95	210	305
Fever			
No	2,066	115,118	117,184
Yes	664	52,363	53,027
Don't know	3,310	339	3,649
Not stated	91	263	354
Cough			
No	2,008	110,863	112,871
Yes	747	56,682	57,429
Don't know	3,283	278	3,561
Not stated	93	260	353

If the mother and child live in different households, about half of the responses are "don't know." The percentage of responses that are "not stated" is also higher for such mothers than for mothers who are co-resident with the child.

The balance of the "yes" and "no" responses is also significantly different for the two kinds of living arrangements. If the mother and child live together, and the mother gives a "yes" or "no" response, the percentage responding "yes" is 17 percent, 31 percent and 34 percent, respectively, for the three questions. If the mother and child are in different households, the corresponding percentages are 9 percent, 24 percent, and 27 percent, respectively. That is, if a mother who *does not* live with her child gives a "yes" or "no" response, she is more likely to say "no" to the question about diarrhea, more likely to say "no" to the question about fever, and more likely to say "no" to the question about cough, compared with a mother who *does* live with her child. It is possible—and perhaps likely—that the symptoms have different prevalence for children who live separately from the mother. However, the data suggest that some of the mothers who live separately from their children are simply guessing. They are more likely to guess that the child *does not* have diarrhea, fever, or cough, because most children do not, and because it is a preferable response.

A more thorough, within-country analysis of the responses of mothers who live separately, with controls for the age of the child, would be desirable, but it is suggested that if the child lives in another household, questions about symptoms of illness in the past two weeks could be skipped. This suggestion would also apply to the questions about liquid and solid foods in the past week, to be discussed in Section 3.2.

In all surveys that included questions about cough, with the sole exception of the Colombia 2000 survey, an additional question was asked about children who had a cough within the last two weeks. This question asked whether the cough was accompanied by "short, rapid breaths," which would indicate a more serious medical condition, perhaps an acute respiratory infection (ARI). The variable has standard recode label h31b, and codes 0 for "no," 1 for "yes," 8 for "don't know," and 9 for "not stated."

In the consolidated file of all index children, a total of 163,072 children were reported to have had a cough in the last two weeks (excluding the Colombia 2000 survey), and therefore should have had a response to the question about short, rapid breaths. Those children, and their responses, are included in Table 3.6. It shows that about half of the children *did* have the more serious cough. The percentage with a serious cough was slightly higher if the cough was reported for the last two weeks but not the last 24 hours.

Overall, less than 2 percent of children received "don't know" or "not stated" codes for h31b, with approximately equal numbers of each (1,357 and 1,635, respectively). Looking at just the last two columns of Table 3.6, the "don't know" response was more likely (and very significantly more likely) than the "not stated" response if the cough persisted into the past 24 hours, rather than being observed earlier in the two-week window but having ended.[5]

Table 3.6 Frequency distributions (unweighted) of the responses to the question about "short, rapid breaths" for all index children reported to have had a cough in the past two weeks, all DHS surveys 1993-2003 that included this question

	Short, rapid breaths				
Had cough in past 2 weeks	No	Yes	Don't know	Not stated	Total
Yes, past 24 hours	3,611	3,127	96	41	6,875
	52.52	45.48	1.40	0.60	100.00
Yes, past 2 weeks	75,262	78,080	1,261	1,594	156,197
	48.18	49.99	0.81	1.02	100.00
Total	**78,873**	**81,207**	**1,357**	**1,635**	**163,072**
	48.37	**49.80**	**0.83**	**1.00**	**100.00**

Most surveys have negligible levels of "don't know" or "not stated" (combined as "missing") codes for h31b. Table 3.7 lists the surveys for which at least 2.0 percent of the responses were in either of these categories. Earlier, in the discussion of diarrhea, fever, and cough, Table 3.1 used a 4.0 percent threshold for the two codes combined, and here they are separated. Table 3.7 also distinguishes between whether the cough was reported to have continued into the last 24 hours or was earlier in the two-week window.

[5] It is possible that the responses depend somewhat on whether the mother and child live in the same household. However, the question about "short, rapid breaths" (h31b) was only asked of women who responded "yes" to the question about a cough in the last two weeks (h31). Of the 55,185 (unweighted) children in the consolidated file to whom this question applied, only 704 lived in a different household than the mother. Of them, 38 had a "don't know" response to h31b and 16 were "not stated." The small number of cases, scattered over several different surveys, prevents any useful analysis.

Table 3.7 indicates that the “not stated” codes are more scattered across surveys than the “don’t know” codes and reach the 2.0 percent threshold in only two surveys, Mali 2001 and Turkey 1993. If the cough occurred in the last 24 hours, not just in the last two weeks, then the “don’t know” responses are also more uniform across surveys and only reach the 2.0 percent threshold in a single survey, Turkey 1993.

Except for the Turkey 1993 survey, the “don’t know” responses tend to be concentrated in specific surveys only when the cough occurred in the last two weeks and not in the last 24 hours. They then exceed the threshold in 10 surveys. In addition to Mali 2001 and Turkey 1993, these include Armenia 2000, Burkina Faso 1998/99, Cote d’Ivoire 1994, Haiti 1994/95, Kazakhstan 1995, Mali 1995/96, Namibia 2000, and Uzbekistan 1998. The last column of Table 3.7 gives the denominator for each survey, that is, the number of index children who had a serious cough and for whom a response to h31b was expected. Two of the surveys, Kazakhstan 1998 and Uzbekistan 1996, had very low frequencies of serious cough, at 117 and 78, respectively. No significance, statistical or substantive, should be attached to their inclusion in the table. The Armenia 2000 survey also had relatively few children for whom a response was expected—only 430.

Table 3.7 Surveys in which 2.0 percent or more of weighted responses to the question about “short, rapid breaths” were coded “don’t know” or “not stated,” all DHS surveys 1993-2003 that included this question

	Past 24 hours		Past 2 weeks		
Survey	Don’t know	Not stated	Don’t know	Not stated	Number of cases
Armenia 2000	10.3	6.7			430
Burkina Faso 1998/99		2.4			1661
Cote d’Ivoire 1994		2.9			1236
Haiti 1994/95		3.7			1630
Kazakhstan 1995		5.3			117
Mali 1995/96		3.6			1585
Mali 2001		7.5		6.4	2797
Namibia 2000		2.3			1064
Turkey 1993		14.3		3.5	955
Uzbekistan 1996		9.6			78

In terms of percentages, the most serious excesses of both “don’t know” and “not stated” responses are in the Mali 2001 and Turkey 1993 surveys. In terms of numbers of children, the Mali 2001 survey is clearly most serious, because of the large number of children who had a cough in the last two weeks—2,797—which was nearly three times the number in the Turkey 1993 survey and far more than in any other survey listed in Table 3.7.

We have been unable to identify any within-survey pattern of variation in these two surveys. Their levels of “don’t know” and “not stated” on h31b do not vary significantly across the categories of any of the covariates used in this report. A more detailed analysis might be able to uncover variations related to location, interviewers, etc. The much higher prevalence of these responses in the Mali 2001 survey than in the Mali 1995/96 survey and the much higher prevalence in the Turkey 1993 survey than in the Turkey 1998 survey suggests the influence of the overall level of training and supervision of interviewers with respect to this specific question.

3.1.1 Patterns of Missing for Treatment of Diarrhea, Fever, and Cough

Next, we look for systematic patterns in the surveys with the highest levels of missing on treatment for diarrhea (h13-15) and fever and cough (h32). The highest levels of missing for diarrhea treatment were found in Bangladesh 1999/2000 (10.5 percent), Nigeria 1999 (14.9 percent), Philippines 1993 (10.4 percent), South Africa 1998 (13.0 percent), and Zimbabwe 1999 (11.9 percent). The highest levels of missing for fever/cough treatment were found in Kazakhstan 1999 (6.6 percent) and Namibia 2000 (6.7 percent), but the numbers of affected surveys and levels of missing are much greater for treatment of diarrhea than for treatment of fever and cough.

A pooling of all 14 surveys that are missing at least 4.0 percent of responses on treatment for diarrhea showed no relationships with the six covariates that are being used here. The pseudo-R^2 values for these logit regressions were all well below 1 percent. In order to reduce heterogeneity, the pooling was then restricted to the five surveys listed in the previous paragraph that are missing at least 10.0 percent of responses for treatment for diarrhea. That subset also did not show any relationship between missing and the six covariates. We then looked at these surveys one at a time. The only statistically significant patterns in relation to treatment for fever and cough were found for the Kazakhstan 1999 survey. The level of missing for this question is found to be significantly (at the .01 level) related to the age of the mother and the age of the child. Figures 3.9 and 3.10 show that the probability of missing generally increases with age of mother, but with a spike for ages 30-34, and increases with the age of the child from ages 0 through 3, but drops for age 4.

In the Kazakhstan 1999 survey, about 12 percent of children were reported to have fever in the past two weeks, and about 12 percent were reported to have a cough in the past two weeks. There was a strong association between these two symptoms; about one–third of the children who had one symptom or the other had *both* symptoms. For most kinds of possible treatments, the treatment is most likely when the child had both symptoms, and when only one symptom or the other was manifested, treatment was slightly more likely for fever than for cough.

Figure 3.9 Percentage of responses to question about treatment for fever and cough in the past two weeks (h32) that are missing in the Kazakhstan 1999 survey, by age of mother

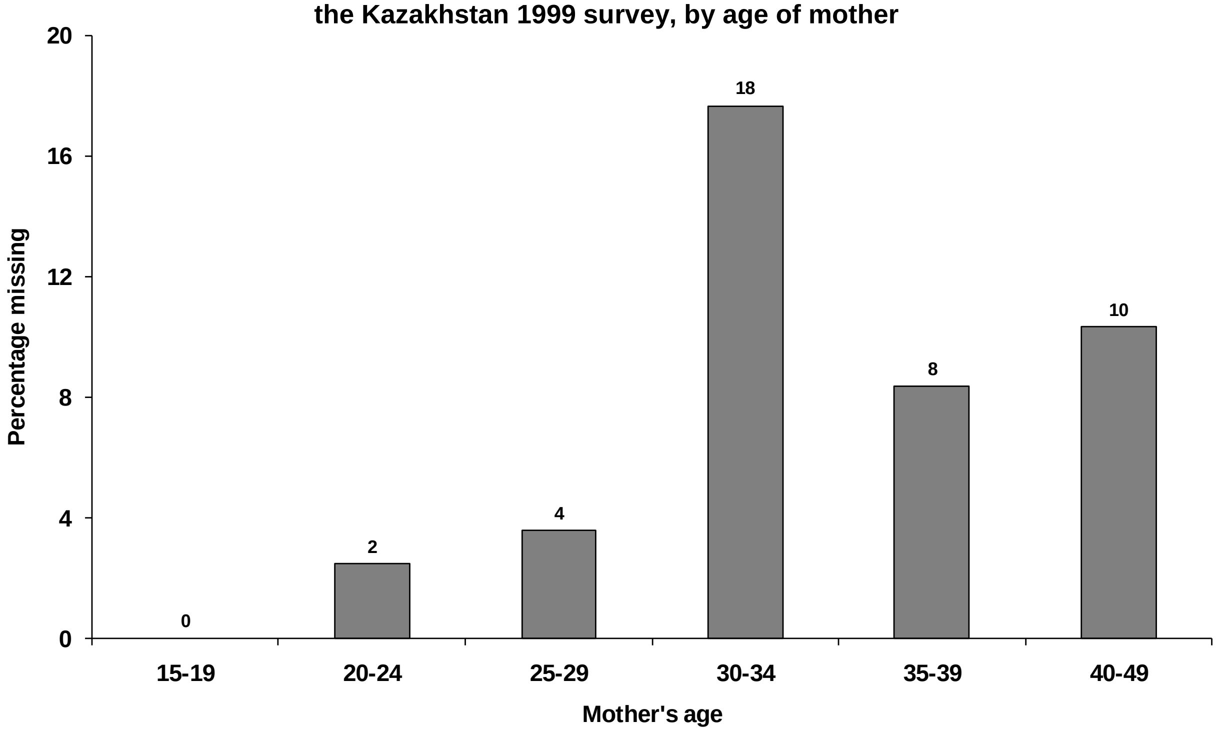
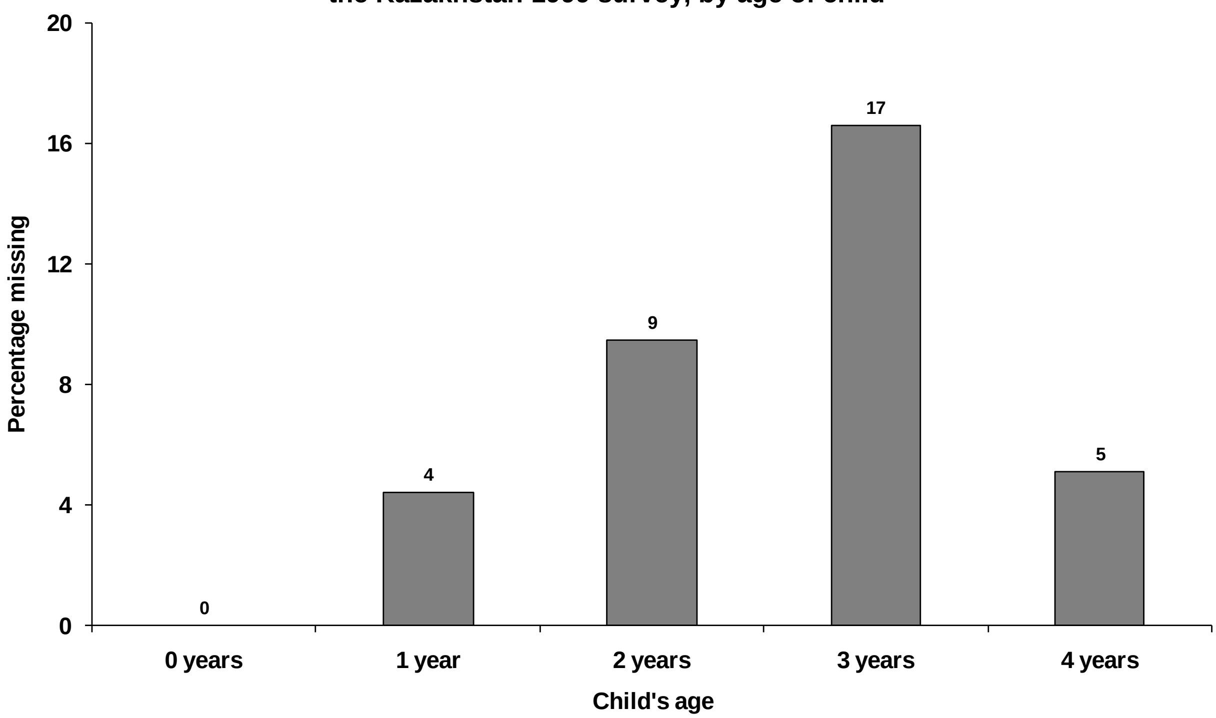


Figure 3.10 Percentage of responses to question about treatment for fever and cough in the past two weeks (h32) that are missing in the Kazakhstan 1999 survey, by age of child



3.1.2 Seasonality in Diarrhea, Fever, and Cough

It is likely that these symptoms are related to rainfall, temperature, and other seasonal factors that influence water quality, breeding of insects, and so on. If this is the case, then the timing of a survey may produce estimates of the prevalence of diarrhea, fever, and cough that are not simply an average for the calendar year but that are artificially high or low just because of the months when the fieldwork was conducted. If, say, a subsequent survey were conducted at a different time of year, a lower or higher prevalence than in the previous survey could easily be misinterpreted. This type of distortion does not reflect on data quality as such, but will be considered here because of its potential effect on the analysis.

To investigate seasonality, we look at whether the probability of a "yes" response to any of the three symptoms, in turn, is significantly (at the .01 level) related to a categorical predictor for the calendar month of the interview (v006). No other predictor or control variables are included.[6] The statistical model is logit regression, which produces a test statistic with an asymptotic chi-square distribution with degrees of freedom equal to the number of months of fieldwork minus one. It would be possible to calculate a similar chi-square statistic just from a two-way table in which the children were classified according to month of interview and whether or not their mother gave a "yes" response. Logit regression simply adjusts this chi-square statistic for sampling weights and sample clusters.

The expected value of any chi-square statistic is its degrees of freedom. In order to identify the surveys with the greatest amount of month-to-month variation, we have calculated the ratio of the calculated chi-square to its degrees of freedom. This ratio will be affected by the size of the survey, as well as the monthly variations in prevalence, but essentially measures the strength of the relationship between month and prevalence in terms of statistical significance. All of the specific surveys to be discussed here are highly significant, at much better than the .01 standard.

We find very strong evidence of seasonality. Virtually all surveys show differences across months that are statistically significant (at the .01 level) for at least one of the three symptoms.

Table 3.8 lists the 21 surveys in which at least one of the ratios of chi-square to degrees of freedom is greater than 10.00.[7] To facilitate the inspection of the table, ratios less than this arbitrarily high level are blanked out. Seventeen surveys exceed a ratio of 10 for cough, five exceed that level for diarrhea, and three exceed it for fever, giving a rough guide to relative degree of seasonality in the three symptoms. Cough clearly has the most seasonality. The countries are widely dispersed and not concentrated in any particular region. As just mentioned, the index is affected by the size of a survey. The table includes two of the three largest surveys, in terms of numbers of surviving children, Indonesia 2002 and Peru 2000,[8] but does not include the largest survey of all, India 1998/99.

[6] It is conceivable that a country with strong seasonality of births, and a strong concentration of the symptoms in certain ages, could show a spurious seasonality of symptoms, but we doubt that a control for month of birth would affect the conclusions. In some countries there was a pattern to the fieldwork (perhaps beginning in the capital city and then moving to rural areas and more remote regions) that would motivate adding controls for region or type of place of residence.

[7] The Kenya 2003 survey indicates highly significant seasonality on fever and cough but is omitted from this list because the large chi-squares are largely due to a sharp drop in prevalence during the last month of interviewing, during which there were relatively few interviews.

[8] The largest survey of all, India 1998/99, did have very significant seasonality for all three kinds of symptoms, but does not exceed the threshold of 10 on this index.

Table 3.8 Surveys with strongest evidence of seasonality of diarrhea, fever, and cough, all DHS surveys 1993-2003 that included these questions

Survey	Diarrhea	Fever	Cough
Benin 2001	18.69		
Bolivia 2003/04			22.96
Egypt 1995	14.31		11.65
Egypt 2000	15.47		35.55
Ghana 1998			13.11
Haiti 2000			13.11
Indonesia 2002		14.61	21.30
Madagascar 1997			11.51
Malawi 2000			10.91
Mozambique 1997	14.10		
Namibia 2000		12.75	
Nepal 1996			10.17
Nepal 2001			12.22
Niger 1998			13.29
Peru 2000			11.41
Philippines 1993		10.40	26.94
Philippines 1998			16.05
Rwanda 2000			12.61
Turkey 1998	10.09		
Uzbekistan 1996			22.09
Zambia 1996			56.37

Note: The entries in the table are the ratio of a chi–square test statistic to its degrees of freedom, when that ratio is greater than 10.00.

Tables 3.9 through 3.11 illustrate the patterns of seasonality for the surveys and symptoms listed in Table 3.8. Many other surveys show similar patterns but are not included because they did not exceed the threshold of 10 for this particular index. The rows refer to the calendar months January (month 1) through December (month 12). Table entries give the reported prevalence of the symptom in the last two weeks for index children, so there is a small lag due to using the calendar month of interview rather than the calendar month of the symptom. The fieldwork for several surveys began late in one calendar year and extended into the next calendar year.

As would be expected because of the wide geographical separation of most of these countries, the patterns differ considerably from one survey to another, but they generally follow one of four patterns: monotonically increasing, monotonically decreasing, a "U" shape, or an "inverted U" shape. When there is monotonicity, there can be doubt as to whether either the "worst" or the "best" month of the year was included in the fieldwork. A consistent "U" shape will generally identify the "best" month (i.e., the month with lowest prevalence) but not the "worst," and an "inverted U" shape will generally identify the "worst" month but not the "best." (If prevalence has more than one peak over the course of a year then the generalizations in the previous sentence may not even hold.)

For example, in the Indonesia 2002 survey there is a minimum for all three symptoms (not just fever, as shown in Table 3.10 and cough, as shown in Table 3.11) in February[9] but it is not at all clear that the calendar month when all three systems are most common, October, is the worst month of the year for these surveys, because October was the first month of fieldwork.

[9] There were only 225 children with data for February 2003 in this survey. These tables include a dashed line (–) for months with fewer than 100 children because they will tend to be most affected by sampling error, but that could come into play for this cell too. The next lowest month for this survey, for all three symptoms, is January, for which there are reports on 4,507 children, so the minimum prevalence is almost certainly in the January-February interval.

Benin provides an excellent example of the kind of misinterpretation that can result from seasonality. August was the worst month for all three symptoms in Benin 2001 (not just diarrhea, as shown in Table 3.9), but that was the first calendar month of fieldwork, and because of the downward trend after August, it is possible that the prevalence in earlier months was even higher than was observed for August. This possibility can be investigated by looking at the survey conducted in Benin in 1996 during the months of June, July, and August. The reported prevalence of diarrhea in those three months of 1996 was 27.4, 28.9, and 21.5, respectively. The prevalence of 21.5 in August 1996 is almost a perfect match with the prevalence of 21.7 in August 2001, and suggests that August was not a peak in 2001, but was simply a point on a downward seasonal trend.

Table 3.9 Reported prevalence of recent diarrhea among index children, by calendar month of interview, for surveys with the strongest evidence of seasonality of diarrhea, all DHS surveys 1993-2003 that included these questions

Month of interview	Benin 2001	Egypt 1995	Egypt 2000	Mozambique 1997	Turkey 1998
1	.	5.5	.	.	.
2	.	-	11.0	.	.
3	.	.	6.5	38.1	.
4	.	.	6.1	19.9	.
5	.	.	-	13.7	.
6	.	.	.	15.9	.
7	.	.	.	-	.
8	21.7	.	.	.	34.9
9	13.3	.	.	.	28.1
10	9.5	.	.	.	20.8
11	6.0	18.5	.	.	-
12	.	13.1	.	.	.
Total	**13.8**	**16.0**	**7.1**	**20.9**	**30.1**

Note: A dashed line indicates months with data on fewer than 100 children and "." indicates months outside the range of the fieldwork

Table 3.10 Reported prevalence of recent fever among index children, by calendar month of interview, for surveys with the strongest evidence of seasonality of fever, all DHS surveys 1993-2003 that included these questions

Month of interview	Indonesia 2002	Namibia 2000	Philippines 1993
1	20.6	.	.
2	15.0	.	.
3	35.4	.	.
4	32.3	.	29.4
5	.	.	25.3
6	.	.	21.7
9	.	.	-
10	.	35.8	.
11	36.1	21.4	.
12	32.0	20.4	.
12	23.7	16.0	.
Total	**26.2**	**21.9**	**25.5**

Note: A dashed line indicates months with data on fewer than 100 children and "." indicates months outside the range of the fieldwork.

Table 3.11 Reported prevalence of recent cough among index children, by calendar month of interview, for surveys with the strongest evidence of seasonality of cough, all DHS surveys 1993-2003 that included these questions

Sub-Saharan Africa

Month of interview	Ghana 1997	Madagascar 1998	Malawi 2000	Niger 1998	Rwanda 2000	Zambia 1996
1	21.5	.	.	-	.	-
2	17.8	.	.	43.0	.	.
3	.	.	.	36.9	.	.
4	.	.	.	24.0	.	.
5	.	.	.	22.5	.	.
6	.	.	.	18.5	50.0	.
7	.	.	51.0	-	44.3	62.6
8	.	-	52.3	.	40.2	57.2
9	.	52.5	46.7	.	36.2	50.2
10	.	46.3	39.9	.	33.1	40.5
11	34.5	36.6	47.9	.	33.8	38.9
12	28.7	30.9	.	.	.	30.8
Total	**26.3**	**42.7**	**47.4**	**25.6**	**36.8**	**45.7**

Asia

	Indonesia 2002	Nepal 1996	Nepal 2001	Philippines 1993	Philippines 1998
1	18.0	41.5	.	.	.
2	14.6	36.5	54.2	.	.
3	37.2	.	48.7	.	41.7
4	35.6	.	43.7	25.4	32.9
5	.	.	43.5	21.1	-
6	.	.	27.5	17.7	.
7	.	.	26.2	-	.
10	40.2	45.1	.	.	.
11	31.6	52.4	.	.	.
12	23.8	53.0	.	.	.
Total	**26.0**	**48.2**	**42.2**	**21.4**	**38.0**

Other Regions

	Bolivia 2003/04	Egypt 1995	Haiti 2000	Haiti 2000	Peru 2000	Uzbekistan 1996
1	23.1	27.5	.	.	.	.
2	.	57.1	30.1	65.7	.	.
3	.	.	18.1	73.6	.	.
4	.	.	13.7	62.6	.	.
5	.	.	-	64.2	.	.
6	.	.	.	66.1	.	-
7	.	.	.	51.2	54.5	6.1
8	48.5	.	.	.	44.4	3.4
9	41.9	.	.	.	43.8	-
10	37.9	.	.	.	39.3	-
11	38.4	48.1	.	.	35.6	.
12	30.7	44.6	.	.	.	.
Total	**37.2**	**46.1**	**19.4**	**63.6**	**43.4**	**4.9**

Note: A dashed line indicates months with data on fewer than 100 children and "." indicates months outside the range of fieldwork.

The overall mean prevalence of diarrhea from the Benin 1996 survey, based on the months of June through August, was 26.2. The mean from the Benin 2001 survey, based on the months of August through November, was 13.8, nearly half of the 1996 level. The month-by-month comparison suggests that there was no real improvement in diarrhea prevalence between 1996 and 2001; the lower rate is completely attributable to the timing of the fieldwork.

One way to deal with seasonality is to include calendar month of interview as a covariate in tables that describe these symptoms, and to make some effort to have repeated surveys during approximately the same months of the year. Another possibility would be to use information from other public health sources about variations during the year, and then try to develop a seasonally-adjusted measure. Otherwise, comparisons between surveys, either in different countries or in the same country, will be risky if the fieldwork was done at different times of year. At present, DHS reports do not generally refer to seasonality as an issue or emphasize the timing of the survey when reporting prevalence levels.

Many differentials between subgroups will probably be fairly stable throughout the year, and relatively unaffected by seasonality. However, subgroups that are more exposed to unsafe water and other risk factors would probably show greater seasonality than other subgroups, an effect that would alter differentials between subgroups. Further investigation of this topic would be helpful.

3.2 Child Feeding Practices

All surveys have questions about duration and months of breastfeeding for each index child (m4 and m5, respectively). The difference between these two measures was described earlier for postpartum amenorrhea and abstinence. Many have detailed questions about liquids and foods, specifically about types of liquids given to the child during the first three days after birth (m55, in 19 surveys) and types of liquids and foods given to the child during the seven days before the interview (m40, in 51 surveys). The overall levels of missing are low, particularly for breastfeeding and liquids in the first three days. Only 0.9 percent of the responses about months of breastfeeding are missing; 0.3 percent of responses about liquids in the first three days after birth are missing. The highest level of missing is for liquids and foods during the past seven days, for which 6.5 percent of the responses are missing. These overall percentages give equal weight to each survey.

Following the general practice in this report, "missing" includes "don't know" responses. For months of breastfeeding (m5), "missing" also includes inconsistencies that occurred because the stated duration of breastfeeding (m4) exceeds the length of the open interval (b11). The response "never breastfed" is included as a valid response.

Figures 3.11 to 3.13 give the distributions of the percentage missing on these three items, respectively. Note that the horizontal scales on these three figures are very different. The levels of missing for these questions are given in Table 3.12. This table uses a 2 percent threshold for months of breastfeeding and types of liquids given to the child in the first three days after birth. For liquids and foods given in the past seven days, a higher threshold is used, namely 5 percent, because so many surveys exceeded the 2 percent level. If the level of missing is less than the specified threshold, the entry in Table 3.12 is blanked out.

The highest level of missing for months of breastfeeding is for the Nigeria 1999 survey, at 5.7 percent, and the highest level for liquids in the first three days after birth is for the Peru 2000 survey, at 2.3 percent. These levels are trivial by comparison with liquids and foods in the past seven days, however. For 19 of the 51 surveys that included these questions, the level of missing exceeded 5 percent, and for four surveys it exceeded 20 percent. There is not a clear regional pattern to the missing responses on this question; the list includes surveys from Latin America and Central Asia as well as sub-Saharan Africa.

Figure 3.11 Distribution of the percentages of cases that are missing months of breastfeeding (m5), all DHS surveys 1993-2003 that included these questions

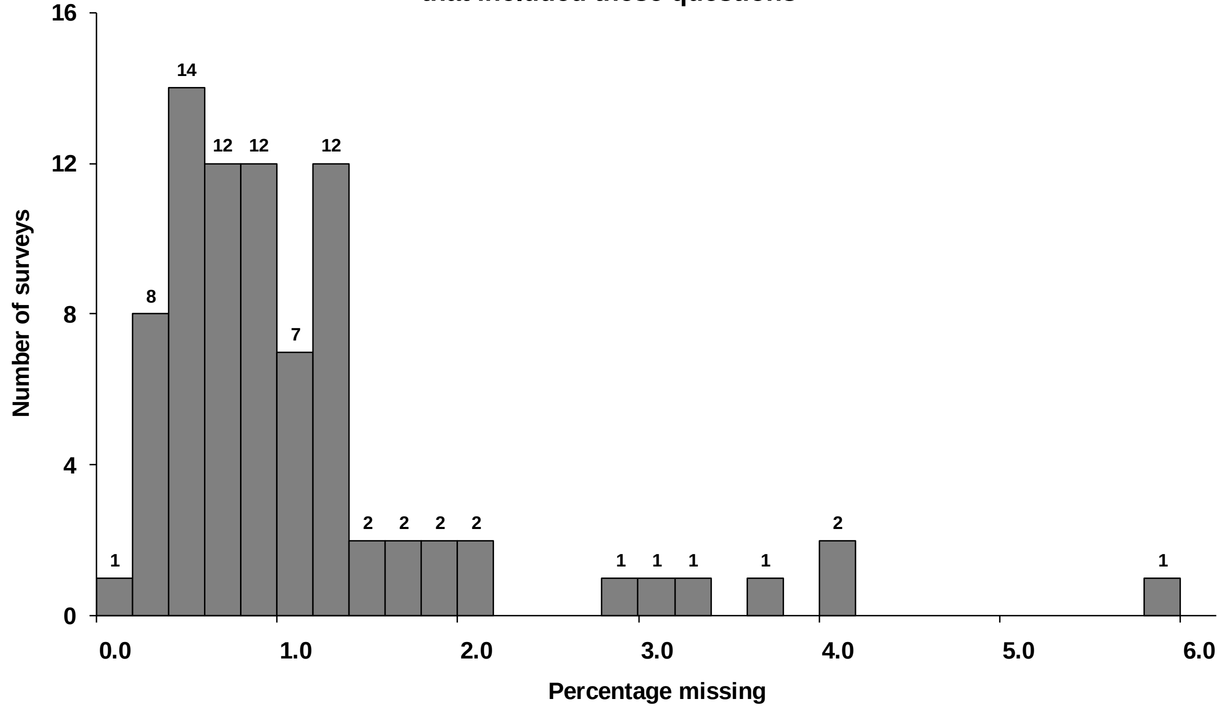


Figure 3.12 Distribution of the percentages of cases that are missing data on liquids in the first three days after birth (m55), all DHS surveys 1993-2003 that included these questions

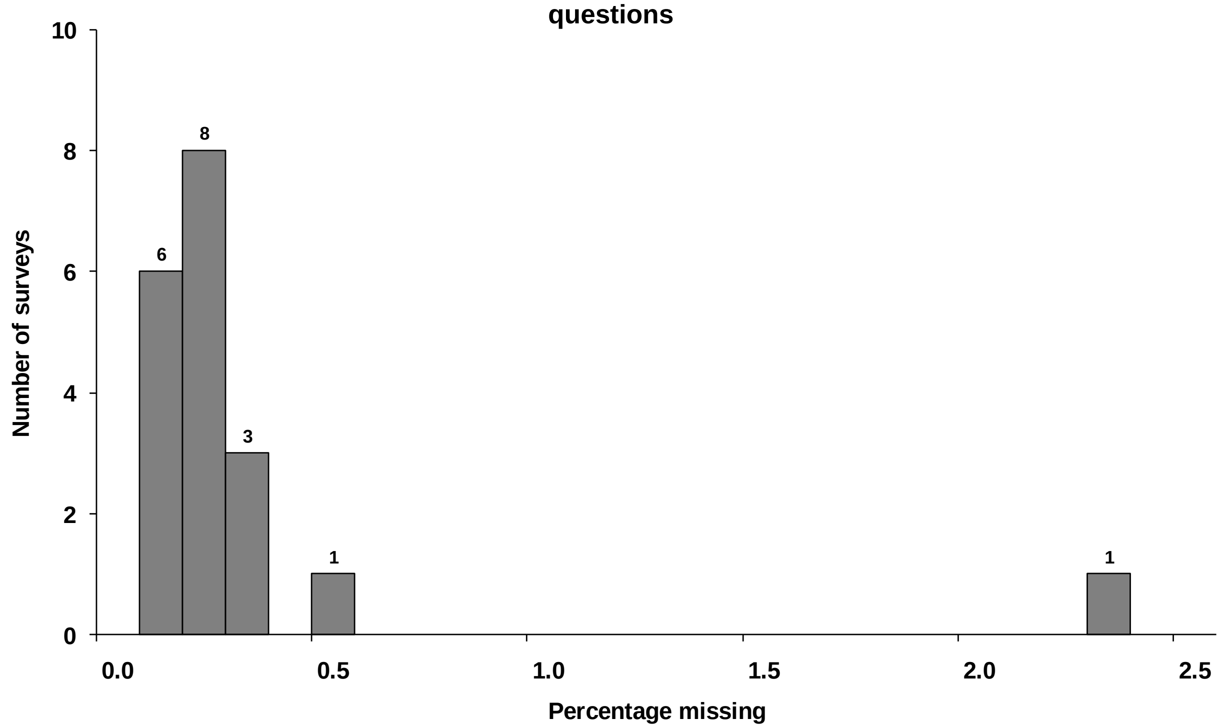


Figure 3.13 Distribution of the percentages of cases that are missing data on liquids and foods in the past seven days (m40), all DHS surveys 1993-2003 that included these questions

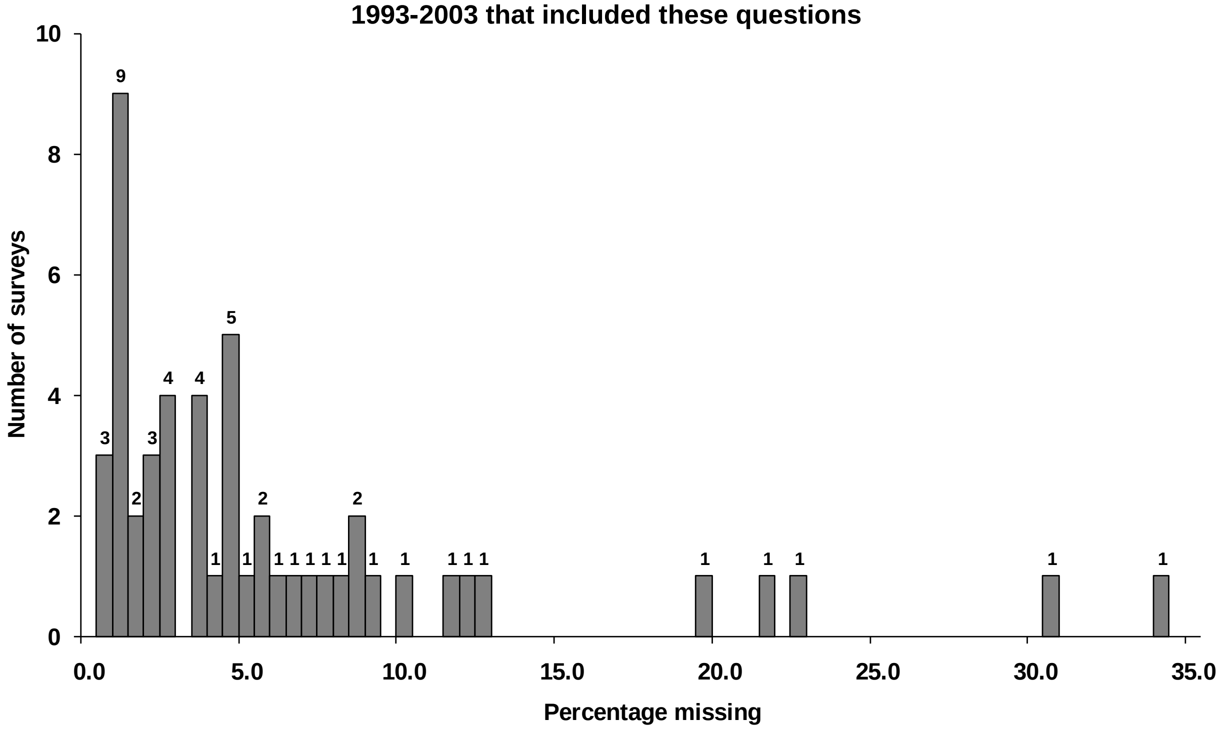


Table 3.12 Surveys that are missing at least 2 percent of responses about months of breastfeeding, at least 2 percent of responses about liquids in the first three days after birth, or at least 5 percent of responses about liquids and foods in the past seven days, all DHS surveys 1993-2003

Survey	Months of breastfeeding (m5); at least 2.0 percent missing	Liquids in first three days (m55); at least 2.0 percent missing	Liquids and foods in past seven days (m40); at least 5.0 percent missing
Benin 1996			7.2
Brazil 1996			21.1
Central Africa Republic 1994/95			19.4
Comoros 1996	2.0		8.6
Dominican Republic 1996			8.1
Dominican Republic 1999			10.0
Dominican Republic 2002			5.8
Guinea 1999	3.9		
Kazakhstan 1995			11.6
Kazakhstan 1999			7.6
Mali 1995/96	3.1		8.0
Mali 2001	3.6		
Mozambique 1997	4.0		30.0
Mozambique 2003	2.9		6.3
Nicaragua 1997/98			33.6
Nigeria 1999	5.7		22.1
Peru 2000		2.3	
South Africa 1998	2.8		
Togo 1998			5.2
Uganda 1995			6.8
Zambia 1996			12.5
Zimbabwe 1994			5.4
Zimbabwe 1999			11.1
All surveys	**0.9**	**0.3**	**6.5**

The relationship between the incidence of missing cases and the set of six covariates is different for these three measures. For the eight surveys that have at least 2.0 percent of cases that are missing or inconsistent on months of breastfeeding, such responses are related to two covariates with a pseudo-R^2 that is greater than .01 and statistically significant: age of child and number of index children in the window. The patterns are shown in Figures 3.14 and 3.15. As might be expected, recall is less accurate if the birth was longer ago or if there were more children in the window. These two characteristics are related; if a child is older, s/he is more likely to have a younger sibling.

Figure 3.14 Percentage of responses to months of breastfeeding (m5) that are missing or inconsistent in the eight surveys with more than 2 percent missing, by age of child

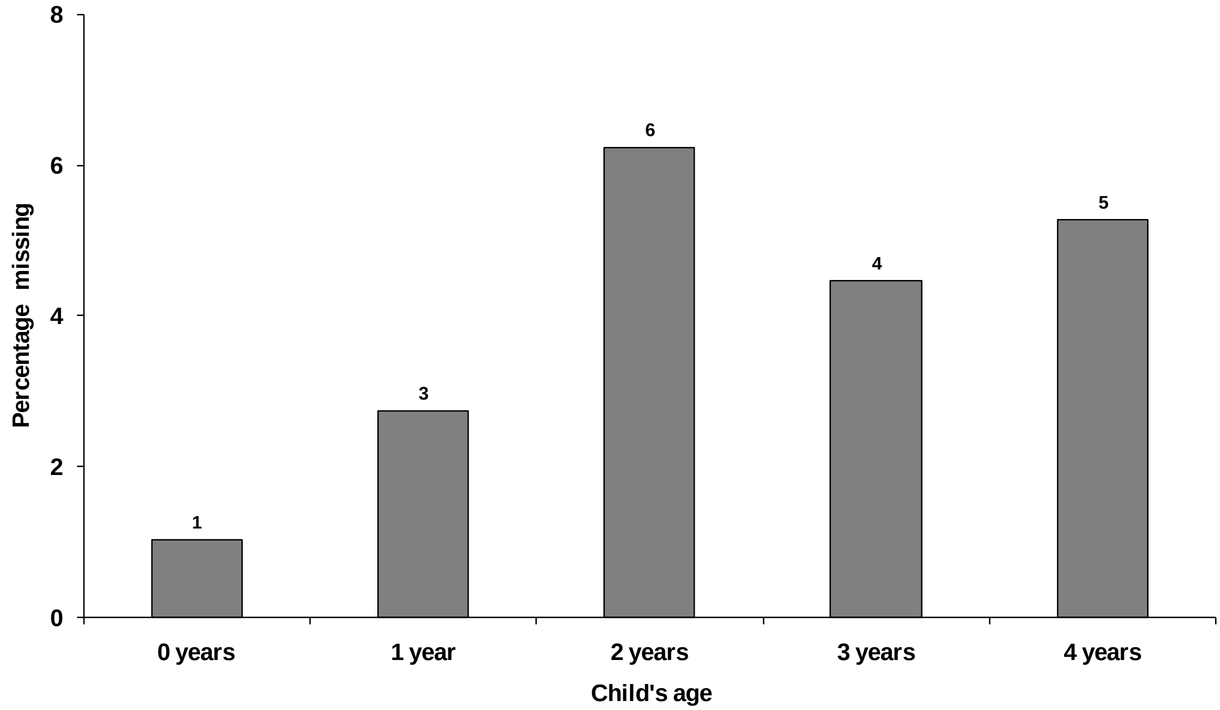
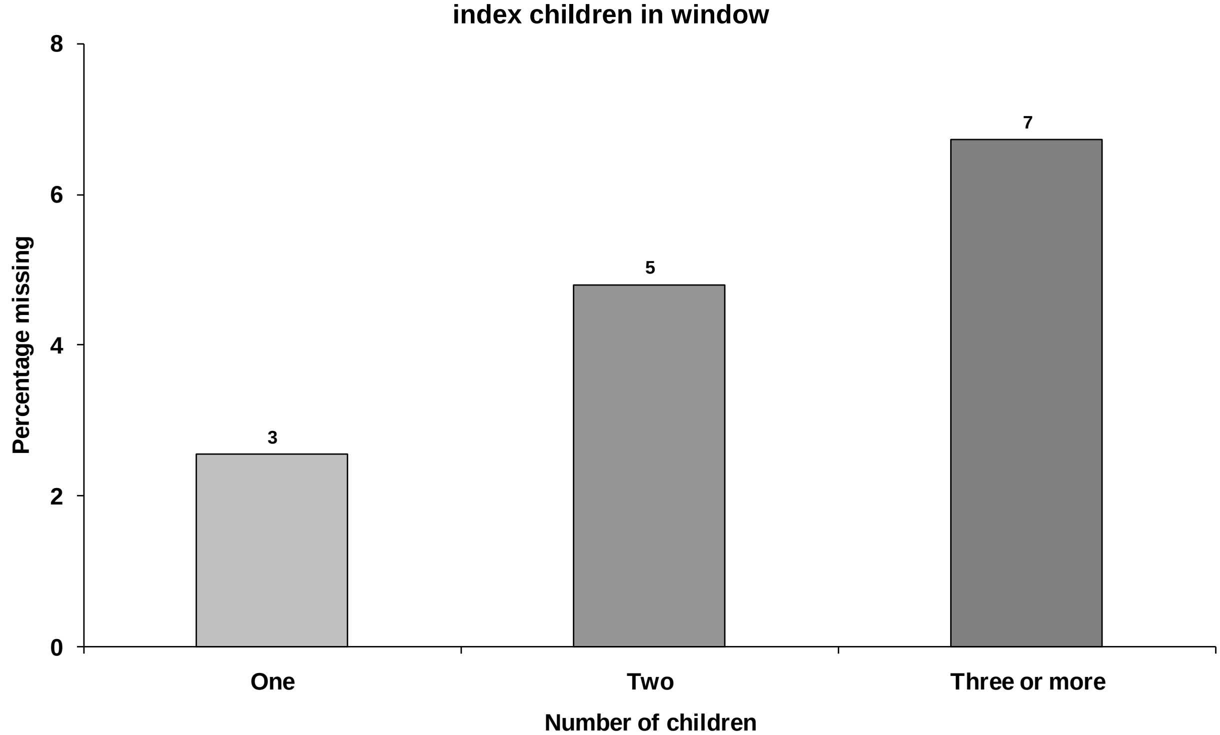


Figure 3.15 Percentage of responses to months of breastfeeding (m5) that are missing or inconsistent in the eight surveys with more than 2 percent missing, by number of index children in window



For the single survey, Peru 2000, that has more than 2.0 percent of cases missing on liquids in the first three days after birth, missing is related (significantly and with pseudo-R^2 greater than .01) to three covariates: type of place of residence, mother's level of education, and number of children in the window. The pattern is the reverse of what might be expected. The data are more likely to be missing in urban areas than in rural areas, for mothers with secondary or higher levels of education, and for women with only one child in the window.

Peru 2000 was the only survey that allowed a "don't know" response (code 8) to the question about liquids in the first three days after birth. Our general practice is to pool "don't know" with "not stated," and in this survey the "don't know" response is much more prevalent than the "not stated" response, with 190 and 12 cases, respectively. The categories for which "don't know" is a more common response are those in which alternatives to breast milk are more readily available. We speculate that the response "don't know" reflects uncertainty about the actual date when the milk substitute was first used. She may be unsure as to whether it was first used in the first three days after birth, or soon after that interval, especially if the child was born in a hospital and may have been given liquids by hospital staff. For such births, "don't know" would be a plausible response to the question.

The survey for Peru 2000 is instructive because it shows that the "don't know" response may be worth including in surveys. But if it is included, it should be expected that it may be used more often by women who used breast milk substitutes fairly soon after the birth.

Despite the much higher level of missing on questions about liquids and foods in the past seven days, a dataset consisting of the 19 surveys with at least 5.0 percent missing on this variable did not show any significant relationships between missing and the covariates.

We will look in more detail at heaping on stated duration of breastfeeding (m4). Responses to the question about duration of breastfeeding are well known for heaping at multiples of six months and, to a lesser degree, multiples of three months. Although such heaping is found in virtually all surveys, it is far worse in some than in others. Figure 3.16 illustrates the wide degree of variation. For three surveys (Cameroon 1998, Philippines 1993, and Nepal 1996), Myers' Index is less than 20, and for two surveys (Bangladesh 1996/97 and 1999/2000) it is greater than 60.

The most extreme example of heaping on stated duration of breastfeeding, for Bangladesh 1999/2000, is shown in Figure 3.17. More than two-thirds of the responses were at multiples of six months, and fully one-quarter of responses were at 24 months. When heaping approaches this level, the stated response probably has little value and should be replaced with current status information. However, most of the surveys have a Myers' Index less than half that of the Bangladesh 1999/2000 survey, and for them we would expect little difference between the two approaches as described earlier for postpartum amenorrhea and abstinence.

Figure 3.16 Myers' Index for duration of breastfeeding (m4), all DHS surveys 1993-2003

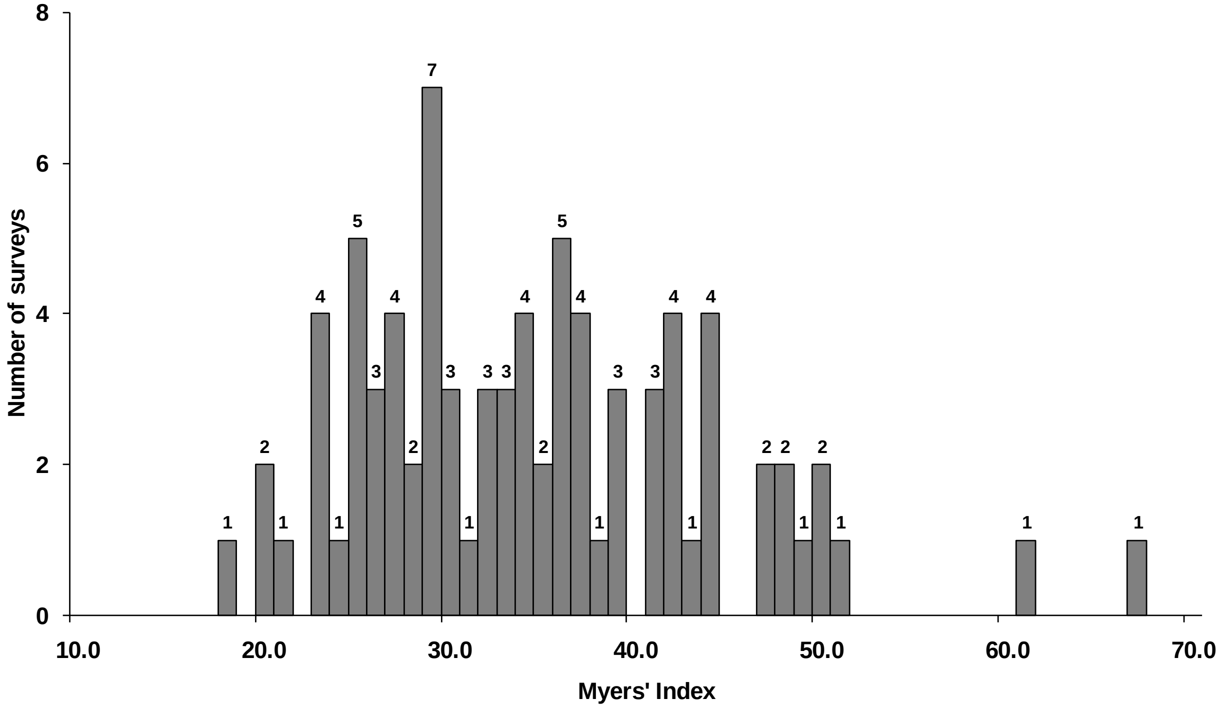
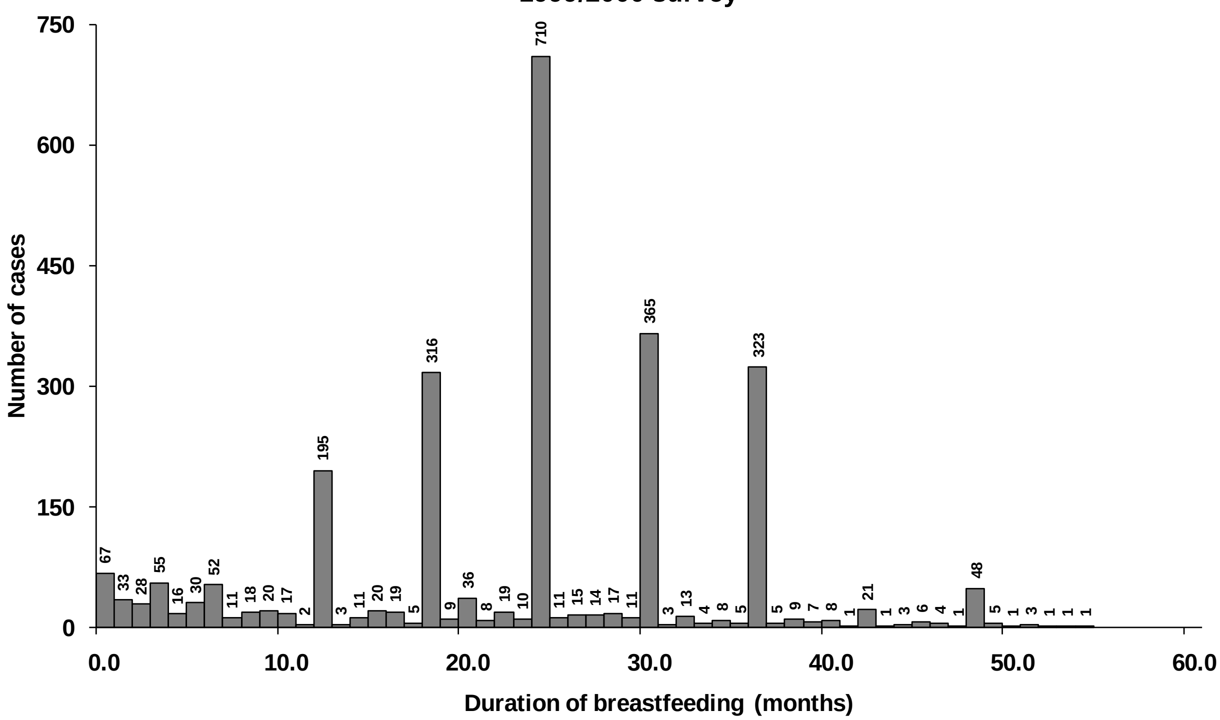


Figure 3.17 Distribution of stated duration of breastfeeding (m4) for the Bangladesh 1999/2000 survey



Because the level of unusable or missing cases is highest for the questions about liquids and foods in the last seven days, it is worthwhile to separate them into the two components: “don’t know” and “not stated.” For each of the many individual liquids and foods that are asked about in the m40 block of questions, code 8 is used for “don’t know” and code 9 for “not stated.”

In the consolidated file, the questions about liquids and foods in the last seven days were to be asked about 275,719 children (unweighted). We summarize the overall response as “don’t know” if code 8 was given for *any* of the specific items and as “not stated” if code 9 was given for *any* of the specific items.

Because the specific items were constructed from a checklist in the questionnaire, the correct coding would be completely consistent, so that code 8 should be given for all items (or for none) and code 9 should be given for all items (or for none). In practice, this was not always the case. A total of 14,342 children received codes 8 or 9, and were counted as “missing” in Table 3.12 and the discussion above. However, 85 of those children received a mix of codes 8 and codes 9 on the different items, which should never happen. Of those children, 21 were in the Mozambique 1997 survey, 17 were in the Nigeria 1999 survey, and the remaining 47 were scattered across 20 other surveys, with one to nine cases per survey. It is puzzling that these 85 cases were able to evade forced consistencies during data processing.

Putting these 85 cases aside, because it is not clear whether they should have been coded 8 (for all items) or 9 (for all items), there were a total of 6,619 “don’t know” (code 8) cases and 7,638 “not stated” cases out of the remaining 275,719 + 85 = 275,634 children. This is a fairly even split; 46 percent were “don’t know” and 54 percent were “not stated.”

When b16 is available and it is known that the mother and child do not live in the same household, the level of missing or unusable responses, as well as the balance between “don’t know” and “not stated” responses, is dramatically different. As noted in Section 3.1, b16 was coded 0 for 6,131 children. A substantial number of these children, 4,611, were in surveys that did not include the m40 block of questions. Of the 1,520 children for whom b16 = 0 and the questions about liquids and foods in the past seven days were asked, 417 children (27 percent) were given usable responses, 1,043 (69 percent) were given the “don’t know” response, and 60 (4 percent) were given the “not stated” response. For the 104,054 children who lived with their mother and for whom the questions applied, the corresponding figures were 102,148 (98.2 percent) usable responses, 1,452 (1.4 percent) “don’t know” responses, and 454 (0.4 percent) “not stated” responses.

It would be possible to compare the 417 usable responses for the children with b16 = 0 with the usable responses for children who were known to live in the same household as their mother. We shall not do so, however, because the number of cases is too small for a comparison that should take account of the context and the age of the child. However, it appears likely that we should have less confidence in these responses. Based on that likelihood and the high percentage of explicit “don’t know” responses, we suggest that questions about liquids and foods during the past week should not be asked if the mother and child reside in different households.

3.3 Immunization Status

DHS surveys routinely obtain information about nine childhood vaccinations. The vaccinations and the schedule for obtaining them as recommended by the World Health Organization (WHO) for African countries[10] are as follows:

- one vaccination against tuberculosis (referred to as the BCG or Bacillus Calmette-Guerin vaccination), recommended at birth (h2);
- four oral polio vaccinations (Polio 0, 1, 2, 3), recommended at birth, 6 weeks, 10 weeks, and 14 weeks (h0, h4, h6, h8);
- three vaccinations against diphtheria, pertussis, and tetanus (DPT 1, 2, 3), recommended at 6 weeks, 10 weeks, and 14 weeks (h3, h5, h7); and
- one vaccination against measles, recommended at age nine months (h9).[11]

This section will include all children in the consolidated child file, even those who are too young to have received a specific vaccination. The interest here is not in whether or when a vaccination was received, but in whether that information is missing.

Ideally, vaccinations and their dates are recorded on a health card for each child. At the beginning of the section of the questionnaire that deals with vaccinations, there is a question about whether the child has a health card (h1). The possible codes for the response, and the overall weighted percentages for each of these codes in the surveys conducted from 1993 through 2003 are as follows:

0: No card; never had one	16.1 percent
1: Yes, and the card is shown to the interviewer	50.1 percent
2: Yes, but the card is not shown to the interviewer	28.3 percent
3: The child had a card but no longer has one	5.1 percent
9: Missing	0.4 percent

The health questions, including h1, were omitted in the Senegal 1997 survey. Otherwise, h1 is mandatory for all surviving children in the window in all surveys and there are no cases with code “.”. The level of missing for the question on the health card is very low.

Information about each of the nine vaccinations is then obtained in a sequence of two questions. The first question (coded as h0, h2, h3, h4, h5, h6, h7, h8, h9) asks whether the child has had the vaccination and has the following possible codes:

0: No vaccination
1: Yes, and vaccination date is given on card
2: Yes, reported by mother
3: Yes, vaccination marked on card but no date
8: Don’t know
9: Missing

[10] See http://www.afro.who.int/afropac/vpd/schedule.html; other public health organizations, such as the U.S. Centers for Disease Control and Prevention (CDC), recommend somewhat different schedules.
[11] For Latin America, the Pan-American Health Organization recommends that the measles vaccination be given at age 12-14 months.

Second, the dates of these vaccinations, if given on the health card, are entered. These are coded by day, month, and year (for example, as h0d, h0m, h0y).[12]

Some surveys completely omit questions about one or more of the vaccinations, producing a code of ".". For example, questions about BCG (h0) were omitted for Armenia 2000; Bangladesh 1993/94, 1996/97, and 1999/2000; Bolivia 2003/04; Cote d'Ivoire 1994; Egypt 1995 and 2000; Ethiopia 2000; Ghana 1993; Indonesia 2002; Kenya 1993; Nepal 1996; Nicaragua 1997/98 and 2001; Philippines 1993, 1998, and 2003; Senegal 1997; Turkey 1993 and 1998; Vietnam 2002; Zambia 1996; and Zimbabwe 1994 and 1999. *All* children in those surveys have a code "." for the BCG vaccination. Otherwise, *no* children have code "." for that vaccination. Similarly for other vaccinations, "." is consistently used for "not applicable." Senegal 1997 was the only survey that omitted all of the questions about vaccinations and the health card.

To assess the level of incomplete data, we combine codes 8 and 9, "don't know" and "not stated." The level of missing data is generally very low. When all surveys are combined, with equal weight for each survey, the overall percentage of responses that are missing ranges from 0.4 percent (on h0, h1, and h2) to 1.5 percent (on h5, h7, and h9).

As would be expected, there is a high correlation between levels of missing on related questions. The four polio items are inter-correlated with one another at a level of approximately 0.6. Polio2 and Polio3 are both missing or both present for all children. The three DPT items are inter-correlated at a level of about 0.8. DPT2 and DPT3 are both missing or both present for virtually all children. Those two sets of items (Polio and DPT) also tend to be positively correlated with one another, at levels in the range of 0.3 to 0.4. In order to reduce repetition, we focus on the levels of missing for five items rather than all ten: having a health card (h1), BCG (h2), the first polio vaccination (h0), the first DPT vaccination (h3), and the measles vaccination (h9).

Table 3.13 lists the surveys in which the percentage missing is at least 1.0 percent for health card, BCG, the first polio vaccination or the first DPT vaccination, or at least 2.0 percent for the measles vaccination. Values below these thresholds are blanked out to make the table easier to read.

Only six surveys have more than 1.0 percent missing on the initial question about having a health card, and the worst of those, Nigeria 1999, only reaches a level of 4.1 percent. Only six surveys are missing more than 1.0 percent on the BCG vaccination, with a maximum of only 1.8 percent for Turkey 1998. Four surveys exceed the 1.0 percent level for the first Polio vaccination; all are Phase 3 surveys, and the maximum is only 1.8 percent, for Kazakhstan 1995. The DPT and measles vaccinations have the highest levels of missing. Eighteen surveys exceed the 1.0 percent level for the first DPT vaccination; Turkey 1998 and Kazakhstan 1995 reach the highest levels, with 4.3 percent and 7.3 percent, respectively. A higher threshold is used for the measles vaccination, because it has the highest levels of missing. Twenty-three surveys have at least 2.0 percent missing on this vaccination, with the highest levels for Haiti 1994/95, Kazakhstan 1995, Turkey 1998, and Burkina Faso 1998/99, at 5.0 percent, 5.1 percent, 5.3 percent, and 6.8 percent, respectively. The only surveys that appear more than twice in Table 3.13 are Ghana 1993, Haiti 1994/95, Kazakhstan 1995, Mali 2001, Nigeria 1999 and 2003, South Africa 1999, and Turkey 1998.

[12] In the consolidated file of children, 13 to 40 percent have code 2 for specific vaccinations, and 89 to 98 percent of those children have codes 0, 2, or 3 for h1. In these cases, the mother's statement that the child had the vaccination is accepted, but the mother is never asked to provide a date. Dates must come from a health card.

Table 3.13 Surveys in which 1.0 percent or more of weighted responses to questions about a health card, BCG, Polio (first dose), or DPT (first dose), or more than 2.0 percent of weighted responses to questions about measles vaccinations, are coded “don’t know” or “not stated,” all DHS surveys 1993-2003 that included these questions

Survey	Card	BCG	Polio	DPT	Measles
Benin 2001					3.2
Bolivia 1994	1.1				
Bolivia 1998				2.0	
Burkina Faso 1998/99		1.4			6.8
Burkina Faso 2003				1.0	2.3
Colombia 2000					2.1
Comoros 1996					2.4
Dominican Republic 1996					2.3
Ethiopia 2000					2.3
Gabon 2000				1.4	
Ghana 1993		1.6		1.9	2.5
Ghana 1998				1.6	2.2
Guatemala 1995	1.0				
Guatemala 1998/99	1.2				
Guinea 1999					2.3
Haiti 1994/95			1.5	1.4	5.0
Haiti 2000				1.3	3.1
India 1998/99				1.0	2.7
Indonesia 1994			1.1		
Indonesia 1997			1.2		
Kazakhstan 1995			1.8	7.3	5.1
Mali 1995/96				1.1	2.2
Mali 2001	1.6	1.0		1.8	3.6
Namibia 2000					2.0
Nicaragua 1997/98					2.5
Nigeria 1999	4.1			1.7	2.9
Nigeria 2003	1.3	1.2		1.8	
Rwanda 2000				1.3	
South Africa 1998		1.0		2.2	2.6
Tanzania 1996					2.1
Turkey 1993				1.1	2.8
Turkey 1998		1.8		4.3	5.3
Uzbekistan 1996				1.6	

In order to see if there is a non-random pattern to the missing in those surveys where it is more common, we apply the logit regression strategy with the six key covariates to three specific surveys and questions: missing on the question about the health card (h1) in Nigeria 1999, missing on the first DPT vaccination (h3) in Kazakhstan 1995, and missing on the measles vaccination (h9) in Burkina Faso 1998/99. We will only present the results for covariates that have a highly significant association with non-response ($p < .01$) and that account for a non-negligible proportion of variation in the outcome (pseudo-$R^2 > .01$). All the results are weighted and adjusted for clustering.

Figures 3.18 and 3.19 show the levels of missing (including “don’t know”) for the question about the health card (h1) for the Nigeria 1999 survey for the only covariates that reach the specified level of importance: mother’s level of education and the number of children in the window. Both show the expected monotonic pattern, with highest levels of missing for women with no education or three or more children in the window. Figure 3.20 shows the distribution of responses to the question, first as observed, and second as imputed on the basis of the pattern of missing (using all the covariates, not just those that are significant). The greatest difference is that the missing cases are estimated to be much more likely to have no card (49.1 percent, versus 39.8 percent for the observed cases). The index of dissimilarity between the two distributions is 9.8 percent. However, the index of dissimilarity between the observed distribution of responses to h1, and the combination of observed and estimated, is only 0.4 percent, which is negligible.

Figure 3.18 Percentage of cases missing on the question about the health card (h1) in the Nigeria 1999 survey, by mother's level of education (pseudo-R^2 = .022)

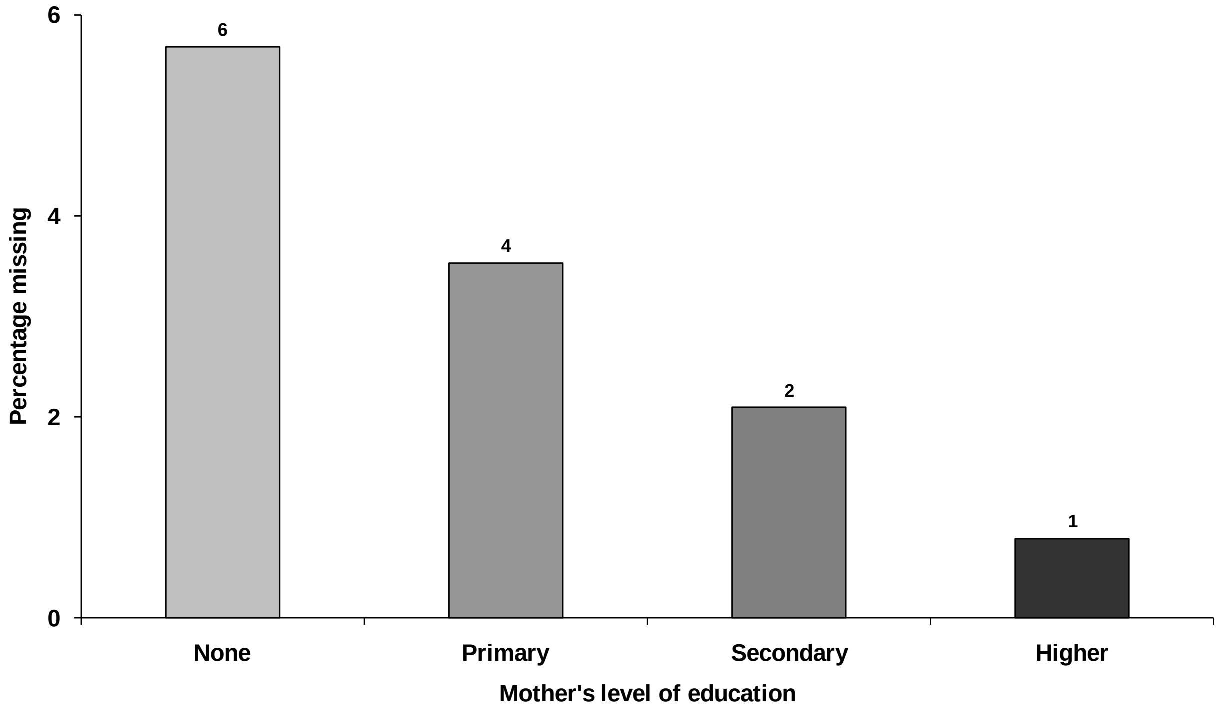


Figure 3.19 Percentage of cases missing on the question about the health card (h1) in the Nigeria 1999 survey, by number of index children in the window (pseudo-R^2 = .020)

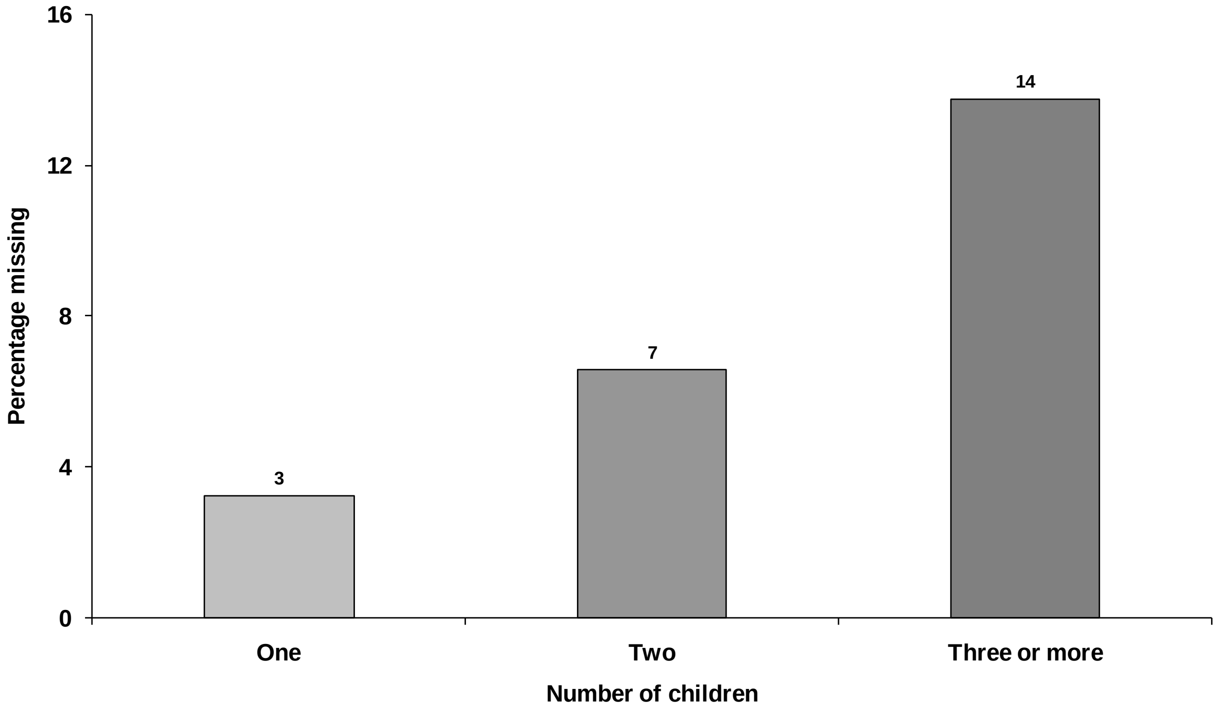
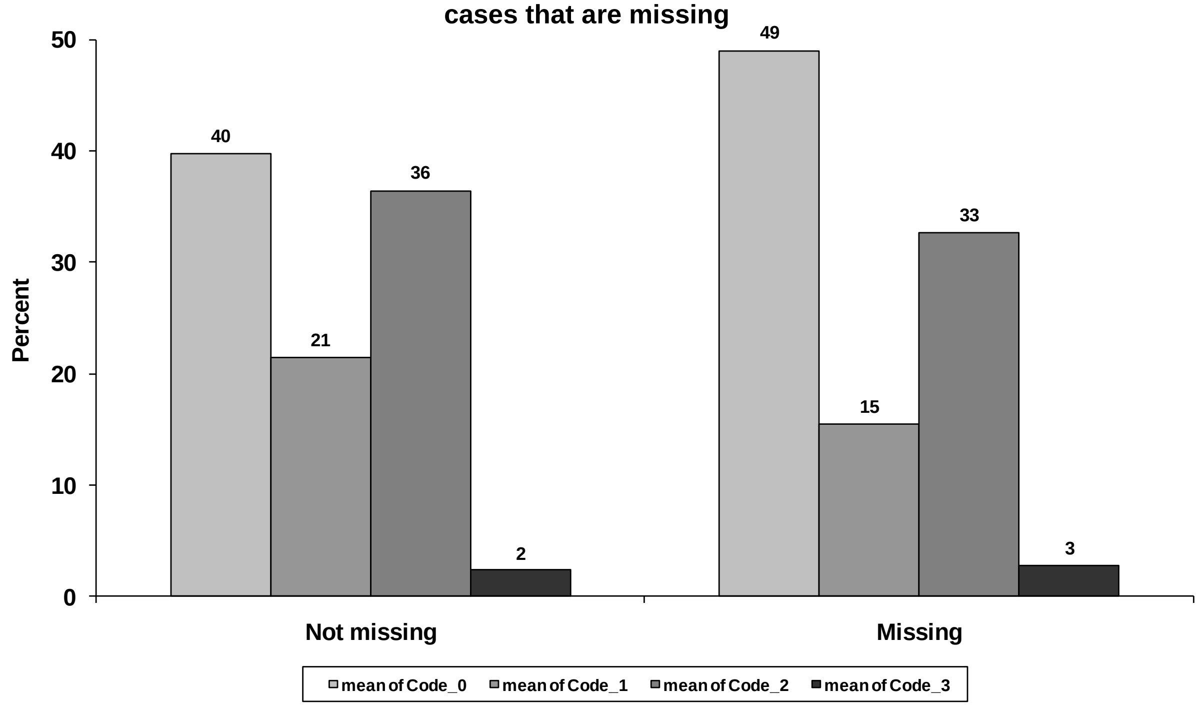


Figure 3.20 Distribution of responses to the question about the health card (h1) in the Nigeria 1999 survey, observed for the cases that are not missing and estimated for the cases that are missing



Notes: Code_0 is "No vaccination;" Code_1 is "Yes, vaccination date given on card;" Code_2 is "Yes, vaccination date given by mother;" Code_3 is "Yes, vaccination date given on card but no date."

When the covariates are checked for a relationship to the pattern of missing on the first DPT vaccination (h3) in Kazakhstan 1995, we find no relationships that satisfy the dual requirements for magnitude and significance. Only one of the covariates is sufficiently related to the incidence of missing on the question about measles vaccination (h9) in the Burkina Faso 1998/99 survey: the age of the index child. The response is least likely to be missing if the child is 0 years of age and most likely if the child is age 3 or 4 (Figure 3.21). This is the pattern that would be expected, because if the vaccination is received, it is normally at age 0. As Figure 3.22 shows, there is some difference between the observed and estimated distributions of responses for the non-missing and missing cases, respectively. The index of dissimilarity to describe the difference between these distributions is 5.70 percent. However, the difference between the observed and combined distributions is only 0.39 percent, which is negligible.

Finally, this section will look at the association between the responses to the question about having a health card (h1) and the four indicators of specific vaccinations (h2, h0, h3, and h9). As listed earlier, there are five possible responses to h1 and six possible responses to the specific indicators, for a total of 30 combinations. Because of the logical connections between the questions, implemented with the skip pattern in the questionnaire, 10 combinations do not occur. For example, if the respondent has a card and shows it to the interviewer (h1 = 1, the most desirable outcome for h1), then "don't know" and "not stated" (codes 8 and 9 for specific vaccinations) are impossible. Of the remaining 20 combinations, ten involve "not stated" on h1 and/or "don't know" or "not stated" on a specific vaccination question. Tables 3.14.1 to 3.14.4 give the weighted percentages of cases in these 10 combinations for the four specific vaccinations. (The marginal percentages for these tables may not agree with one another or with the overall percentages given at the beginning of this section because some surveys omitted some vaccinations.) The irrelevant cells in these tables are blanked out. The column for h1 = 1 is omitted. Total percentages for rows and columns are retained to permit a check for proportionality of the cell percentages to these totals (the row totals omit the column for h1 = 1). Expected percentages under the null hypothesis of independence are given in parentheses.

Figure 3.21 Percentage of cases missing on the question about measles vaccination (h9) in the Burkina Faso 1998/99 survey, by age of the index child (pseudo-R^2 = .016)

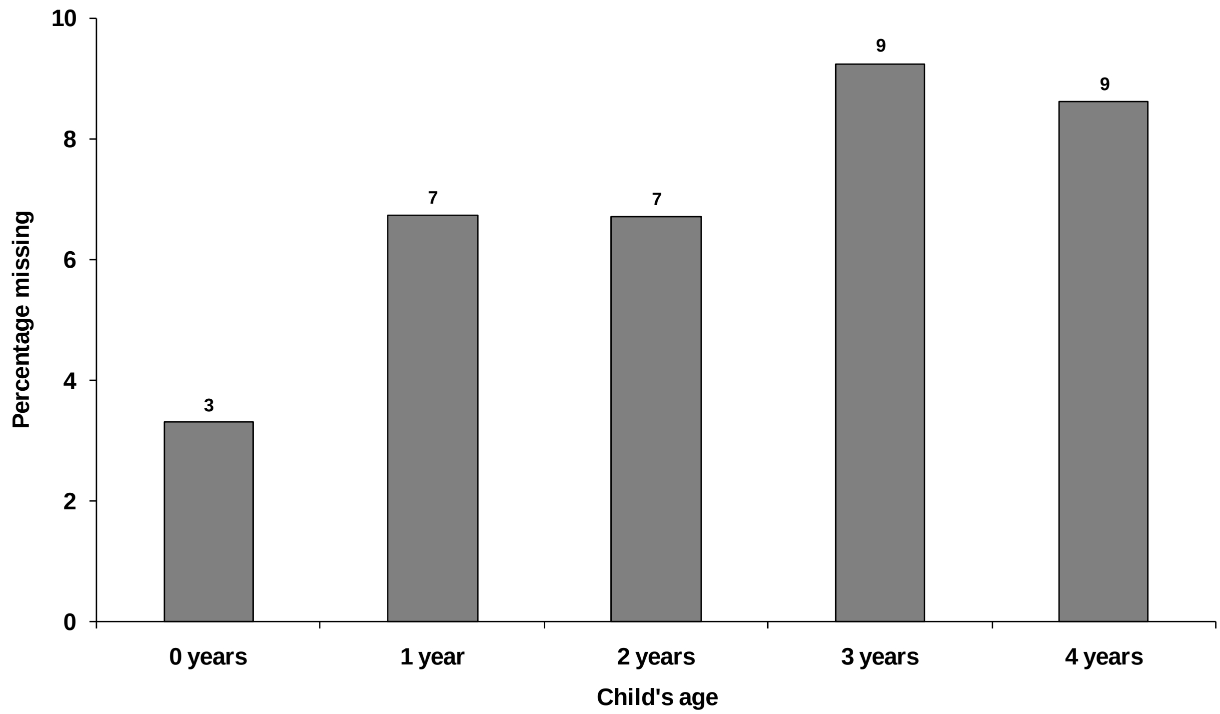


Figure 3.22 Distribution of responses to the question about measles vaccination (h9) in the Burkina Faso 1998/99 survey, observed for the cases that are not missing and estimated for the cases that are missing

Tables 3.14.1 to 3.14.4 show low levels of missing throughout, but they also show that “don’t know” (code 8) is three to four times as common as “not stated” (code 9). The previous analysis combined these two codes. It is evident that the findings pertain primarily to the “don’t know” code. The “don’t know” response, in turn, can be traced predominantly to respondents who say they have a health card, but are unable to show it to the interviewer (“yes, but not seen”). About half of the “don’t know” or “not stated” cases for specific vaccinations can be attributed to this particular cell. The percentage of cases in this combination is somewhat more than expected, given the row and column percentages, for polio, DPT, and measles vaccinations (h0, h3, and h9) but not for BCG (h2). Otherwise, the percentages in Tables 3.14.1 to 3.14.4 are small and close to what would be expected. There is no clear reason to disaggregate these tables or look at specific surveys.

Table 3.14.1 Weighted percentages of responses in combinations of “not stated” for the health card (h1), and “don’t know” or “not stated” for the BCG vaccination (h2), all DHS surveys 1993-2003 that included these questions

Had BCG vaccination	No card	Yes, but not seen	No longer has card	Missing	Total
No				0.52	31.33
Reported by mother				0.17	67.94
Don’t know	0.20 (0.20)	0.33 (0.35)	0.08 (0.06)	0.00 (0.00)	0.62
Not stated	0.04 (0.04)	0.05 (0.07)	0.01 (0.01)	0.01 (0.00)	0.12
Total	**32.22**	**56.78**	**10.29**	**0.70**	**100**

Note: Respondents who had a health card and showed it to the interviewer are omitted. Expected percentages are given in parentheses.

Table 3.14.2 Weighted percentages of responses in combinations of “not stated” for the health card (h1), and “don’t know” or “not stated” for the first polio vaccination (h0), all DHS surveys 1993-2003 that included these questions

Had polio 0 vaccination	No card	Yes, but not seen	No longer has card	Missing	Total
No				0.72	78.20
Reported by mother				0.08	21.09
Don’t know	0.13 (0.19)	0.37 (0.31)	0.06 (0.01)	0.00 (0.00)	0.56
Not stated	0.03 (0.05)	0.09 (0.08)	0.01 (0.01)	0.01 (0.00)	0.15
Total	**34.13**	**55.16**	**9.89**	**0.82**	**100**

Note: Respondents who had a health card and showed it to the interviewer are omitted. Expected percentages are given in parentheses.

This section has found a level of missing for the questions on vaccinations that is generally quite low. A selective examination of the surveys with the highest incidence of missing cases indicates that the mother's education, number of children in the window, and the elapsed time since the normal age at vaccination are sometimes related to the incidence of missing, but the impact of missing cases on the overall distribution of responses is negligible. Most of the missing responses come from the "don't know" code and from the cases for which the respondent claims to have a health card but is unable to show it to the interviewer.

Table 3.14.3 Weighted percentages of responses in combinations of "not stated" for the health card (h1), and "don't know" or "not stated" for the first DPT vaccination (h3), all DHS surveys 1993-2003 that included these questions

Had DPT 1 vaccination	No card	Yes, but not seen	No longer has card	Missing	**Total**
No				0.54	34.36
Reported by mother				0.15	64.04
Don't know	0.30 (0.44)	0.91 (0.78)	0.17 (0.14)	0.00 (0.01)	1.38
Not stated	0.05 (0.07)	0.14 (0.12)	0.02 (0.02)	0.01 (0.00)	0.22
Total	**32.22**	**56.78**	**10.29**	**0.70**	**100**

Note: Respondents who had a health card and showed it to the interviewer are omitted. Expected percentages are given in parentheses.

Table 3.14.4 Weighted percentages of responses in combinations of "not stated" for the health card (h1), and "don't know" or "not stated" for the measles vaccination (h9), all DHS surveys 1993-2003 that included these questions

Had measles vaccination	No card	Yes, but not seen	No longer has card	Missing	Total
No				0.57	45.12
Reported by mother				0.11	51.86
Don't know	0.46 (0.74)	1.52 (1.31)	0.32 (0.24)	0.01 (0.02)	2.31
Not stated	0.10 (0.23)	0.52 (0.40)	0.08 (0.07)	0.02 (0.00)	0.71
Total	**32.22**	**56.78**	**10.29**	**0.70**	**100**

Note: Respondents who had a health card and showed it to the interviewer are omitted. Expected percentages are given in parentheses.

4 Anthropometric Measurements

DHS makes a considerable investment in accurately measuring the height and weight of all women age 15-49 years, as well as all surviving children born in the window. The purpose is to infer, at least at an aggregate level, the nutritional status of women and children.

This chapter will describe and assess the measurements of height and weight and also some of the indexes that are constructed from them. Sections 4.1 and 4.2 concern women age 15-49 but are limited to women who are mothers with at least one index child in the window for child health questions. Section 4.1 concerns primarily the measurements of mothers' height, but makes some references to weight and raises many issues that affect the other sections of this chapter. Section 4.2 concerns mothers' weight and its relationship to height. Section 4.3 considers the heights and weights of children.

During the 1993-2003 interval, DHS used two different strategies to obtain anthropometric data.[13] About two-thirds of the surveys, and virtually all of those early in the interval, collected the data during the survey of women and recorded them in the women's questionnaire. About one-third of the surveys, and especially the later surveys, collected the data as part of the household survey and recorded them in the household schedule. When the household approach was used, the information was appropriately transferred to the women's records during data processing, using the correspondence between line numbers in the household survey and the identification codes and child index numbers in the women's data. The two strategies for data collection have some subtle differences with respect to the coding and interpretation of eligibility for measurement, missing data, etc., that will be mentioned where relevant to this report.

The main reason why DHS changed from a woman-based strategy to a household-based strategy for obtaining children's heights and weights is that the earlier approach omitted entirely the measurement of children who were maternal orphans, a substantial and important group of children in countries with high levels of HIV/AIDS. Because the present report organized the data around mother-child combinations, omitting children without mothers and women without children, it is effectively weighted toward the earlier, woman-based strategy.

4.1 Measurements of Maternal Height

This report's limitation to the mothers of index children, and the exclusion of other women age 15-49, was made largely because of data processing considerations. The 81 surveys include 385,586 mothers of index children, a large number in itself, and a file of all women would have been approximately twice as large. The construction of the consolidated files included linking children with their mothers, with the perspective that the health of the children was of primary interest. In terms of data quality, we expect that the height and weight measurements of women without index children would be consistent with the women analyzed here.

For women, the standard variables reviewed in this section are as follows:

v437 Respondent's weight (kilos-1d)
v438 Respondent's height (cms-1d)
v439 Ht/A Percentile (resp.)

[13] The author is grateful to Jerry Sullivan for providing details about the two data collection procedures.

v440 Ht/A Standard deviations (resp.)
v441 Ht/A Percent ref. median (resp.)

Weight of the woman, v437, is measured in kilograms (kg), coded in tenths of a kilogram. Height, v438, is measured in centimeters (cm), coded in the file in tenths of a centimeter (i.e., millimeters). The measurements of height show considerable heaping.

The heaping on preferred digits, in this case primarily whole and half centimeters, with some additional heaping at multiples of five and ten centimeters, is indicated by the deviation of the frequency with which each final digit (0 through 9) occurs from what would be expected in the absence of any digit preference. Each digit should occur 10 percent of the time, but we observe that final digit 0 occurs 19.36 percent of the time and final digit 5 occurs 13.49 percent of the time in the consolidated file.[14] Heaping is greatest for the Nigeria 1999 survey, for which these percentages are 22.95 percent and 18.03 percent, respectively. Low deviations from 10 percent are found for the surveys in Bangladesh, India, and Nepal, which is noteworthy. In the context of age, these countries are well-known for high levels of heaping, but that does not carry over into the measurements of height in the fieldwork.

It is important to bear in mind that the units of measurement for height, particularly for adult heights, are fairly small. Indeed, it is unreasonable to expect accuracy down to millimeters. Rounding to a centimeter, the bulk of the heaping for height, is not in itself a serious measurement issue. It may reflect on the overall quality of the care taken during the fieldwork, but this kind of rounding will certainly not have a serious impact on any kind of analysis.

If a case is missing on v437, it is given a code 9999; if missing on v438, it is given code 9999 (one case in the Rwanda 2000 survey had v438 = 9997; we have treated this as 9999). The Stata code ".", when it appears, virtually always (see the next two paragraphs) indicates that the measurement was omitted for an entire survey. As stated earlier, "." in Stata is equivalent to a blank in the ASCII version of the data and can be interpreted as "not applicable," whereas "9999" (or similar) is a pre-coded option for "not stated." Implausible values of v437 and v438 are not flagged, but lead to flagged codes for the constructed variables.

Of the 81 surveys assessed in this report, 15 (in nine countries) omitted weight and height measurements of women. These were the following surveys: Bangladesh 1993/94; Dominican Republic 1999 and 2002; Indonesia 1994, 1997, 2002; Namibia 2000; Philippines 1993, 1998, and 2003; Senegal 1997; South Africa 1998; Tanzania 1999; and Vietnam 1997 and 2002. For these 15 surveys, all cases were coded "." on v437, v438, and all the related constructed variables.

Contrary to expectation, three surveys were coded "." for a *subset* of cases. India 1998/99 had this code for 8.9 percent of women; Mali 1995/96 for 0.9 percent of women; and Nigeria 1999 for 0.1 percent of women. In these three surveys, the "." code was almost certainly used incorrectly and should have been "9999." We have looked at the distributions of "." and "9999" within clusters in these three surveys. It appears that in the India survey, v437 and v438 were completely omitted in a few clusters and were coded "." for all women in those clusters. In these clusters there may have been a decision to uniformly omit weight and height, for reasons that we do not know, in which case "." would have been the correct code. In most clusters in that survey, however, there appear to be some cases that should have been coded "9999" but were incorrectly coded ".".

[14] These percentages apply to the range $1280 \le v438 \le 1994$, which will be shown later to be the valid range of v438 for women age 18+. The calculation uses the sample weights, modified to give equal weight to each survey.

In the Mali 1995/96 survey, almost all clusters show consistent use of only "9999" or only ".". A handful of clusters show use of "9999" for some cases and "." for some cases. In the Nigeria 1999 survey, the two codes were apparently used correctly, except for one cluster in which "." was consistently used in place of "9999" and one other cluster in which both codes were used. The coding errors in these three surveys could have been avoided with better training or removed with field or computer edits, but they probably had no substantive effects. Our analysis will retain the original coding.

Of the 297,506 mothers in the consolidated file for which neither v437 nor v438 was coded ".", there are 8,221 cases for which both were coded "9999." There are also 524 cases in which weight was coded "9999" and height received another numerical code. These were scattered over many surveys; the only ones with 20 or more such cases were Mali 2001 (117 cases), India 1998/99 (81 cases), Rwanda 2000 (49 cases), Ghana 2003 (42 cases), and Brazil 1996 (23 cases). There are another 594 cases with the reverse pattern, missing height (v438 = 9999) but not weight. The only surveys with 20 or more such cases were Bolivia 1994 (85 cases), Brazil 1996 (76 cases), Malawi 2000 (43 cases), India 1998/99 (38 cases), Nicaragua 2001 (32 cases), Ghana 2003 (29 cases), and Peru 1996 (27 cases). The fact that the two possible combinations of missing and non-missing occur about equally often implies that any difficulties in taking the measurements are about the same for height and weight.

During the preparation of the standard recode files, indices are constructed that can be used to compare the measurements of women's height and weight with international norms. Three measures describe the woman's height relative to an international normative distribution. v439 identifies her percentile in the normative distribution; in the literature this is sometimes called HAP. v440 expresses her height as a z score, the number of standard deviations below or above the median normative height, sometimes called HAZ. v441 gives her height as a percentage of the normative median for women of her age. These three variables are coded 9999 or 99999 if v438 is coded 9999. The three variables are flagged (v439 = 9998, v440 = 9998, and v441 = 99998) for 933 cases because of implausible values of height.

We will first examine the incidence of unacceptable values of height and/or weight. Table 4.1 identifies those surveys in which at least 2.0 percent of cases were missing on weight or height. It gives the percentages missing, as well as the percentages flagged for implausible or inconsistent values on various measures. These include v439 through v441, the constructed variables that use height only, and also v442 through v444, v444a, and v445 through v446, other constructed measures that use both height and weight and that will be described further in Section 4.2. All estimates are weighted, with equal weight for each survey. It is clear that most deficiencies take the form of v437 = 9999 or v438 = 9999. In most surveys, relatively few cases—rarely more than 1 percent—are flagged.

To simplify the analysis, we construct an index that is 1 if a case is missing ("9999") on height or weight, or flagged as implausible or inconsistent on any of the constructed measures, and 0 otherwise. This will simply be referred to as the error index, and the percentage of cases with value 1 in a group will be referred to as an error rate.

Table 4.1 Weighted percentages of women missing on weight (v437 = 9999) or height (v438 = 9999) or flagged for inconsistent values, DHS surveys 1993-2003 with at least 2.0 percent of cases missing (code 9999) on v437 or v438

	Variable and code					
Survey	v437 9999	v438 9999	v439 9998	v442 99998	v444a 99998	v445 99998
Armenia 2000	2.5	2.4	0.0	0.0	0.0	0.0
Bangladesh 1996/97	2.9	2.9	1.0	2.5	2.5	1.4
Bolivia 1994	12.4	14.0	0.0	0.3	0.3	0.0
Bolivia 1998	4.5	4.4	0.2	0.7	0.7	0.2
Brazil 1996	6.3	8.7	0.2	0.3	0.3	0.1
Burkina Faso 1998/99	2.5	2.4	0.5	0.6	0.6	0.5
Cameroon 1998	6.6	6.7	0.1	0.4	0.4	0.3
Colombia 1995	5.7	5.5	0.2	0.5	0.5	0.3
Colombia 2000	3.5	3.5	0.0	0.2	0.2	0.0
Comoros 1996	4.7	4.5	0.1	0.6	0.6	0.1
Dominican Republic 1996	4.2	4.6	0.6	0.8	0.8	0.7
Egypt 1995	3.3	3.3	0.2	0.4	0.4	0.3
Gabon 2000	7.2	7.2	0.1	0.2	0.2	0.0
Ghana 1993	2.2	2.4	0.1	0.3	0.0	0.1
Ghana 2003	3.3	3.8	0.2	0.3	0.3	0.1
Guatemala 1995	4.7	4.5	0.1	3.2	3.2	0.0
Guatemala 1998/99	8.9	8.9	0.3	3.2	3.2	0.3
Guinea 1999	4.5	4.5	1.2	1.6	1.6	1.1
Kazakhstan 1999	51.7	51.7	0.0	0.2	0.2	0.2
Kenya 1993	5.4	5.4	0.2	0.5	0.0	0.3
Kenya 1998	2.7	2.4	0.1	0.3	0.3	0.0
Kenya 2003	4.7	4.9	0.0	0.1	0.1	0.1
Madagascar 1997	2.8	3.0	0.2	0.3	0.3	0.1
Mali 2001	4.7	4.0	0.3	0.3	0.3	0.2
Mozambique 1997	2.4	2.1	0.0	0.2	0.2	0.0
Mozambique 2003	4.0	4.0	0.1	0.3	0.3	0.0
Nicaragua 1997/98	2.9	3.0	0.2	0.6	0.6	0.3
Nicaragua 2001	2.1	2.7	0.1	0.3	0.3	0.0
Nigeria 1999	8.4	8.8	11.1	14.5	14.5	11.0
Nigeria 2003	2.0	2.1	0.0	0.2	0.2	0.0
Peru 1996	4.9	5.1	0.1	0.9	0.9	0.1
Peru 2000	3.4	3.4	0.0	0.7	0.7	0.0
Tanzania 1996	4.5	4.7	0.8	1.1	1.1	0.9
Turkey 1993	4.0	4.0	0.3	0.4	0.0	0.2
Turkey 1998	4.0	3.6	0.0	0.0	0.0	0.0
Uganda 1995	2.0	2.0	0.7	1.1	1.1	0.9
Uganda 2000/01	5.8	5.9	0.2	0.5	0.5	0.3
Zimbabwe 1994	2.1	2.3	0.3	0.4	0.4	0.3
Zimbabwe 1999	4.7	4.7	0.0	0.4	0.4	0.0
Total	**3.1**	**3.1**	**0.3**	**0.7**	**0.7**	**0.3**

4.1.1 Correlates of the Error Rate

The very high level of missing cases for weight and height in Kazakhstan 1999, more than 50 percent, is by design. According to the report on this survey (page 147), measurements were taken in one-half of the households. The selection of households for measurement appears to have been random. In a check for a pattern to the missing cases, none of the covariates approach statistical significance. There is also no significant variation across the 13 regions, ethnicity, or religion. These missing cases definitely should not be classified as errors.

We will briefly investigate possible factors behind other missing and flagged cases. Three surveys stand out as most problematic: Nigeria 1999, Bolivia 1994, and Guatemala 1998/99 (with error rates of 23.9 percent, 14.5 percent, and 12.2 percent, respectively). These are the surveys in which more than 10 percent of cases are missing, implausible, or inconsistent on weight and/or height. For each of these four surveys, the error index was regressed on categorical variables for type of place of residence, woman's age, woman's level of education, and number of interviewer visits, using logit regression with sampling weights and corrections for clustering. A variable will be judged to be an important covariate of errors if the pseudo-R^2 is at least .01 and the regression is statistically significant at the .01 level or better.

When the covariates are applied to Nigeria 1999, Bolivia 1994, and Guatemala 1998/99, the only significant ($p < .01$) covariation is with the woman's level of education in Guatemala (pseudo-$R^2 = .016$); women with no education are significantly more likely than other women to have some kind of problem with their report of weight and/or height. Education is statistically significant in Bolivia 1994 as well, but the pseudo-R^2 is only .009, which is less than the threshold of .010 generally used in this report. No other covariate is significant at the .01 level in any of these four surveys.

For most of the health-related variables reviewed in this report, it is reasonable to expect that characteristics of the respondent will be strongly associated with problematic responses. However, measurements of weight and height are not responses from the woman but are objectively made by an interviewer. It is reasonable to look for variation in data quality that is more specifically related to the interview process. Four potential sources of error are:

v008 Month of interview
v028 Interviewer identification
v029 Keyer identification
v030 Field supervisor

Because v028, v029, and v030 have so many categories, the importance of all four of these variables will be assessed simply with cross-tabulation of the variable with the binary indicator of weight and height problems, and the usual chi-square test of independence. This approach does not adjust for sample weights or clustering.

As stated above, only half of the households in the Kazakhstan 1999 survey were subsampled for height and weight measurements, so the missing cases are not errors. However, there was highly significant variation by some other survey characteristics. Month of interview is significantly related to the subsampling rate ($p = .001$). During the first month of interviews, July, the percentage skipped was 38.6 percent, in August it was 53.3 percent, and in September, the third and final month, it was 54.3 percent.

In the survey with the next highest level of missing on weight and age—and not by design—Nigeria 1999, month of interview is extremely significant ($p < .001$). In March the error rate was 41.8 percent, in April it was 29.4 percent, in May it was 15.8 percent. The measurements were poor in all months but the improvement over time was dramatic. Interviewer identification was marginally significant with a .01

criterion (p = .012), and keyer identification was very important (p < .001) but supervisor was not (p = .255). As Table 4.1 shows, most of the problems in this survey were flagged values, rather than missing values, which is consistent with the importance of the keyer. Hardly any keyers had an acceptable error rate.

In the Bolivia 1994 survey, survey-related variables are highly significant but in a different combination: interviewer identification (p < .001) and field supervisor (p < .001). Month of interviewer and keyer are not important (p = .234 and p = .156). Although field supervisor is a statistically significant source of variation, the data file indicates that only three people filled this role; one of them supervised 77 percent of the interviews and had the lowest error rate. One of the other two supervisors had nearly twice the error rate of the main supervisor.

Finally, in the Guatemala 1998/99 survey, the only other survey with an error rate above 10 percent for maternal weight and height, all four of these variables are highly significant. Months of interviews (p = .001) extended from November 1998 through April 1999. The error rates for these six successive months were 13.5 percent, 9.0 percent, 12.3 percent, 14.3 percent, 16.4 percent, and 9.5 percent, respectively; there is no trend. Interviewer identification (p < .001) shows that the survey employed 46 different interviewers, 11 of whom did fewer than 20 interviews. These 11 interviewers conducted 95 interviews, an average of only 8.6 each. A pooling of these interviewers gives an error rate of 17.9 percent. A pooling of the other 35 interviewers, who did an average of 83.9 interviews each, gives a much lower error rate, 12.7 percent. A possible inference is that more experienced interviewers have a lower error rate, although this is not the only possibility. Keyer identification (p = .004) and field supervisor (p < .001) show great variation in the incidence of problems.

It must be noted that the Guatemala 1998/99 survey was greatly affected by a tropical storm in November 1998, the very month when fieldwork commenced.[15] These conditions are described in the survey report (page 4). Fieldwork was somewhat delayed and the sequence of regions was changed from the original plans.

This review of the association between selected characteristics of the data collection and data entry process, and apparent errors in measurements of maternal weight and height, has at least two important implications. The first is that better training of interviewers, keyers, and even field supervisors may produce better measurements. The second implication, most clearly shown with the Nigeria 1999 survey, is that the tracking of apparent errors early in the fieldwork can lead to improved performance in later months.

4.1.2 Criteria for Flagging Height

Height (v438) is not flagged directly, but if the reported value of the variable is implausible, then the constructed variables are flagged. As mentioned above, three constructed variables relate the woman's height to a normative distribution for height. v439 gives her percentile in the normative distribution, v440 gives the number of standard deviations above or below the normative median height, and v440 expresses her height as a percentage of the normative median. Although their labels, which include "Ht/A," suggest that the woman's age is taken into account, the normative distribution appears to be identical for all women age 18 and above. Labels including "Ht/A" may be appropriate for children, but are misleading for the adult measures.

We will now look in greater depth at the implied criteria for flagging the constructed variables because of implausible or inconsistent measurements of height, v438. During the construction of the three variables,

[15] The author is grateful to Shea Rutstein for providing details about the circumstances of this survey.

933 cases are flagged as too extreme. What are the criteria for flagging a case? This is an important question because if a low height is flagged incorrectly, the degree of stunting will be underestimated, and if it is not flagged, but should have been, the degree of stunting will be overestimated.

A detailed review of the data indicates that if the reported value of v438 is less than 1280, i.e. 1.280 meters (or 50.39 inches), *for a woman age 18 or above*, then v439, v440, and v441 were intended to be flagged, and assigned code 9998 or 99998. The file contains 876 such cases, and 867 of them were indeed flagged. The nine cases that were not flagged were effectively dropped by being assigned code 9999 or 99999. These nine cases had v438 ≤ 1000, one meter or less, but we see no reason for assigning them codes 9999 or 99999 rather than 9998 or 99998.

Women age 15, 16, or 17 years appear to have had a threshold for flagging that was slightly lower than 1.280 meters. It appears to be 1.240 meters, but because of the very small number of mothers with ages 15, 16, or 17 who are near the threshold, it is impossible to infer the exact threshold.

The 876 cases flagged on v439, v440, and v441 for implausibly low height comprised only 0.30 percent of the 288,691 women with non-missing values of v438. For only three surveys did this percentage reach the level of 1.0 percent or more: Nigeria 1999, 305 cases or 10.78 percent; Guinea 1999, 46 cases or 1.19 percent; and Bangladesh 1996/97, 46 cases or 1.03 percent. These three surveys had the highest numbers, as well as the highest percentages, of such cases, and account for 45 percent of the 876 cases. By contrast, for 15 surveys, only one or two cases at the low end of height were flagged, and for most surveys the number was less than 10 cases.

Height was also flagged if implausibly large. It appears that the upper end of the distribution was defined as v438 > 1994, i.e., 1.994 meters (or 78.50 inches). DHS assigned v440 = 9998 for the 65 cases with v438 > 1994. Forty-seven of these cases were in the Nigeria 1999 survey, and four were in the Rwanda 2000 survey; otherwise, one or two cases were scattered in each of 11 other surveys. No cases in the consolidated file with 1280 ≤ v438 ≤ 1994 were flagged on v439, v440, and v441.

The selection of upper and lower values for plausible heights is necessarily somewhat arbitrary. However, looking at the full distribution of v438, there are no natural breaks at 1.280 meters or 1.994 meters that would lead to the selection of those boundaries. These values were selected to keep the HAZ, or height for age z-score, v440, in the range of six standard deviations from the normative median height. That is, heights less than 1.280 meters would produce a v440 value of -600 or less, and heights greater than 1.994 meters would produce a v440 value of 600 or more.

The consolidated file of mothers contains 515 cases with 1280 ≤ v438 ≤ 1994 that were coded 9999 on v439, v440, and v441, apparently because their weight was missing, that is, they had v437 = 9999. In other words, these 515 women had valid values of height but were excluded from the constructed indices of height because their weight was missing. These cases were scattered over 52 surveys but were concentrated in Mali 2001 (116 cases), India 1998/99 (80 cases), Rwanda 2000 (49 cases), and Ghana 2003 (42 cases). We see no reason why these women had to be dropped, and suggest that such cases be retained for the height indices, although of course they must be dropped from the weight-for-height indices.

Careful inspection of the 932 flagged values of v438 outside the range 1280 ≤ v438 ≤ 1994, of which a total of 351 can be traced to the Nigeria 1999 survey, suggests that many of the flagged values are due to certain types of data entry errors, particularly the omission of leading or final digits. For example, 404 of them occur in the range 280 to 994, which is consistent with a leading digit "1" having been dropped. Because all legal values of v438 would have had a leading digit "1," it is plausible that this digit would have been occasionally and inadvertently omitted. Outside of the Nigeria 1999 survey, this was by far the most common type of illegal value, accounting for 350 out of 581 flagged values. Another 251 flagged values occur in the range 128 to 199, which is consistent with the final digit having been dropped. In the Nigeria 1999 survey, this was the most common type of illegal value, accounting for 181 out of 351 flagged values. This mistake could easily occur during data entry but could also occur in the field. Generally, in countries that use the metric system, heights are measured with three significant digits rather than four. The file even contains five cases (three of which are from the Nigeria 1999 survey) isolated in the range 12 to 19, consistent with dropping the final two digits. It may be useful to know that the most common type of error appears to be dropping a leading or terminal digit.

The file contains a few other mysterious low values of v438, specifically v438 = 57 (Benin 1996) and the following (all in Nigeria 1999): 71, 72, 100, 110, 112, 120 (four times), 121, 123, and 127. There are a similar number of mysterious high values of v438, specifically v438 = 2510 (Haiti 2000) and the following (all in Nigeria 1999): 2200 (four times), 2305, 2500 (twice), 2505, 2600, and 2750. They are almost certainly the result of rather simple data entry errors, but what is most remarkable is that there are so few such cases.

All of the remaining flagged values on v438 actually have the appearance of being legitimate, although improbable. Below the lower legal boundary there is an uninterrupted string of 187 cases ranging from 999 to 1277. (The 999 may be an incorrectly entered 9999, but there is also a case with v438 = 1000.) There is also an uninterrupted string of 22 cases above the upper legal boundary, ranging from 1995 through 2085. The first seven of these cases, with values 1995 through 1999, are outside Nigeria 1999; the other 15 are from Nigeria 1999.

To conclude this section, we find remarkably low levels of missing or implausible heights in almost all DHS surveys except for the generally exceptional case of the Nigeria 1999 survey. A few other surveys have been mentioned several times in this section but, in view of the large numbers of women whose heights were recorded, the quality of the recorded data is very good.

4.2 Measurements of Maternal Weight

This section will assess the quality of the data on maternal weight and the constructed measures of weight that adjust for the woman's height. Some of these measures also take into account her age and pregnancy status. The relevant variables in the standard recode files are the raw measurements of weight (v437), height (v438), and six constructed measures, as follows:

v437 Respondent's weight (kilos-1d)
v438 Respondent's height (cms-1d)
v442 Wt/Ht Percent ref. median (DHS)
v443 Wt/Ht Percent ref. median (Fogarty)
v444 Wt/Ht Percent ref. median (WHO)
v444a Wt/Ht Std deviations(resp) DHS
v445 Body mass index for respondent
v446 Rohrer's index for respondent

As stated in Section 4.1, weight is measured in kilograms, coded in tenths of a kilogram. Thus, for example, a weight of 50.0 kg is coded as 500. There is considerably less heaping on final digits for weight than was observed for height. In the weighted consolidated file, final digit 0 occurs 12.98 percent of the time, and final digit 9 occurs 8.86 percent of the time. Otherwise, all final digits occur 9.24 to 10.43 percent of the time, relatively minor deviations from the expected 10 percent. It appears that most transfers are due to an upward shift of one-tenth of a kilogram, from final digit 9 to final digit 0, with little net effect. Even the Nigeria 1999 survey, which generally sets the standard for most evidence of reporting error, is close to a uniform distribution for the final digit of v437.

If a case is missing on v437, it is given a code 9999. Section 4.1 included some analysis of the missing codes. In this section we will focus on three constructed variables that combine weight with height: v442, v444a, and v445. Their construction is described in the DHS recode documentation and in Nestel and Rutstein (2002).

4.2.1 Discussion of v442, "Wt/Ht Percent Ref. Median (DHS)"

During the preparation of the standard recode files, indices are constructed that can be used to compare the measurements of women's weight with international norms. Six measures describe the woman's weight relative to her height, using normative distributions that also refer to age. Three of these, v442, v443, and v444, give her weight as a percentage of a normative median for women of her age and height. Age is in standard five-year intervals, 15-19,…, 45-49 (coded as v013). These three measures use somewhat different standards but are correlated with one another at an average level of about 0.95. The three Wt/Ht percentages are not usually used in the main country reports, but because they are made available to other analysts, we will include a brief assessment of them, beginning with the one that is specifically attributed to DHS itself, v442.

v442 is coded as 99998 if an implausible value would otherwise be produced. The plausible range can be inferred to be 5500 ≤ v442 ≤ 24000, that is, from 55 percent through 240 percent of the normative weight, given height, age, and pregnancy status. How does this range match with the original raw values of weight and height?

We find that valid codes for v442 were assigned for heights as small as 1.371 meters (v438 = 1371) or as large as 2.510 meters (v438 = 2510). This is inconsistent at the upper end with the acceptable range of heights, 1.280 to 1.994 meters, as inferred in Section 4.1. It is not at all clear why women whose height was given to be greater than 1.994 meters were considered to have implausible values for the purposes of the height indices, v439, v440, and v441, but to be acceptable for the weight-for-height indices, v442, v443, and v444. There are only 32 such cases, with recorded heights between 1.995 and 2.510 meters; 24 cases were in the Nigeria 1999 survey, two each in Haiti 2000 and India 1998/99, and one each in Burkina Faso 2003, Madagascar 1997, Nicaragua 2001, and Rwanda 2000. We suggest that the software be modified so that cases coded 9998 on v439 through v441 will *always* be coded 99998 on v442 through v444.

Such a rule was generally followed, however, with the exception of those 32 cases. Of the cases with v442 = 99998, 870 were coded 9998 on v439 through v441 because recorded height was less than 1.280 meters; another 45 cases were similarly coded because recorded height was greater than 1.994 meters. Similarly, 916 out of the 2,376 cases flagged on v442 were also flagged on v439 through v441, because of recorded heights outside the range 1.280 meters to 1.994 meters (i.e., because of v438 < 1280 or v438 >1994).

Of the cases flagged on v442 (v442 = 99998), 2376 – 916 = 1460 had plausible values of height, i.e., had 1280 ≤ v438 ≤ 1994, and apparently were discarded because the weight was implausible given the combination of height, age, and months of pregnancy. Many of these cases could have been flagged just because the recorded weight was implausible regardless of the reported height, age, and months of pregnancy. Indeed, the recorded weights for these cases range from 3.1 kg. (v437 = 31) to 906.3 kg (v437 = 9063). It is difficult to establish a plausible range for weight by itself, but it appears that many of the flagged and extreme values are due to the kinds of data entry errors described earlier for heights such as a dropped leading digit, a dropped terminal digit, or a mistyped leading digit.

We can see, for example, whether there is some low value of weight, v437, below which v442 is always flagged, and whether there is some high value above which v442 is always flagged, regardless of the woman's height, etc. It is readily found that the lowest accepted weight is 26.5 kg (v437 = 265). There are 245 cases with a lower recorded weight, and all of them are flagged on v442. More than half of those cases, 138, are in the Nigeria 1999 survey. The only other surveys with more than 10 cases with implausibly low weights are in the Bangladesh 1996/97 survey (15 cases) and 1999/2000 survey (11 cases). Another 31 surveys have at least one case with implausibly low weight. The highest accepted weight is 178.0 kg (v437 = 1780). There are 39 cases with a higher recorded weight, and all of them are flagged on v442. Two-thirds of these cases, 26, are in the Nigeria 1999 survey. Ten other surveys have one, two, or three cases with implausibly high weight.

The only further description of these 1,460 cases, which were probably flagged for other good reasons, will be in terms of their distribution across surveys. Most surveys had at least a handful of such cases, but six surveys accounted for 58 percent of them: Bangladesh 1996/97, 60 cases; Guatemala 1995, 221 cases; Guatemala 1998/99, 98 cases; India 1998/99, 224 cases; Nigeria 1999, 78 cases; and Peru 1996, 96 cases. These 1,460 cases are trivial when compared with the 285,792 cases with valid codes on v442. It is very unlikely that misinterpretations could be made because of incorrect flagging. In any case, as mentioned before, v442 and the other two Wt/Ht percentages, v443 and v444, are not normally cited in DHS reports.

4.2.2 Discussion of v444a, "Wt/Ht Std Deviations(Resp) DHS"

We now turn to the two Wt/Ht indices that are mainly used to assess wasting. The first of these, v444a, converts the woman's weight to a z score, the number of standard deviations below or above the median normative weight for women of her age and weight, sometimes call WHZ. This variable, like v442, uses a construction developed specifically by DHS.

v444a is intended to have a normative mean of 0 and standard deviation of 100. Women with scores below -200 are "wasted" and those with scores below -300 are "severely wasted." Scores below -400 and scores above 600 are flagged (v444a = 9998).

This variable is flagged (v444a = 9998) for 2,343 implausible combinations of weight and height. These are exactly the same as the 2,376 cases that are flagged for v442, v443, and v444, except for 33 cases that were flagged for those three variables, and should have been flagged for v444a, but instead were assigned to missing (v444a = .). These cases were confined to three surveys: five cases in Ghana 1993, 17 cases in Kenya 1993, and 11 cases in Turkey 1993. It appears that there was an error in the algorithm for constructing v444a that was corrected after the processing of these three surveys, all of which were at the beginning of the time interval for this assessment.

Before describing the algorithm by which v444a is calculated we will first illustrate what it accomplishes. Consider, for example, women with a reported height of 150 centimeters, or v438 = 1500. This is equivalent to 4 feet, 11 inches, and is at the 22nd percentile of the distribution of height for the consolidated file. The file includes 2,103 women at this exact value of v438 who are not missing on v444a.

Of these 2,103 women, nine were correctly assigned to 9998 or 9999 according to the coding rules. The reported weight of five women (with v438 values 5.0 kg, 5.1 kg, 6.3 kg, or 140.0 kg, [twice]) was implausible, so v444a was assigned the value 9998. For four women, the reported weight was missing (v437 = 9999), so v444a was assigned the value 9999.

The remaining 2,094 women with v438 = 1500 have a reported weight that ranges from 32.4 kg (v437 = 324), or 71.3 pounds, to 100.8 kg (v437 = 1008), or 221.8 pounds. There is a clear preference for coding v437 with final digit 0, i.e., weight in whole kilograms; 20.1 percent of these women have final digit 0, with a virtually uniform distribution across other final digits (e.g., there is no preference for final digit 5). This level of heaping at final digit 0 is characteristic of the full distribution of weight.

The distribution of weight, v437, for women at this height is given in Figure 4.1. Each interval consists of a kilogram, with v437 rounded to the nearest kilogram. The distribution is skewed to the right; large deviations above the mean are more common than large deviations of the same magnitude below the mean. The distribution has a mean of 51.8 kg, a median of 50.4 kg, a standard deviation of 9.50 kg, and a skew of 1.09.

The distribution of these same women on v444a is given in Figure 4.2. The distribution is conspicuously more symmetric than the distribution of weight itself. The construction of v444a largely compensates for the positive skew in the distribution of weight. For example, if a woman has a weight of 35.1 kg (v437 = 351) and a height of 1.500 m (v438 = 1500), then her standardized weight (v444a) is -268, more than two standard deviations below normal for her height. Such a woman would be "malnourished."

Figure 4.1 Unweighted distribution of physical weight (v437) for all 2,094 mothers with height (v438) equal to 1.500 meters and a valid code for v444a, all DHS surveys 1993-2003

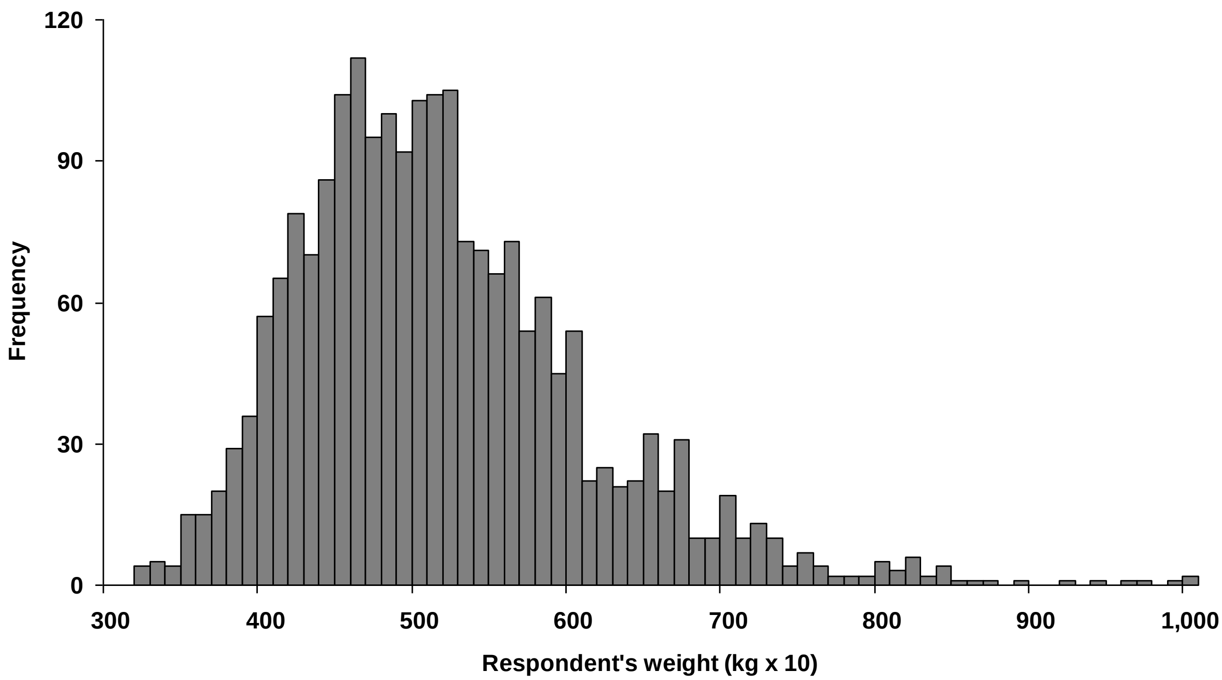
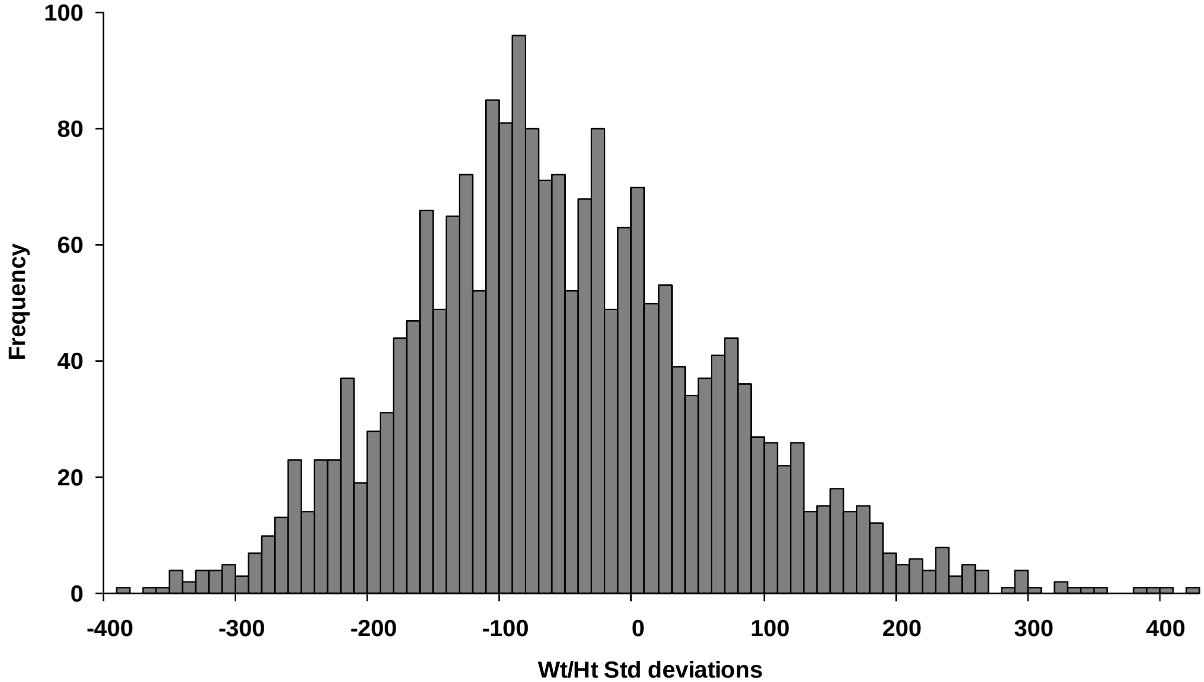


Figure 4.2 Unweighted distribution of standardized weight (v444a) for all 2,094 mothers with height (v438) equal to 1.500 meters and a valid code for v444a, all DHS surveys 1993-2003



4.2.3 Discussion of v445, Body Mass Index

v445 is the conventional BMI, calculated simply as weight (in kg) divided by the square of height (in meters). More specifically, in order to produce the desired number of digits, v445 is coded with a factor of 10^7 such that v445 = 10^7 [v437/(v438^2)], and BMI = v445/100. With standard terminology for this measure, if the BMI is less than 16.0, the woman is "severely thin;" if 16.0 to 16.9, she is "moderately thin;" if 17.0 to 18.4, she is "mildly thin." The entire range below 18.5 is "thin." If the BMI is 25.0 to 29.9, the woman is "overweight;" if 30.0 or higher, she is "obese." For example, if v437 = 400 (40.0 kg) and v438 = 1500 (1.500 meters) then v445 = 1778, the BMI is 17.78, and the woman is "mildly thin." For the "malnourished" woman described in the last paragraph of section 4.2.2, with a weight of 35.1 kg and a height of 1.500 meters, the BMI is 15.6, so she is "severely thin."

There is little concern in DHS reports with the "overweight" or "obese" categories. There is good evidence that obesity is a risk factor for complications related to pregnancy, and some other negative outcomes, but the main concern is with evidence of malnutrition. Note that the BMI does not take account of age or pregnancy status, unlike v442 and v444a.

The BMI distribution for the 2,142 mothers in the consolidated file with height 1.500 meters is given in Figure 4.3. Because they all have the same height, the distribution looks the same as that of weight itself in figure 4.1, showing the same skew to the right.

The BMI distribution for all 287,193 mothers in the consolidated file is given in Figure 4.4 (omitting, of course, the flagged cases). The distribution has the same basic shape as the one in Figure 4.3, showing how the BMI is able to successfully combine a full range of weights and heights. The full distribution is remarkably regular, but has a long tail on the right (with 111 cases between 50.00 and 60.00, and 893 cases between 40.00 and 50.00) with cases that mostly should have been flagged.

Figure 4.3 Body mass index (BMI) for all 2,142 mothers with height equal to 1.500 meters, all DHS surveys 1993-2003

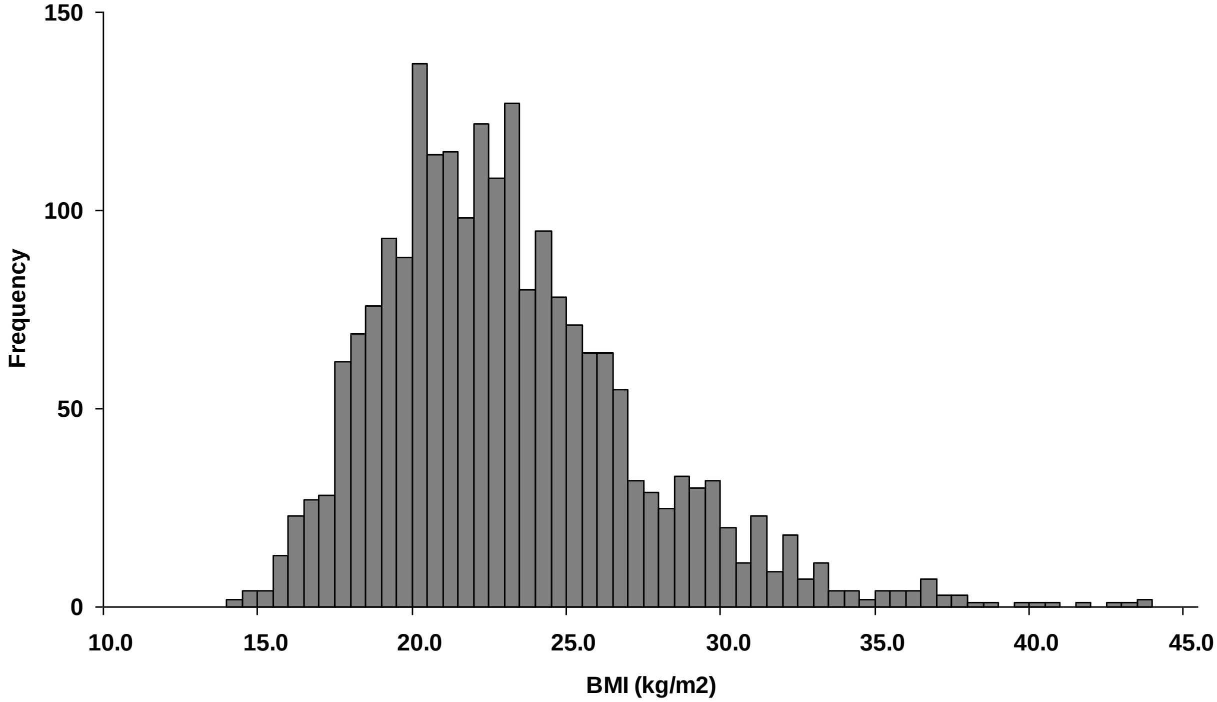
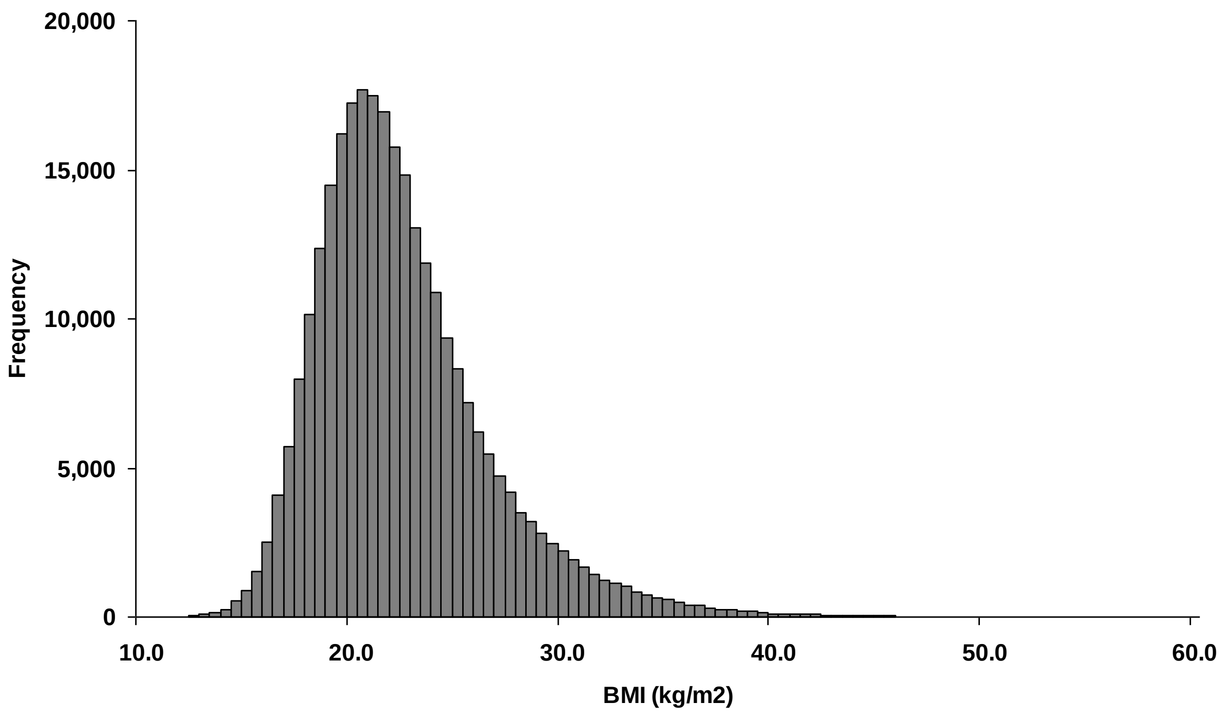


Figure 4.4 Body mass index (BMI) for all 287,193 mothers, all DHS surveys 1993-2003



v445 in the consolidated file appears to have been calculated exactly according to the formula, except for rounding error, with the exception of 1,135 cases in the Bolivia 2003/04 survey. For these women (out of a total of 7,200 women in the survey with valid codes for v445), the value of v445 in the data is one to five points below the value given by the formula. Thus, for the woman with the largest discrepancy, v437 = 647 and v438 = 1455, implying that v445 should be 3056 (BMI = 30.56). However, for this case, v445 has the value 3051 (BMI = 30.51). This difference of five points in v445 translates to a difference of only .05 in the BMI and is negligible.

The difference between calculated and coded values of the BMI in the Bolivia 2003/04 survey can be traced to a deduction for the estimated weight of heavy clothing (for women in traditional dress) before making the BMI calculation.[16]

v445 is missing (code 99999) if either v437 or v438 is coded 9999; 9,338 cases have code 99999. They are flagged (code 9998) for 975 implausible combinations of weight and height. As with v442 and v444a, cases are flagged on v445 for combinations of weight and height that would produce implausibly low or implausibly high values of the index, but the criteria are more tolerant, so to speak, for v445 than for v442 and v444a.

The criterion for flagging on v445 appears to be that the BMI must not be below 12.00 (183 flagged cases) or above 60.00 (791 flagged cases). The single flagged case that does not fit either of these restrictions is in the Ethiopia 2000 survey, with v437 = 160 (weight = 16.0 kg) and v438 = 532 (height = 0.532 meters). Both of these measurements are obviously incorrect, and implausibly low, but they would combine to produce a very large BMI, 56.53. This case is flagged on all the constructed variables, but seems to fall outside the rules described above for flagging v445. Several cases that would have similar values of the BMI are *not* flagged.

Well over a thousand cases are flagged for v442 and/or v444a but are *not* flagged for v445. We recommend more stringent flagging for v445, perhaps starting the upper range of flagged cases at 50.00, or even 40.00, rather than 60.00.

v446 is a less-frequently encountered measure, Rohrer’s Index, calculated as weight divided by the cube of height. It is coded missing or flagged for exactly the same cases as v445, presumably using criteria for v445. This measure is found in the literature but is rarely used. We will not go into the interpretation of different ranges for this index.

There are large international variations in the mean ratio of weight to normative median weight, v442. Substantial differences appear even after a restriction to women with secondary or higher levels of education. Table 4.2 is limited to such women and to surveys in which the deviation of their mean from the overall mean (for the better-educated women in all surveys) is at least 4 percent. All differences are very significant.

Better-educated women in all the South Asian surveys are 10-15 percent below the overall average ratio. Better educated women in the Ethiopia 2000 survey have a similar negative deviation. By contrast, better-educated women in the two surveys of Egypt are 14 percent and 20 percent above the overall average. Several surveys show smaller deviations.

The basic DHS strategy of assessing nutritional status from measurements of height and weight, which are then converted to indexes and compared with established thresholds, is consistent with WHO and CDC recommendations. There is certainly no basis for taking issue with the basic strategy. Of the

[16] Again, the author acknowledges information about the fieldwork provided by Shea Rutstein.

measures discussed in this section, v442 and v444a are based on a model in which ideal weight is approximately proportional to height, v445 is based on the assumption that ideal weight is approximately proportional to the square of height, and v446 on the assumption that ideal weight is approximately proportional to the cube of height. The BMI (v445) appears to be the most widely used of these different kinds of measures and is currently the one most often used in DHS country reports.

We fully support the inclusion of a variety of measures in the standard recode files, and also support the emphasis on BMI in reports. We recommend that DHS expand the flagging of cases on the BMI so that the flagged cases will better correspond with the flagged cases on v442 and v444a. We would be very cautious about deviating from the standard and recommended practices of WHO and other agencies, but it is possible that the international standards do not fit all contexts, cultural as well as biological, equally well. The use of different standards for different regions might be considered.

A handful of surveys include hemoglobin measurements. These data will not be included in this assessment.

Table 4.2 DHS surveys in which the mean of v442, the ratio of mother's weight to normative median weight, given height, age, and pregnancy status, deviates from the overall mean by 4 percent or more, all DHS surveys 1993-2003 that include height and weight measurements. Restricted to women with secondary or better education.

Survey	Mean deviation
Bangladesh 1996/97	-15.34
Bangladesh 1999/2000	-13.06
Benin 2001	4.39
Bolivia 1998	4.96
Bolivia 2003/04	6.60
Burkina Faso 2003	6.00
Central African Republic 1994/95	-6.74
Egypt 1995	14.30
Egypt 2000	20.06
Ethiopia 2000	-12.46
Ghana 1998	-4.96
Guatemala 1995	4.63
Guatemala 1998/99	10.11
India 1998/99	-13.55
Kenya 1993	-4.03
Kyrgyz Republic 1997	-4.74
Madagascar 1997	-12.57
Nepal 1996	-12.41
Nepal 2001	-10.30
Nicaragua 1997/98	4.93
Nicaragua 2001	8.24
Peru 1996	4.47
Peru 2000	6.60
Togo 1998	-5.32
Turkey 1993	5.14
Turkey 1998	5.11
Uganda 1995	-5.48
Uzbekistan 1996	-5.47
Zambia 1996	-4.81
Zambia 2001/02	-5.40

Note: The overall mean of the ratio of mother's weight to normative median weight, given height, age, and pregnancy status, and expressed as a percentage, is 99.38 percent.

4.3 Measurements of Children's Height and Weight

As described at the beginning of this chapter, DHS used two strategies for collecting height and weight data during the interval 1993-2003. When the anthropometric data were collected as part of the survey of women, the age (in months) of each surviving child in the window (whether or not available to be measured) was calculated from the birth history as hw1; height and weight measurements were hw2 and hw3, respectively; and a result code, hw13, was coded. The main survey reports for Ghana 1993 (page 223), Ghana 1998 (page 231), Kenya 1998 (page 261), India 1998/99 (pages 426-427), and Zimbabwe 1994 (page 270) provide some details of this approach. The result codes were almost always as follows: 1, measured; 2, child sick; 3, not present; 4, child refused; 5, mother refused; and 6, other.

There were some modifications when the data were collected as part of the household survey. The variable hw1, age (in months) at time of measurement, was then computed and assigned for all living children in the window who were listed in the household schedule, along with height and weight (hw2 and hw3) and a response code, hw13, which in this strategy was generally coded as follows: 1, measured; 2, not present; 3, refused; and 6, other.

During data processing, hw1, hw2, hw3, and hw13 were transferred to the data for the correct woman and child, using the line number links between the household survey and the survey of women. If a mother was included in the household roster, but her child was not (generally meaning that the child did not live in the same household as his/her mother), these variables were given "missing" codes. Below, we will see that in some surveys there are substantial numbers of children with "missing" codes for the hw variables. It is likely that the bulk of such cases are due to the absence of the child from the household, rather than due to, say, an interviewer's failure to measure a child who was present.

As stated earlier, this analysis is restricted to children who were matched with mothers, and thus omits children in the later surveys who were measured during the household interview but did not have a mother in the household. There are reasons to believe that the omitted children would differ from the children who were included in terms of nutritional status, but we would not expect them to differ in terms of data quality.

We first look at the coverage of these measurements. Some surveys omitted the children's height and weight measurements entirely. Some omitted them (as well as the other child health data) for children at ages three and four—that is, used a three-year window rather than a five-year window. Such children should have been coded "." on all the hw variables, beginning with hw1, age in months. The natural indicator of whether the measurements applied to a specific child is hw1; the measurements should apply if, and only if, hw1 was not coded ".". The file contains some violations to this rule, most importantly 311 children in the Ethiopia 2000 survey who received valid numerical codes for hw1 but were later assigned code "." on the constructed variables. However, there were no children who were assigned "." on hw1 but subsequently received valid measurements of height and weight. We shall use hw1, supplemented with the response code hw13, to identify those children who should have been measured.

The 12 surveys in the following list omitted these measurements entirely: Bangladesh 1993/94; Dominican Republic 1999; Indonesia 1994, 1997, and 2002; Philippines 1993, 1998, 2003; Senegal 1997; South Africa 1998; and Vietnam 1997 and 2002.

Of the 485,715 children in the consolidated file, 97,980 had a "not applicable" code, ".", for both hw1 and hw13; 90,234 of these children were in the 12 surveys that omitted the measurements entirely. An additional 7,746 children were eligible for the measurements, on the basis of hw13, but were assigned the "not applicable" code for hw1. Virtually all of those 7,746 children were probably children whose measurements would have been taken in the household survey but they simply were not in the same household as the mother. The absence of their measurements definitely does not suggest poor fieldwork. Table 4.3 shows their distribution across 21 surveys. In several surveys there were from two to several hundred children who did not have these measurements taken. The greatest number, 2,322, was in the India 1998/99 survey.

Table 4.3 Numbers of children who were assigned "." on hw1 and did not have their heights and weights measured, but were eligible for measurement on the basis of their age and survey, DHS surveys 1993-2003

Survey	Number of children	Percent of children in survey
Benin 2001	158	3.3
Bolivia 2003/04	153	1.6
Burkina Faso 2003	234	2.5
Cote d'Ivoire 1994	2	0.1
Dominican Republic 2002	723	6.6
Egypt 2000	59	0.5
Ghana 2003	130	3.7
Haiti 2000	362	6.0
India 1998/99	2,322	7.5
Kenya 2003	258	4.7
Malawi 2000	384	3.7
Mali 2001	427	3.8
Mozambique 2003	299	3.3
Namibia 2000	639	16.9
Nepal 2001	55	0.9
Nicaragua 2001	184	2.7
Nigeria 2003	205	4.0
Rwanda 2000	225	3.3
Uganda 2000/01	462	7.3
Zambia 2001/02	192	3.2
Zimbabwe 1999	273	8.2
Total	**7,746**	

Apart from the surveys in which the items were missing for all children, the level of missing exceeds 5 percent in six surveys: Dominican Republic 2002, Haiti 2000, India 1998/99, Namibia 2000, Uganda 2000/01, and Zimbabwe 1999. The Namibia 2000 survey had by far the highest incidence of such cases, 16.9 percent. Later in this section we will return to these cases.

Fundamentally, there are three *potential* problems with height and weight data for children: incorrect measurement of height; incorrect measurement of weight; and incorrect measurement of age, which is particularly important for determining whether a child is stunted or underweight.

This section will first describe the reports of age and date of measurement, and how they correspond with the information about the date of the main interview and the child's birth date. Next, we review the measurements of weight and height themselves, and how they are converted to indices of nutritional status. This description will lead to some basic indicators of data quality. We will look at variations in data quality across surveys in order to identify the surveys with the most problematic data and to identify the covariates of data problems.

4.3.1 Correspondences in Ages and Dates

The hw block of variables contains some information about ages and dates that should match up with information in the b block and the date of interview. Specifically, the variables

hw1	Age in months	[0-59]
hw16	Day of birth of child	[1-31, 98 99]
hw17	Date measured (day)	[1-31, 99]
hw18	Date measured (month)	[1-12]
hw19	Date measured (year)	

should match with

b1	Month of birth	[1-12]
b2	Year of birth	
b3	Date of birth (CMC)	
b8	Current age of child	[0-4]

and

v006	Month of interview	[1-12]
v007	Year of interview	
v008	Date of interview (CMC).	

Here, CMC is an abbreviation for century month code and is calculated as CMC = 12 x (y – 1900) + m, where y is a calendar year and m is the ordinal number of a month. For example, a birth in March 2003 would have century month code 12 x (2003 – 1900) + 3 = 1239.

To avoid any ambiguity as to the definitions of these variables, the allowed codes for some of them are given in brackets. The codes in the standard recode files are always consistent in the sense that

b3 = 12 x (b2 – 1900) + b1;

v008 = 12 x (v007 – 1900) + v006; and

b8 = truncated integer value of (v008 – b3) / 12.

If the ages and dates in the hw block are consistent, then the estimated CMC of birth obtained by subtracting hw1 from the date of measurement, that is,

hw_cmc_birth = 12 x (hw19 – 1900) + hw18 – hw1,

should be the same as b3. Table 4.4 shows that these two numbers usually (for 368,902 children) agree exactly. They occasionally differ by one month (for 1782 + 45 = 1827 children), which may not be incorrect, because the days of the month are not specified. For 336 children, they differ by more than one month, with the difference ranging up to 48 months.

These 336 cases are distributed across 20 surveys, but only six surveys had 10 or more cases. Table 4.5 shows that approximately half (175) are found in the survey of Egypt 1995. India 1998/99 had the next largest number, 54, although almost all of the discrepancies in that survey are at only two months (47 at 2 months, three at 3 months, four at 4 months). The Nigeria 1999 survey, the most often-cited survey in this assessment, is next, with 29 discrepancies. The five largest discrepancies, for 26, 30, 32, 36, and 48 months, are all found in the Rwanda 2000 survey.

The preceding check does not take into account the information about days of the month. A more thorough consistency check will go into the actual construction of hw1, months of age on the day when the measurements were taken. We will briefly go into more detail on the computer construction of hw1, age of the child in months at the time of measurement.

Table 4.4 Unweighted frequency distribution of the difference between the date of birth estimated from hw1, hw18, and hw19, and the date of birth given as b3 (h_dev = 12 x (hw19 – 1900) + hw18 – hw1 – b3), all DHS surveys 1993-2003

Deviation	Frequency
-12	2
-10	3
-5	1
-3	5
-2	3
-1	45
0	368,902
1	1,782
2	260
3	10
4	28
5	6
6	10
10	2
12	1
26	1
30	1
32	1
36	1
48	1
Total	**371,065**

The household questionnaire includes the age, in years, of everyone in the household, even if the household does not include any eligible respondents for the survey of women. All women age 15-49, and all children under a cutoff age that covers the time interval for the child health questions, are included. Then, within the section of the household questionnaire that pertains to height and weight, the day, month, and year of the child's birth are recorded. (This sequence allows for a possible discrepancy between the age implied by the birth date and the age recorded earlier in the questionnaire, but the incidence of such discrepancies will not be assessed here.)

Although the day, month, and year of the child's birth are recorded at the point in the household survey when the measurements are made, only the day of the month is coded into the standard recode file, as hw16. If the child's mother is alive and in the household and is an eligible respondent for the survey of women, then the child's month and year of birth can be obtained from the mother's birth history as b1 and b2, respectively. Thus, for these children, the day of birth is obtained once, but the month and year are obtained twice. The file does not include the month and year as they were given at the time of height and weight measurement, and we can *only* obtain them as b1 and b2. We do not know whether DHS checks that b1 and b2 agree with the month and year given at the time of height and weight measurement, and do not know what is done if they disagree. To repeat, the exact date of birth can only be constructed from the data file using hw16 (day), b1 (month), and b2 (year).

The day, month, and year of measurement are given on the cover page for the household interview. In the household file, they are given as hv016 (day), hv006 (month), and hv007 (year). We assume (but have not confirmed; this appears to be the only possible source for the exact date of measurement) that these match with hw17, hw18, and hw19, respectively.

Table 4.5 Unweighted distribution, across surveys, of the number of discrepancies between two estimates of the child's birth date, when the discrepancy in Table 4.4 exceeds one month

Survey	Frequency
Bangladesh 1999/2000	10
Bolivia 1994	13
Bolivia 1998	5
Brazil 1996	1
Burkina Faso 1998/99	1
Cameroon 1998	4
Chad 1996/97	2
Colombia 1995	1
Comoros 1996	2
Egypt 1995	175
Ethiopia 2000	5
Gabon 2000	12
India 1998/99	54
Kenya 1993	3
Mozambique 1997	5
Nigeria 1999	29
Peru 2000	3
Rwanda 2000	5
Zambia 1996	5
Zimbabwe 1994	1
Total	**336**

Thus, we expect that hw1 is constructed as the number of completed months from the date of birth to the date of measurement, using the following variables:

	Date of birth	Date of measurement
Day	hw16	hw17
Month	b1	hw18
Year	b2	hw19

Specifically, we expect that hw1 is calculated as

$$(hw\ 18 - b1) + 12(hw\ 19 - b2) \quad \text{if } hw\ 16 \leq hw\ 17,$$

and

$$(hw\ 18 - b1) + 12(hw\ 19 - b2) - 1 \quad \text{if } hw\ 16 > hw\ 17.$$

An ambiguity can arise with this rule because hw16 can be missing (hw16 = 99) or don't know (hw16 = 98), and hw17 can be missing (hw17 = 99). hw16 has code 98 for 46,737 children, scattered over many surveys; about 30 percent are in the surveys of South Asia. hw16 has code 99 for 11,326 children, also scattered over many surveys but with no clear concentration. hw17 has code 99 for only one child in the entire file, in the Nepal 1996 survey. (hw17 is also coded 99 for 178 children in the Nigeria 2003 survey, but these children had hw18 = . and hw19 =., so the correct code for these 178 children would have been hw17 =.) The typical demographer would assign day 15 to these cases. In no case, however, should hw1 in the data file, and hw1 estimated as above, differ from one another by more than one month.

There are 318,564 children for whom hw16 and hw17 are given codes 1–31 (rather than 98 or 99), and for whom hw1 in the data file should agree exactly with hw1 calculated as above. However, the agreement is exact for only 178,220 children. hw1 is one month greater than expected for 139,731 children. This particular discrepancy occurs so often because it appears that *day* of birth and *day* of measurement, hw16 and hw17, were usually ignored in the calculation of hw1. Other discrepancies are harder to understand. hw1 is one month less than expected for 455 children. For another 158 children the discrepancy is greater, by an amount as large as 47 months. The pattern of these deviations looks very much like those in Tables 4.4 and 4.5. The most extreme discrepancies are the following from the Rwanda 2000 survey:

Birthdate	Date of measurement	hw1	Expected hw1
March 3, 1998	November 27, 2000	2	32
January 6, 1997	July 18, 2000	10	42
July 2, 1996	August 25, 2000	13	49
May 30, 1996	November 29, 2000	6	53

It is difficult to understand how inconsistencies such as these can occur, especially since hw1 must be constructed by a computer algorithm. However, given the size of the consolidated file, their incidence and effect are trivial. We simply recommend that DHS determine how they could have occurred and try to prevent further occurrences.

We expect that most of these inconsistencies, especially the largest ones, are the result of data entry errors. The error may be in either the hw variables (hw1 and hw16 – hw19) or in the b variables (b1 and b2) but more likely the hw variables because the b variables undergo considerable checking. The date of measurement of the child's height and weight shows other symptoms of occasional errors. Normally, the household survey precedes the survey of eligible women age 15–49 by a few days or is at virtually the same date. When hw18 and hw19 are compared with v008, 99.77 percent of all interviews of women were within one month of the measurement of the children's heights and weights (unweighted, across all surveys). But for 868 children, the interview of the woman occurred 2 to 12 months after those measurements or 2 to 10 months *before* those measurements. These puzzling cases are scattered across 35 surveys, but 45 percent of them, 389, can be traced to the Ghana 2003 survey, and 20 percent of them, 175, were in the Egypt 1995 survey. We recommend tighter field checks on the accuracy of the dates of the height and weight measurements.

If the birth date was coded correctly in the birth history, then these discrepancies will have no effect on the analysis, because the birth date is always taken to be b3, the century month code constructed from b1 and b2. However, if the correct age and date at time of measurement are actually given in the hw variables, and the birth date is incorrect in the birth history, then there could be a minor effect on estimates. In particular, the weight/age and height/age calculations could lead to extreme or flagged values. There could be additional repercussions for any analyses using birth dates. We will not pursue these cases further because there are so few of them.

4.3.2 Heights of Children

As stated earlier, the measured height of the child is hw3. As for mothers, height is measured in tenths of a centimeter, that is, millimeters. Although the use of millimeters may give an exaggerated sense of precision, accuracy for children is more important than for mothers, because the difference between the normative median and the threshold for malnutrition is much smaller for children.

hw3 has a numerical code for 387,237 children; of them, 25,237 (6.52 percent) have code 9999, which is missing in the sense of "applicable but not measured." This is more than twice the incidence of code 9999 that was observed for mothers. (In Rwanda 2000, one child had hw3 = 9997; in Bangladesh 1999/2000, two children had hw3 = 9998; and in Namibia 2000, one child had hw3 = 9998. These cases might appear to be data entry errors for 9999, but they had valid codes for hw2, weight, and will not be recoded to 9999.)

Whenever hw1 is not coded ".", that is, whenever it has a numerical code, then the height and weight questions apply to the case. It should never happen that hw1 is numerical but hw3 (or hw2) is ".". As noted earlier, this rule is violated for 311 cases in the Ethiopia 2000 survey and two cases in the Malawi 2000 survey; those cases were coded hw3 = . when they should have been coded hw3 = 9999. In the Ethiopia 2000 survey there is not a significant pattern of variation in the incidence of this error across office or field editors or other characteristics of the fieldwork. However, the pattern is strongly related to the age of the child. The odds that this error will occur are about 5.5 times greater for children age 2-4 than for children age 0-1 (within those two groups there are only small increases in the odds for single years of age). We infer that in this survey, there may have been general instructions to interviewers that if a child was difficult to measure, he or she could be exempted and receive the code ".". The code 9999 was also used in this survey, and we believe that the cases with code "." should have been assigned code 9999.

In order to relate the child's height to any norms, it is necessary to know the child's sex and age at the time of measurement. The variable "age in months" described above, hw1, is not used for this purpose by DHS. Rather, we have verified that DHS uses the recorded date of birth and date of measurement directly. To repeat, hw1 is not used in the construction of any of the height (or weight) indices. The only inputs for determining the *relative* height of the child are the following variables:

hw3	Height in centimeters (1 dec.)
b4	Sex of child (1 = male, 2 = female)

hw16	Day of birth of child	[1-31, 98 99]
b1	Month of birth	[1-12]
b2	Year of birth	
hw17	Date measured (day)	[1-31, 99]
hw18	Date measured (month)	[1-12]
hw19	Date measured (year)	

DHS standard recode files include three variables to describe the relative height of the child:

hw4	Ht/A Percentile
hw5	Ht/A Standard deviations
hw6	Ht/A Percent of ref. median

These three indices are calculated using the eight input variables and the "nutchildren" (for child nutrition) routine in the Epi-Info package distributed by the CDC. The package and routine can be downloaded free of charge from CDC. We have done so and have verified, using a sample of cases in the consolidated file of children, one case at a time, that this is the procedure used by DHS. It is clear that DHS obtained the routine in a different format and was able to enter the input information, and extract the computed indices, automatically during the construction of the standard files.

Nutchildren gives users the choice of two reference populations or standards, referred to in the menu as CDC/WHO 1978 and CDC 2000. The one used by DHS, throughout the interval 1993-2003, is CDC/WHO 1978.

We have determined that if hw16 or hw17 is coded 98 or 99, then the day is imputed to be 15. As mentioned above, this is the standard substitution. It will cause a child to receive values for hw4, hw5, and hw6 that are all slightly too high (if the true birthdate was earlier than the 15th of the month) or all slightly too low (if the true birth date was later than the 15th of the month), but such errors should tend to cancel one another in the aggregate.

The DHS files include another variable that will not be discussed in any detail, hw15, "Height: lying or standing." If a child over 24 months of age is measured lying down, then height is calculated as length minus one centimeter.

hw4 is the child's percentile in the normative distribution of height, given age and sex, and is coded such that at the 50th percentile, hw4 = 5000. hw5 is a z-score, or HAZ score, and gives the child's relative position in the normative distribution as a standard normal variable, coded to have mean 0 and standard deviation 100. hw6 is the ratio of the child's height to the normative median height, given age and sex, scaled such that the ratio is 1 if hw6 = 10000.

Out of 387,421 children in the consolidated file who have numerical values for hw4, hw5, and hw6, the three variables are coded "not stated" (hw4 = 9999, hw5 = 9999, hw6 = 99999) for exactly the same 31,817 cases (8.21 percent of cases). They are "flagged" (hw4 = 9998, hw5 = 9998, hw6 = 99998) for exactly the same 15,062 cases (3.89 percent of cases). We would expect "not stated" codes on these constructed variables when the component variables used in their construction are missing, and "flagged" codes when the components are present but implausible, either individually or in combination. This is largely, but not completely, what is observed.

Consider the circumstances under which hw4, hw5, and hw6 are coded 9999 or 99999. We first note that there is one child, in the Haiti 2000 survey, for whom the variables should have received these codes but instead were coded "." on all three variables. This child had hw1 = 50, hw2 = 999, and hw3 = 9999, and should have received codes 9999 or 99999 on the constructed variables. This is a puzzling (and, fortunately, inconsequential) discrepancy and may have arisen during some case-specific editing. Otherwise, there are 25,236 cases that were coded 9999 or 99999 on these three variables because hw3 = 9999, and an addition 6,581 children who received those codes because hw2 = 999, even though hw3 was not coded 9999. (These circumstances account for all 31,817 cases that for which hw4, hw5, and hw6 are coded 9999 or 99999; 25,236 + 6,581 = 31,817.)

Thus, as with the mothers, a child is coded missing, in effect dropped, on the Ht/A codes if height was measured but weight was not. This practice seems to be an unnecessary omission of 25,236 cases. The practice may be based on the assumption that the height data will be of poor quality if the weight data were not collected, but we do not believe that assumption is supported by the data.

Now consider the circumstances under which hw4, hw5, and hw6 are coded 9998 or 99998. These are readily found to be the cases for which the HAZ score, hw5, would be less than -600 or greater than +600, that is, more than six standard deviations away from the median normative height. This is the same standard of implausibility that is used for the measured heights of women age 15-49.

In addition, hw4 through hw6 are flagged if the weight-for-age z-score, or WAZ (hw8, to be discussed in the next section), is more than six standard deviations from the median normative weight given age, or if the weight-for-height z-score, or WHZ (hw11, also to be discussed below), is more than six standard deviations from the median normative weight given height.

As was mentioned in the review of mothers' heights, we fully endorse DHS's practice of including the measured heights themselves in the data file, and flagging the constructed variables rather than the raw variables. However, for children it is not possible for us to determine whether the case was flagged because it was deemed implausible on the basis of hw5, hw8, or hw11. We can only apply the nutchildren routine on a case-by-case basis, and do not have the software to do an automated check. We are unable to determine the frequency of the three different kinds of implausibility, or whether there is a tendency for cases that are implausible on height to be implausible on weight as well.[17]

The total number of children who should have valid values of hw4, hw5, and hw6, but were missing because the height and/or weight measurement was not made in the field, or because their measured height was implausible given their age and sex, was 31,817 + 15,062 = 46,879, or 11.84 percent of all children who should have valid codes. This is much higher than was found for women's heights. Later in this section we will go into some of the characteristics of these lost cases.

Children with hw5 below -200 are stunted; if hw5 is less than -300, they are considered to be severely stunted. DHS country reports generally tabulate the percentages of children in a survey who are stunted or severely stunted, with these definitions, within categories of variables such as age, sex, birth order, size at birth, residence, mother's education, mother's age, and wealth quintile. The reports note that within the normative population, and in the absence of malnutrition, some children would be more than two or three standard deviations below the median simply by the definition of a normal distribution. Specifically, about 2.27 percent will have hw5 < -200, and about 0.14 percent will have hw5 < -300. Most surveys find percentages below these thresholds that are far greater than would be implied by a normal distribution.

[17] The University of Massachusetts does distribute a package called NutStat which can import data from a Microsoft Access data file and produce the same output as nutchildren in an automated way.

4.3.3 Weights of Children

Weight (hw2) is measured in kilograms and is supposed to be coded in tenths of a kilogram. For example, a recorded weight of 20 kg should be coded as 200. Code 999 was used for missing values. (hw37 and hw38, for the mother, and hw3, height of the child, use code 9999 for missing. It is unfortunate that 999, rather than 9999, is the missing code for hw3.) For the most part, cases that are missing on weight are also missing on height (hw3). Table 4.6 shows that of the 4,855 children who were missing one measurement or the other, 3,730 were missing height, but not weight. A smaller number, 825, were missing weight but not height. It appears that weight was somewhat easier to measure than height.

Table 4.6 Frequencies of combinations of codes for child's weight (hw2) and child's height (hw3), all DHS surveys 1993-2003

	Child's height (hw3)			
Child's weight (hw2)	<9999	9999	"."	Total
<999	361,360	3,730	0	365,090
999	825	21,507	0	22,332
"."	0	0	98,293	98,293
Total	**362,185**	**25,237**	**98,293**	**485,715**

Weight is the crucial component for six constructed variables, as listed below:

hw7	Wt/A Percentile
hw8	Wt/A Standard deviations
hw9	Wt/A Percent of ref. median
hw10	Wt/Ht Percentile
hw11	Wt/Ht Standard deviations
hw12	Wt/Ht Percent of ref. median

These six variables are in two groups of three each, corresponding exactly with the three height-for-age variables reviewed earlier. Within each group, the first variable is a percentile within the normative distribution, the second is z-score, giving the number of standard deviations from the normative median, and the third is a ratio to the normative median. All of them are obtained automatically from the nutchildren routine of the Epi-Info software package distributed by CDC.

The six constructed variables hw7 through hw12 have exactly the same pattern of missing and flagged cases as hw4 through hw6. All of them are coded missing (9999 or 99999) for the same 26,057 cases and are flagged (9998 or 99998) for the same 15,525 cases. Thus, hw7 through hw9, like hw4 through hw6, are missing or flagged whenever either height or weight, or their combination, is deemed implausible. Because of this strategy for flagging, more cases than necessary have been discarded. We suggest that cases with plausible values of weight could have been retained for hw7 through hw9, just as cases with plausible values of height could have been retained for hw4 through hw6.

hw8 is considered implausible and is flagged if the calculated value would be less than -600 or more than 600. This is the same criterion as for hw5. However, hw11 is flagged if the calculated value would be less than -400 or more than 600. hw7 through hw9, like hw4 through hw6, depend on the child's sex as well as age. The normative weight, as well as the normative height, is slightly higher for boys than for girls. In addition, hw10 through hw12 depend on both sex and age.

If a child is more than two standard deviations below the normative median weight, given age and sex, as indicated by hw8, s/he is described as underweight. If more than three standard deviations below the normative median, the child is severely underweight. The condition of being more than two standard deviations below the normative median weight, given age and sex *and height*, as indicated by hw11, is termed "wasting." If more than three standard deviations below the normative median, the child is "severely wasted." All three of the conditions being assessed, namely being stunted, underweight, or wasted, may be indicative of malnutrition, particularly if they occur more often, at an aggregate level, than in the normative population.

4.3.4 Distribution and Correlates of Not Applicable, Missing, and Inconsistent Values

We now look further into the factors that affect the completeness and accuracy of the raw and constructed measures of children's heights and weights. First, there is a specific question in the household schedule that applies if a child is given a missing code, that is, if hw2 = 999 or hw3 = 9999. Table 4.7 gives the distribution of this variable, hw13, for all children with the code for missing on either hw2 or hw3; there are 26,062 such children.

The categories and distribution in Table 4.7 are strange, with 406 cases coded either "measured" or "7," and nearly 8,000 cases coded either "other" or "not stated." Perhaps the most useful information is that the largest single category is "not present," encompassing about 40 percent of all missing cases. The index children are only age 0-4, and should not be attending school, but for one reason or another they were apparently not in the household at the time when the measurements were taken or they resided in a different household. Even infants were sometimes not present; 927 children age 0 have this response, but the frequencies actually peak at age 2, with 2,924 absences.

Table 4.7 Unweighted frequency distribution of hw13, "reason not measured," for children with hw2 = 999 or hw3 = 9999. index children in all DHS surveys 1993 to 2003

Reason not measured	Frequency	Percent
Measured	208	0.80
Sick	992	3.81
Not present	10,532	40.41
Refused	4,027	15.45
Mother refused	2,203	8.45
Other	2,904	11.14
7	198	0.76
Not stated	4,998	19.18
Total	**26,062**	**100**

Looking in more detail at three of the reasons for not being measured, "not present," "refused," and "mother refused," we find that that all three are disproportionately more likely in the two surveys of Peru, 1996 and 2000. These two surveys account for 10.8 percent of these three responses. In several surveys there are scores or even hundreds of cases of "refused" but very few with "mother refused." For example, in the Uganda 2000/01 survey, there are 209 instances of "refused" and not a single case of "mother refused." "Refused," which presumably means that the child was uncooperative, peaks at age 2 and is a minimum at ages 0 and 4, with ages 1 and 3 intermediate. By contrast, "mother refused" peaks at age 0, is about half as likely at ages 1 and 2, and about a quarter as likely at ages 3 and 4. "Sick" is a relatively infrequent reason but is also most common at age 0 and decreases dramatically by age 4.

There are three possible points at which cases may be dropped for the purposes of providing the indices of stunting, underweight, and wasting (putting aside the 12 surveys that categorically omitted the height and weight measurements). First, a child may be assigned the "not applicable" code, ".", for the entire block of measurements. Second, a child may be assigned the "not stated" code, "999" or "9999," for height and/or weight. We distinguish between these two types of missing data, even though we suspect that the "." code was sometimes used incorrectly in place of "999" or "9999." Third, a case may be dropped because the recorded height and weight measurements were implausible, either separately or in combination.

Apart from the 12 surveys that omitted the height and weight measurements entirely, all nine recoded variables have the same frequencies of "not applicable," "not stated," or "flagged" children: 8,060, 31,817, and 15,062, respectively. Table 4.8 gives the percentages of cases with these codes, within each survey, weighted according to each survey's sampling weights. The first column is the percentage of cases coded "not applicable." (The percentages in this column differ slightly from those in Table 4.3 because Table 4.8 uses the sampling weights, and Table 4.3 did not.) The second column is the percentage coded "not stated," as a percentage of the children remaining after the "not applicable" cases have been dropped. The third column is the percentage flagged, as a percentage of the children remaining after the "not applicable" and "not stated" cases have been dropped. That is, the percentages in the second and third columns are conditional, limited to the children at risk of being classified as missing or flagged, respectively. The final column of Table 4.8 gives the percentage of children who are dropped for any of the three reasons.

As Table 4.8 shows, the combined loss from these three factors was often very large. As noted earlier, the high figure for Kazakhstan 1999 was by design and requires no explanation. Nigeria 1999 lost more than half of all cases. Burkina Faso 1998/99, Guinea 1999, India 1998/99, Mozambique 1997, Namibia 2000, Uzbekistan 1996, and Zimbabwe 1999 are other surveys that lost more than 20 percent of all cases. The only survey with less than 5 percent combined loss was Nepal 2001. For countries that had several surveys, there is not a consistent pattern of improvement. Some, such as Burkina Faso and Nigeria, had much lower losses in later surveys; others, such as Zimbabwe, more than doubled their losses in later surveys. Overall, and for most surveys, missing responses were much more serious than flagging. The surveys of Nigeria 1999 and Uzbekistan 1996 had the greatest losses from flagging.

Two sets of possible influences on these kinds of losses can be distinguished: aspects of the data collection and data entry, on the one hand, and characteristics of the mother and child, on the other.

Table 4.8 Weighted distribution of “not applicable,” “not stated,” and “flagged” codes for children’s height and weight across the DHS surveys, 1993-2003, that included these measurements

Survey	Not applicable	Not stated	Flagged	All
Armenia 2000	0.0	6.9	1.5	8.3[b]
Bangladesh 1996/97	0.0	9.4	6.5	15.2
Bangladesh 1999/2000	0.0	12.1	4.0	15.7[b]
Benin 1996	0.0	11.0	4.6	15.1[b]
Benin 2001	3.4	13.6	2.6	18.7[b]
Bolivia 1994	0.0	11.9	2.7	14.2[b]
Bolivia 1998	0.0	6.4	3.3[a]	9.5[b]
Bolivia 2003/04	1.3	3.5	2.1	6.7[b]
Brazil 1996	0.0	15.1	2.1	16.8[b]
Burkina Faso 1998/99	0.0	25.4	3.9	28.3[b]
Burkina Faso 2003	2.8	3.2	7.3	12.7[b]
Cameroon 1998	0.0	11.3	4.1	14.9[b]
Central African Republic 1994/95	0.0	6.3	4.3	10.4[b]
Chad 1996/97	0.0	9.4	3.0	12.1[b]
Colombia 1995	0.0	9.1	0.9	9.9[b]
Colombia 2000	0.0	8.1	1.0	9.0[b]
Comoros 1996	0.0	8.6	4.6	12.8[b]
Cote d’Ivoire 1994	0.1	5.7	2.8	8.4[b]
Cote d’Ivoire 1998/99	0.0	11.4	3.6	14.6
Dominican Republic 1996	0.0	15.1	1.6	16.5[b]
Dominican Republic 2002	7.1	10.5	1.4	18.1[b]
Egypt 1995	0.0	3.9	4.9	8.6[b]
Egypt 2000	0.6	1.6	4.0[a]	6.1[b]
Ethiopia 2000	3.0	2.6	3.4	8.7[b]
Gabon 2000	0.0	14.8	2.4	16.8
Ghana 1993	0.0	8.2	3.7	11.6[b]
Ghana 1998	0.0	10.1	3.0	12.9[b]
Ghana 2003	4.1	5.5	3.3	12.4[b]
Guatemala 1995	0.0	7.4	2.8	10.0[b]
Guatemala 1998/99	0.0	14.1	3.0	16.7[b]
Guinea 1999	0.0	37.9	6.2	41.7[b]
Haiti 1994/95	0.0	10.4	4.5	14.4[b]
Haiti 2000	6.5	1.7	1.6	9.5
India 1998/99	8.8	5.7	7.5[a]	20.5[b]
Kazakhstan 1995	0.0	6.6	1.5	8.0
Kazakhstan 1999	0.0	53.8	2.2	54.8[b]
Kenya 1993	0.0	12.7	3.4	15.7[b]
Kenya 1998	0.0	5.9	6.4	12.0[b]
Kenya 2003	4.3	5.2	3.2	12.1[b]
Kyrgyz Republic 1997	0.0	6.6	1.5	8.1
Madagascar 1997	0.0	10.8	2.3	12.9[b]
Malawi 2000	3.2	2.4	6.5[a]	11.7[b]
Mali 1995/96	0.0	4.8	6.2	10.7[b]
Mali 2001	3.7	7.6	5.1	15.6[b]
Mozambique 1997	0.0	21.9	4.4[a]	25.4[b]
Mozambique 2003	3.2	5.9	3.9[a]	12.4[b]
Namibia 2000	17.7	3.5	4.0	23.7[b]
Nepal 1996	0.0	7.2	2.0	9.0[b]
Nepal 2001	1.0	1.7	1.5	4.1[b]
Nicaragua 1997/98	0.0	11.1	4.3[a]	15.0[b]
Nicaragua 2001	3.3	6.4	3.2[a]	12.4[b]
Nigeria 1998	0.0	5.7	3.1	8.7
Nigeria 1999	0.0	15.1	45.9[a]	54.1[b]
Nigeria 2003	3.9	6.8	8.5	18.0[b]

Continued…

Table 4.8—*Continued*				
Survey	Not applicable	Not stated	Flagged	All
Peru 1996	0.0	7.7	2.1[a]	9.7
Peru 2000	0.0	9.5	1.5	10.9
Rwanda 2000	3.0	4.4	4.5	11.4[b]
Tanzania 1996	0.0	9.4	4.7	13.6[b]
Tanzania 1999	0.0	8.6	2.5	10.9
Togo 1998	0.0	8.1	4.0	11.7[b]
Turkey 1993	0.0	9.4	1.4	10.8[b]
Turkey 1998	0.0	17.7	1.5	18.9[b]
Uganda 1995	0.0	8.6	4.1	12.3
Uganda 2000/01	6.5	10.0	2.2	17.7[b]
Uzbekistan 1996	0.0	12.8	14.4	25.4[b]
Zambia 1996	0.0	8.3	2.9[a]	10.9[b]
Zambia 2001/02	3.3	3.1	3.9	9.9[b]
Zimbabwe 1994	0.0	6.5	3.0	9.3
Zimbabwe 1999	8.1	8.3	8.5	22.9[b]
Total	**1.4**	**9.8**	**4.3**	**14.9**

Note: The second and third columns are based on reduced denominators; see text for explanation.
[a] Highly significant variation across keyers (v029)
[b] Highly significant variation across interviewers (v028)

Characteristics of the interview, in particular the identity of the interviewer, are likely to have an effect on whether a case will be dropped. All interviewers are supposed to be adequately trained and supervised, but some individual differences may remain. (Note, however, that some interviewer differences, if found, may result simply from the fact that some interviewers were assigned to more challenging households or worked under more difficult circumstances, etc.) Checks for such effects are best done within a single survey, rather than by pooling surveys. The easiest way is to calculate the chi-square statistic to test the null hypothesis of independence in a cross-tabulation with rows for interviewers and columns for the outcomes of the measurements. We use just two outcomes: one column for valid measurements of both height and weight and a second column for unacceptable outcomes, consisting of not applicable, missing, or flagged readings. The cells of such a table contain the numbers of children in each combination of interviewer and outcome. If the chi-square statistic is large, then there is large variation across interviewers in their relative frequency of unacceptable outcomes.

Errors made in the field may actually be attributable to the team rather than to a specific interviewer. There is one set of measuring equipment for each team, and several people take part in the measurements.

Some errors, particularly implausible numerical values of height and weight, can be introduced during the data entry stage. To check for data entry errors, another cross-classification is prepared with rows for keyers and two columns; the first column is for cases that were not NA and not missing and not flagged, and the second column is for flagged cases. If the chi-square statistic for this table is large, then there is large variation across keyers in their relative frequency of flagged outcomes.

These two tables were prepared for each survey that included the height and weight measurements. The calculations use the unweighted frequencies and do not adjust for sampling weights and clustering. Such adjustments can be done with logit regressions but were often defective because the tables have many cells with small frequencies. To compensate for not making these adjustments, significance was set at the .01 level rather than the .05 level.

Table 4.8 includes a superscript “a” in the “flagged” column if the keyer identification, v029, was a significant within-survey covariate of the level of flagging. Similarly, a superscript “b” in the “all” column indicates that the interviewer identification, v028, was a significant covariate of the combined level of dropping. (Separate tests were not applied to the “not applicable” and “not stated” columns.) Nearly all of these associations, but especially those for interviewer identification, were extremely significant, far better than a .01 criterion might suggest. Keyer identification is a significant source of variation in flagged values of height and weight in 12 surveys in 10 different countries.

When a case is flagged, it is impossible to tell, with the information available to us, whether an error resulted in the field as the measurements were recorded, or in the office as they were entered into a computer file. But it is clear from hw2 and hw3 that many weights and heights are recorded incorrectly, simply because some values are much below or much above a plausible range—without taking into account whether they are inconsistent with one another. For example, the combined file contains 121 children with a weight code less than or equal to 10 (including six cases with hw2 = 0!), implying a weight of 1.0 kg or less. The reported ages of these children range from 0 to 58 months. It is doubtful that any of the surviving children, including newborns, have a correct weight less than 1.0 kg. Some of the weights may have been recorded in full kilograms, rather than tenths of a kilogram. This problem of misplaced decimal points was suggested earlier for the Nigeria 1999 survey, and, indeed, 69 of these 121 problem cases are from that survey.

In terms of height (or length), 30 centimeters (approximately one foot) is well below any plausible height for a newborn, but the file includes 769 children with heights implied by hw3 to be between 0 and 30 cm at ages between 0 and 59 months. Most surveys have at least one such value, but 308 of these children are in the Nigeria 1999 survey. Sixty children are allegedly taller than two meters (hw2 > 2000); 43 of them are in the Nigeria 1999 survey. Such cases are clearly the result of dropped digits, added digits, reversed digits, etc., which probably occur more frequently for some interviewers and keyers than for others.

Variation across interviewers appears to have been much more important than variation across keyers. It was significant in all *except* 10 surveys in 10 different countries. In most of the surveys where it was significant, the chi-square test of independence between lost cases and interviewer identification (or, equivalently, the chi-square test of homogeneity of the probability of a lost case, across interviewers), was very large, leading to a p-value of 10^{-6} or less.

To illustrate the variation across interviewers, we will focus on one of the three surveys of Ghana. These surveys were conducted in 1993, 1998, and 2003, with 55, 61, and 54 different interviewers, respectively, and overall loss rates of 11.6 percent, 12.9 percent, and 12.4 percent, respectively. Relative to other surveys, these are not particularly high loss rates, but they are nevertheless substantial.

We will look specifically at the Ghana 2003 survey, not using the sampling weights. It included 3,530 index children, of whom 3,094 have acceptable measurements of height and weight. The losses were due to 105 children who were given a not applicable code, “.”, 201 who were given a missing code for height, “9999” and/or weight “999,” and 130 cases that were flagged because of implausible values of height and/or weight. The 53 interviewers varied widely in the number of children they measured. Seven interviewers measured only one child each. Another eight interviewers measured two, three, five, or nine children. The median number of children per interviewer was 74, and the maximum was 175.

It is somewhat surprising to us that there is so much variation in the workload of the interviewers. One apparent reason for the variation is that interviewers did not work steadily from the beginning to the end of the fieldwork. Twenty–five interviewers, nearly half of the total, did their first interview in the first four days of fieldwork.[18] Two started in the next week, five the week after that, four the week after that, and so on. The date of first interview for the 54 interviewers extended over a range of 72 days.

In an attempt to identify interviewers who disproportionately accounted for losses of height and weight data, Poisson regression can be applied to the number of cases lost by each interviewer for each of the three reasons, with the log of the number of children measured as an offset. The fitted value for each interviewer, for each reason, will be the expected number of cases lost, and this can be compared with the actual number lost. The difference can be thought of as the excess due to a specific interviewer. The difference, divided by the square root of the expected number, is the standardized residual and has an approximately normal distribution. The problematic interviewers are those for whom the difference and the standardized residual are both positive and large.[19] We readily identify five interviewers with a frequency and rate of lost cases that are substantially larger than would occur by chance. (For example, we set thresholds at a level of three more lost cases than would be expected together with a standardized residual greater than three.) The most error-prone interviewers are described as follows.

Interviewer 27, who measured 144 children, and entered "not applicable" for 11 of them, would have been expected to enter this for 4.20 children. There is an excess of 6.80, with a standardized residual of 3.32.

Interviewer 65, who measured 88 children and had 12 flagged, would have been expected to have 2.62 flagged. This is an excess of 9.38, with a standardized residual of 5.80.

Interviewer 67, who measured 96 children and had 15 flagged, would have been expected to have 2.86 flagged. This is an excess of 12.14, with a standardized residual of 7.19.

Interviewer 144, who measured 175 children and entered "not stated" for 37 of them, would have been expected to enter this for 9.96 of them. This is an excess of 27.04, with a standardized residual of 8.56.

Interviewer 145, who measured 161 children and entered "not stated" for 40 of them, would have been expected to enter this for 9.17 of them. This is an excess of 30.83, with a standardized residual of 10.18.

The excess for these five interviewers, added together, is 86.19 cases, or about 20 percent of the total of 436 lost cases in this survey.

Three of these interviewers (#65, #67, and #145) were in the initial group of 25 who started in the first four days of fieldwork. There is no evidence that the initial group was any less prone to losses than the interviewers who started later. Interviewers #144 and #145 measured more children than any other interviewers.

A second step was a logit regression on the date of interview, done separately for each of the error-prone interviewers. For both interviewers #144 and #145, there was a highly significant increase in the rate of entering "not stated" as the interviewing proceeded. In separate logit regressions of the log odds of this entry on the date of interview, the z statistic for the slope is 4.85 and 4.56 for interviewers #144 and #145, respectively.

Looking closer at the residuals from this logit regression, we find that the lost cases were concentrated in the final days of interviewing for these two interviewers. Table 4.9 lists the relevant results for their last

---

[18] We refer here only to the dates of interviews that included measurements of children's heights and weights.

[19] If the difference and standardized residual are large but negative, then the interviewer would serve as a model or standard.

eight days of interviews. The list begins with date 15977 (September 29, 2003). This date was typical for these and other interviewers, with none of the listed cases classified as “not stated.” But beginning one week later with date 15984, *all* listed children were recorded as “not stated” on height and weight. The number of children listed by the two interviewers on date 15985, 15 and 13, respectively, was much greater than the normal workload.

Thus, during their final days on the job, interviewers #144 and #145 coded *all* of the children that they were supposed to measure as “not stated.” One interpretation could be that they were struggling to complete a group of assigned interviews by a deadline. Another possibility is that these interviewers were simply fired at this point, which was well before the end of data collection. Interviewing for this survey continued until day 16028 (November 19, 2003).

Table 4.9 Final days of interviewing for interviewers #144 and #145 in the Ghana 2003 DHS survey

	Interviewer #144 Number of children		Interviewer #145 Number of children	
Date	Listed	Not stated	Listed	Not stated
15977	2	0	2	0
15978	-	-	-	-
15979	4	0	9	2
15981	9	4	5	1
15982	7	1	7	1
15983	3	2	-	-
15984	4	4	10	10
15985	15	15	13	13
15986	3	3	8	8
15987	-	-	1	1

Note: Dates are coded such that January 1, 1960, would be date 1. Date 15977 is equivalent to September 29, 2003. A dashed line indicates that no interviews took place on this date.

There were a few other days in the Ghana 2003 survey with clusters of 10 or more dropped cases, such as those observed above. Those other clusters were not followed by the departure of the interviewers, but the interviewers’ behavior improved afterwards.

Similar investigations of other specific surveys could be undertaken. This somewhat detailed examination of the Ghana 2003 survey is simply illustrative; that survey was selected because of rather typical levels of lost cases rather than unusually high levels. The important finding here is that nearly all surveys show very significant variation across interviewers in the proportion of children whose height and weight measurements are coded “not applicable” or “not stated” or are flagged, and statistical methods can be used to locate problems of this type. It is suggested that DHS try to reduce the variation in these outcomes by identifying the interviewers with higher levels and thereby reduce the overall level. The usual mechanisms of better training, better supervision, and simplification of the measurement procedures are advised. It is recommended that the problematic outcomes be identified immediately by field supervisors and that households be revisited if children are coded “not applicable” or “not stated” on the initial visit. For example, the 54 children missed on dates 15984-15987 by interviewers #144 and #145 in the Ghana 2003 survey should have been revisited.

Digit preference in the recording of height and weight is completely a characteristic of the interviewer. The pattern for children is very similar to that described above for mothers. The range from 3.0 to 19.9 kg ($30 \leq hw2 \leq 199$) includes more than 99 percent of the accepted values of children's weights. An unblended form of Myers' Index, simply the index of dissimilarity of the deviation from a uniform distribution across the final digits 0 through 9, is only 2.76 percent. That is, less than 3 percent of measured weights would have to be shifted to another final digit to achieve a uniform distribution of these weights. 11.64 percent of the weights were reported with final digit 0, and 9.19 percent were reported with final digit 9. These are the largest deviations from 10 percent.

Much greater heaping is associated with recorded heights. The range from 50.0 to 109.9 cm ($500 \leq hw3 \leq 1099$) includes more than 99 percent of children's heights. The unblended Myers' Index for this distribution is 15.50 percent, more than five times as large as the index for weights. The unweighted number of children with final digit 0 is 64,743, or 19.19 percent of all cases, compared with 18,951, or 5.62 percent of all cases, at final digit 9. (The expected frequencies under a uniform distribution would be 33,735.6.) There is massive shifting from 9 up to 0, but there are clearly other transfers as well, leading to conspicuous heaping at 5 and 2, as well as 0.

There is clearly a tendency by interviewers who show digit preference on one measurement to show digit preference on the other one as well. In a cross-tabulation of the final digits for height and weight, there are 9,665 children for whom both measurements end in 0. If heaping on 0 for height were independent of heaping on 0 for weight, we would expect to observe 7,491.6 such cases. The observed number, 9,665, is 29 percent larger than the expected number.

The higher level of heaping on height than on weight is probably attributable to the much larger number of possible values for height. The interval from 30 to 199 that includes almost all accepted heights covers 170 distinguishable values. The interval from 500 to 1099 for height covers 600 distinguishable values. This is probably a level of detail that it is unreasonable to expect from interviewers. As was stated for mothers, for whom the same pattern was observed, the high level of heaping in the measured heights probably has very little, if any, analytic consequence. The process of taking the measurements could probably be accelerated, with equally good results, if the measuring instrument simply had four marks between each full centimeter, rather than nine, and the interviewer was instructed to work only with final digits 0, 2, 4, 6, and 8, or if some other simplifications were made.

We next examine the pattern of lost cases according to the characteristics of the mother and child, using the standard set of covariates used in Chapters 2 and 3 of this report. The covariates are type of place of residence, age group of mother, educational level of mother, age of the child, and the number of index children in the family. (The sixth covariate considered earlier, the number of interviewer visits, is believed to be less applicable here.)

Because of the high levels of lost cases, the analysis first pools all children in all surveys. Using multinomial logit regression, a four-category variable for the outcome (coded 1 if the measured height and weight was acceptable, 2 if "not applicable," 3 if "not stated," and 4 if "flagged") was regressed on these five categorical covariates, in turn. These regressions used the sampling weights and gave equal weight to each survey. Because each regression includes hundreds of thousands of children, each of them is highly significant, but most of them have pseudo-R^2 values well below our usual criterion of .01. The pseudo-R^2 values are .0014 for type of place of residence, .0014 for mother's age, .0207 for mother's education, .0092 for child's age, and .0011 for number of index children in the family. The patterns for two of the covariates—mother's education and child's age—will be described in detail.[20]

[20] The pseudo-R^2 for child's age, .0092, is somewhat below our earlier criterion of .01, but this covariate has a very strong relationship with some specific types of loss.

Table 4.10 comes from a cross-tabulation of mother's level of education with the four-category variable described in the previous paragraph. (Note that the percentages in the columns of Table 4.8 came from binary variables with somewhat different denominators. Therefore there are some discrepancies between the "totals" row of Table 4.8 and the "totals" rows of Tables 4.10 and 4.11, etc.) Each row of the table gives the percentage breakdown of the four-category variable within categories of the mother's level of education. It shows that the incidence of flagging monotonically declines as the mother's level of education increases. It seems likely that mothers will be interested in the results of the measurements, and mothers who are better educated will help to ensure accuracy and correct recording of the measurements. The best educated women are least likely to have the measurements coded "not applicable," presumably because of their interest in the results. However, the incidence of "not stated" (and the sum of "not applicable" and "not stated") is actually highest for these women. The variation in this particular column is difficult to justify.

Table 4.10 Outcome of measurement of child's height and weight, within mother's level of education, all DHS surveys 1993-2003 that included child's height and weight, using sampling weights and giving equal weight to each survey

Woman's (mother's) education	Measurement of child's height and weight: Okay	Not applicable	Missing	Flagged	All
No education	84.12	1.19	10.09	4.60	100
Primary	86.77	1.62	8.22	3.39	100
Secondary	84.58	1.56	10.51	3.34	100
Higher	82.91	0.92	13.46	2.71	100
Total	**85.13**	**1.43**	**9.66**	**3.79**	**100**

In Table 4.11, a similar percentage breakdown of the outcome of the measurements is given within each year of age of the child. (Age is given by variable b8, rather than hw1.) All three categories of lost cases show a steady progression in incidence. "Not applicable" is about 10 times as likely for children age 4 as for children age 0, and "not stated" is about twice as likely for children age 4 as for children age 0. The higher incidence for older children probably reflects the fact that they are more mobile, possibly in pre-school,[21] and perhaps less accessible. Older children are also more likely to reside in a different household than the mother, but, as noted earlier, this effect is difficult to identify because most surveys from 1993 to 2003 did not include the variable b16.

[21] It is also possible that children of better educated mothers are more likely to be in some kind of pre-school, which would explain the unexpected pattern by mother's education in Table 4.10.

Table 4.11 Outcome of measurement of child's height and weight, within years of age of the child, all DHS surveys 1993-2003 that included child's height and weight, using sampling weights and giving equal weight to each survey

	Measurement of child's height and weight				
Age of child (years)	Okay	Not applicable	Missing	Flagged	All
0	86.86	0.34	6.40	6.40	100
1	87.07	0.69	7.85	4.39	100
2	84.01	1.40	11.91	2.69	100
3	83.67	2.78	11.72	1.83	100
4	81.84	3.44	13.01	1.72	100
Total	**85.13**	**1.43**	**9.66**	**3.79**	**100**

Flagging, by contrast, is most likely for infants and declines monotonically for older children. Apart from the kinds of extreme values noted above that probably result from data entry errors, it is difficult to tell whether flagging is due to a defective measurement of height, weight, or age (for example, an inaccurate imputation of months of age). It may sometimes be difficult to obtain accurate measurements for a child 0-23 months because s/he is always supposed to be measured lying down, but the child may be bundled up. Whatever the reason, there is a substantial reduction in the incidence of flagging with every additional year of age until age 3, with little change from age 3 to age 4. The combined effect of these three factors is that about 13 percent of children age 0 and 1, about 16 percent of children age 2 and 3, and about 18 percent of children age 4 are dropped for one reason or another.

The pattern described in Tables 4.10 and 4.11 gives equal weight to each survey. Variations from one survey to another have been examined but will not be discussed here, except to note that in some surveys the incidence of flagging is not related to the age of the child, and occasionally it is related to age but peaks at an age other than 0. For example, in the Uzbekistan 1996 survey, the incidence of flagging is 9.30 percent at age 0, 22.79 percent at age 1, and 10.28 percent at age 2, which was the maximum age included in that survey.

There is a tendency for children to be dropped (given the "not applicable" or "not stated" codes, or flagged) at the family level. That is, if a mother has two or more index children, and one of them is dropped, then there is an increased probability that one or more of the others will be dropped. Table 4.12 describes the child data after they have been collapsed so that all of the index children associated with the same mother have been grouped into a sibship. This is a limited definition of a sibship, because it only includes siblings who should have been measured. The rows of the table give the number of children in the sibship, and the columns give the number of children who were dropped from the height and weight data for any of the three reasons. For example, of the 338 sibships of size four, there were 222 in which no children were dropped, 74 in which one child was dropped, 23 in which two children were dropped, 11 in which three children were dropped, and 8 in which all four children were dropped. The table includes five sibships of size five, but they will not be discussed.

The question of whether children tended to be dropped as a set does not arise for sibships of size one. In these sibships, the chance of being dropped was 29,790 / 213,501 = .1395. In sibships of size two, the number of children at risk of being dropped was 2 x 78,371 = 156,742. The number actually dropped was 12,534 + 2 x 4,577 = 21,688. Hence, the proportion dropped was p = 21,688 / 156,742 = .1384. This is an estimate of the unconditional probability that a child in a two-child sibship will be dropped.

One of the various ways to assess the degree of family-level clustering of the lost cases is to calculate the expected frequencies in row 2 of Table 4.12 *if there were no such clustering*. In this case the assumptions of a binomial distribution would be met, and the expected frequencies would be $(1 - p)^2n$, $2p(1 - p)n$, and p^2n, respectively, where n = 78,371 and p = .1384. The expected frequencies in row 2 would be 58,183.46, 19,687.09, and 1500.46, respectively. These numbers, rounded to the nearest integer, have been inserted into Table 4.12 immediately beneath the observed frequencies.

Applying this approach to sibships of size three, the (estimated) unconditional probability of being dropped is (1,434 + 2 x 481 + 3 x 293) / (3 x 7,952) = 3,275 / 23,856 = .1373. If there were no clustering, the expected frequencies would be $(1 - p)^3n$, $3p(1 - p)^2n$, $3p^2(1 - p)n$, and p^3n, where n = 7,952 and p = .1373.

For sibships of size four, the (estimated) unconditional probability of being dropped is (74 + 2 x 23 + 3 x 11 + 4 x 8) / 4 x 388 = 185 / 1,552 = .1192. If there were no clustering, the expected frequencies would be $(1 - p)^4n$, $4p(1 - p)^3n$, $6p^2(1 - p)^2n$, $4p^3(1 - p)n$, and p^4n, where n = 338 and p = .1192.

First note that the unconditional probabilities of being dropped are .1395, .1384, .1373, and .1192 in sibships of size one, two, three, and four respectively. These are essentially identical (the fourth one comes from only 338 sibships and is not significantly different from the others). It would be plausible to expect that children in larger sibships would be more likely to be dropped—in particular, to be given "not applicable" or "not stated" codes, but this is not observed.

Table 4.12 Number of children dropped (given the "not applicable" or "not stated" codes, or flagged) from the height and weight measurements, within sibships of sizes 1 through 5, all DHS surveys 1993–2003 that included the measurements, unweighted

	Number of children dropped					
Sibship size	0	1	2	3	4	Total
1	183,711	29,790	-	-	-	213,501
2	61,260 (58,184)	12,534 (19,687)	4,577 (1,500)	-	-	78,371
3	5,744 (5,106)	1,434 (2,438)	481 (388)	293 (21)	-	7,952
4	222 (203)	74 (110)	23 (22)	11 (2)	8 0	338
5	5	1	0	0	0	6
Total	**250,942**	**43,833**	**5,081**	**304**	**8**	**300,168**

Note: Expected frequencies, under the assumption of no clustering, are given in parentheses for sibship sizes 2, 3, and 4. A dashed line indicates cell that is logically empty.

Table 4.12 provides strong evidence of clustering, because the observed frequencies, and those expected under the assumptions of the binomial distribution, are very significantly different. In rows 2 and 3, which are most diagnostic for this tendency, there are far more sibships with more than one child dropped, and far fewer with just one child dropped, than would have been expected. In sibships of size 2, for example, we observe 4,577 households in which *both* children were dropped, whereas we would have expected only 1,500.

More analysis would be required to identify possible sources of this clustering. If, say, a mother refuses to have one child measured, it seems likely that she will refuse to have any of her children measured. If a child is absent from the household, it is more likely that siblings will also be absent. If an interviewer is working under stress, there could be a tendency to rush or make errors that would affect all siblings. Such clustering is traceable to the same factors that cause an individual case to be lost. More analysis, perhaps focusing on sibships of size two, would be useful and relatively easy to do.

In earlier chapters of this report, an attempt was made to assess the amount of bias that was produced by systematic variation in the incidence of missing values. It would be desirable to have such an assessment for the heights and weights of children (and similarly for mothers), to determine whether the estimates of malnourished children are affected by the rather high losses of cases in some surveys. Unfortunately, the strategy used earlier for producing such estimates of bias is not feasible in the present context, mainly because the conversion from recorded weight and height, hw2 and hw3, to the Ht/A, Wt/A, and Wt/Ht indices, would require special software. However, we believe that any bias is likely to be negligible, at least for most surveys, because the main covariate of the "not applicable," "not stated," and "flagged" outcomes is age, and all three of the constructed groups of indices are age-adjusted. Even if, for example, four-year olds were transferred out of the window and were under-represented in the final measures, relative to infants, the control for age in the construction of the measures should eliminate any age-related bias.

As a final comment on the three constructed measures that are used by DHS—hw5 (HAZ), hw8 (WAZ), and hw11 (WHZ)—it appears that the third of these is most successful. It combines all three of the components that are subject to error—weight, height, and age. In an examination of the distribution of the three measures for the children of the best educated women in all the surveys, this measure is least likely to produce extreme values that may be evidence of deficient data. All three of the measures are useful, of course, but the third measure is most robust.

Although the BMI is calculated and included on the records for women, it is not included among the standard anthropometric measures for children. The BMI may have been omitted because for children its interpretation depends on sex and detailed age. Virtually all the information that could be conveyed by the BMI is contained in the WHZ and the percentile index that accompanies it.

The anthropometric data include two other variables for children in subsets of surveys. Arm circumference (hw20) was included in one survey (Ghana 1998) and hemoglobin level (hw53) was included in Armenia 2000, Benin 2001, Bolivia 2003/04, Burkina Faso 2003, Egypt 2000, Ghana 2003, Haiti 2000, Mali 2001, Peru 2000, and Uganda 2000/01. These measurements are not included in this analysis.

To summarize this assessment of the measurement of children's heights and weights, we have found that the level of dropped data—which takes the form of the "not applicable," "not stated," and "flagged" values—is generally higher than for other variables in the DHS surveys. Most of the "not applicable" codes can be traced to the absence of children from their mother's household, or to subsampling, and do not reflect on the quality of the data. Highly significant variation across interviewers in most surveys implies that the incidence of dropped cases could be substantially reduced by better training and tighter supervision of interviewers. It may also be possible to simplify the task, and make it less demanding, by reducing the number of possible values for height. Accuracy to the nearest fifth or even half of a centimeter should be a sufficient goal. An expectation of accuracy to the nearest millimeter simply invites heaping and rounding, as well as frustration in attempting to measure an uncooperative infant or child.

5 Effects of Survey Design on Summary Measures of Child Health

5.1 Effect of Limiting the Number of Surviving Children in the Window

This chapter will briefly consider the effects, upon summary measures of child health, of two optional features of the typical DHS survey design. We will first examine the value of obtaining health information about all surviving children in the window, the current general practice, versus a possible alternative of only collecting this information for the two youngest children.

Using all 81 surveys, Table 5.1 gives the unweighted numbers of children by current age (in years) and their index or sequence value. The index is 1 for the youngest child, 2 for the next youngest, and so on. Of 485,715 children, only 13,045, or 2.69 percent, had an index value of three or more.

Table 5.1 Unweighted numbers of children by age and sequence, for all DHS surveys 1993-2003

	Index to birth history						
Current age of child	1	2	3	4	5	6	Total
0	113,693	1,211	6	0	0	0	114,910
1	104,327	5,899	88	2	0	0	110,316
2	78,368	28,049	713	25	0	0	107,155
3	39,716	35,840	3,121	150	4	0	78,831
4	29,445	36,122	8,308	592	35	1	74,503
Total	**365,549**	**107,121**	**12,236**	**769**	**39**	**1**	**485,715**

The only child with an index value of 6 occurred in the Dominican Republic 2002 survey. The 39 children with birth order five were scattered over 25 surveys; the 769 children with birth order four were scattered over 56 surveys. Eighteen surveys had a maximum of three children, either by design or by chance. Eight surveys (Gabon 2000; India 1998/99; Indonesia 1994, 1997, and 2002; South Africa 1998; Uzbekistan 1996; and Vietnam 2002) had a maximum of two index children, by design.

Most (57) surveys had a five-year window for the child health questions, but many had a shorter interval. The window was reduced to four years for one survey (Uganda 1995), and three years for 23 surveys (those surveys were listed in Section 1.2). Of course, the maximum value of the child index tended to be lower for surveys with a shorter window.

In view of the small percentage of children with index 3 or above, DHS might consider a general policy of only obtaining health information on the two youngest surviving children, as in the eight surveys listed above. There are at least three potential reasons for doing this. The first reason would be some savings in field and data processing expenses, although it is beyond the scope of this report to estimate the amount of savings. A second possible justification is that a maximum of two children in the window might reduce any motivation by interviewers to displace some children out of the window in order to reduce their workload. Most age transfers affect the children whose correct age is just below the age cutoff; they tend to be the higher index children for whom the health information would no longer be required. The result would be better estimates of recent fertility. Third, additional children from the same mother tend to duplicate the information of the other children.

We shall use the Guatemala 1998/99 survey, which had the highest percentage of children with index value 3 or more, to illustrate the potential effect on some key results if those children were dropped.

This exercise will involve three scenarios for selected results:

Scenario 0: Results as tabulated from all index children, following current practice;

Scenario 1: Results tabulated from just the two youngest children;

Scenario 2: Results tabulated from just the two youngest children, but using weights to adjust for the fact that higher index children are older than the two youngest children.

The weights for scenario 2 were calculated as follows. An unweighted tabulation of surviving children born in the window was prepared, giving age in months by index. In the Guatemala 1998/99 survey, the mean ages of successive children, in months, are as follows: first child, 22.6 months; second child, 41.0 months; third child, 51.2 months; fourth child, 56.5 months. (These means, and all other measures given here for the Guatemala survey, are weighted with v005.) There were 98 children at 44 months of age, for example. The number of children at index value 1, 2, 3, and 4, was 40, 51, 6, and 1, respectively. To compensate for dropping the seven children with index 3 or 4, the 51 children with index 2 were given a weight factor of 1 + (7/51) = 1.1373. The sampling weight (v005) was multiplied by this factor. A similar process was used for all other months of age. The weights for children with index value 1 were not adjusted.

Table 5.2 compares the three scenarios for five key variables that are strongly related to age and would therefore be relatively sensitive to omission of the older children: having a health card (h1), DPT 2 (h5), and recent experience of diarrhea (h11), fever (h22), and cough (h31). In addition to giving the distributions of these variables for each of the scenarios, the table also gives the index of dissimilarity measuring the difference between the baseline scenario 0 and scenario 1, and between the baseline scenario 0 and scenario 2.

For all five of these key variables, the distributions under scenarios 1 and 2 are very close to the baseline distribution. The index of dissimilarity for the difference between scenarios 0 and 1 has a maximum value of 0.75 percent and the index is consistently less, with a maximum value of 0.26 percent, for the difference between scenarios 0 and 2. These differences are well within the range of sampling error and suggest that a restriction to just the two youngest children would have very little effect on estimates.

Another possible consideration is the extent of maternal clustering when more than two children from the same household (or, indeed, more than one child from the same household, but we would certainly not recommend a limitation to one child) are included in a survey. Studies of infant mortality, for example, have found similarities in outcomes for children of the same mother. Such children share many household risk factors, genetic and environmental, that affect health and survival. Because of these similarities, each child will partially duplicate the information from other children of the same mother. As with cases from the same primary sampling unit (PSUs) or cluster, the effective sample size is less than the nominal sample size.[22]

[22] As stated elsewhere in the text, all standard errors and test statistics in this assessment are adjusted for clustering on the primary sampling unit, v001.

Table 5.2 Weighted percentage distributions of selected child health variables in the Guatemala 1998/99 survey using three scenarios for index child three and above

	Codes for selected variable				Index of dissimilarity	
Variable	0	1	2	3	0 vs 1	0 vs 2
HEALTH CARD (h1)						
Scenario 0	7.97	59.06	27.32	5.65	0.62	0.19
Scenario 1	8.04	59.61	26.90	5.46		
Scenario 2	7.81	59.12	27.45	5.62		
DPT 2 (h5)						
Scenario 0	23.90	45.64	27.58	2.89	0.55	0.23
Scenario 1	24.23	45.77	27.03	2.98		
Scenario 2	23.70	45.60	27.71	2.99		
DIARRHEA (h11)						
Scenario 0	86.55	13.45			0.62	0.26
Scenario 1	85.93	14.07				
Scenario 2	86.29	13.71				
FEVER (h22)						
Scenario 0	72.95	27.05			0.38	0.12
Scenario 1	72.56	27.44				
Scenario 2	73.07	26.93				
COUGH (h31)						
Scenario 0	66.66	33.34			0.75	0.15
Scenario 1	65.91	34.09				
Scenario 2	66.51	33.49				

Notes:
Scenario 0: Includes all index children
Scenario 1: Includes only the two youngest children
Scenario 2: Includes only the two youngest children, but adjust the weight for the second child

The ratio of the nominal sample size (that is, the number of cases for whom the data were actually collected) to the effective sample size (the size of an equivalent simple random sample) is the design effect. This effect varies from one estimator to another. It is usually calculated as the ratio of the variance (square of the standard error) for an estimator *with* a correction for clustering to the variance of the same estimator *without* a correction for clustering.

High maternal clustering would have implications for the survey design and analysis. In terms of the survey design, it would imply that there is little gain in information from taking more than two children from the same mother. In terms of the analysis, this kind of clustering tends to mean that standard tests have less significance than claimed and standard confidence intervals are too narrow. Adjusted standard errors would produce test statistics that are closer to zero and confidence intervals that are wider compared with unadjusted standard errors.

To investigate the importance of this issue, we will use three binary indicators, a subset of the five variables discussed earlier in this section. For each of these, not applicable, don't know, and missing are omitted (coded "."). These questions are asked only about children who survived to the date of interview.

They are:

- has_card01, constructed from h1, coded 1 if the child has a health card, 0 if s/he does not, and "." otherwise;
- diarrhea01, constructed from h11, coded 1 if the child had diarrhea in the past two weeks, 0 if s/he did not, and "." otherwise; and
- cough01, constructed from h31, coded 1 if the child had cough in the past two weeks, 0 if s/he did not, and "." otherwise.

It would be reasonable to expect maternal clustering on each of these indicators. Children of the same mother probably tend to receive the same kind of medical care and to be similar in whether they have a health card. Symptoms such as diarrhea and cough could be similar within a household because of a shared source of water, contagion, etc.

It is difficult to assess clustering effects when clusters are small, as with the number of children from the same mother. In particular, sibships with only one surviving index child provide no information at all about maternal clustering. There are also two situations in which it is impossible to distinguish between clustering by sibship and clustering by primary sampling units, and these arise fairly often with DHS data because many clusters contain only a few women and children. Maternal and primary sampling units clustering will be confounded if only one woman in the primary sampling units had more than two children, or if all the children in the cluster had the same outcome—for example, if no child in the cluster had diarrhea symptoms during the past two weeks.

Table 5.3 gives the estimated design effects for maternal or sibship clustering, calculated for those children for whom the issue is relevant—that is, limited to children of women who had at least two surviving children, who were in a primary sampling units with at least two such women, and for whom there was variation in the outcome within the primary sampling units.[23] To distinguish maternal clustering from psu clustering, we calculate the maternal design effects *within* psu within each survey. The strategy is to do two logit regressions for each of the binary outcomes, without covariates, within each primary sampling units. In the first regression, there is no adjustment for within-psu clustering. In the second regression, an adjustment is made for maternal clustering (using the mother's identification code, "caseid," and the option "cluster(caseid)"). The maternal design effect within the psu will be the ratio of the variance of the intercept in the second regression to the variance of the intercept in the first regression. Sampling weights are ignored.

Within each psu, the effective sample size is calculated by dividing the nominal psu size by the design effect. The effective sample sizes are then added up for the entire survey and divided into the sum of the nominal psu sizes, producing an overall design effect for the survey.

Table 5.3 presents the results for those surveys in which at least one of the three characteristics had a maternal design effect in excess of 1.5. The design effects are given in the third column under each indicator. Of the 81 surveys, 28, about one-third, are listed in the table. A design effect of 1.5 means that three children from the same mother would provide approximately the same information as two children who were randomly sampled.[24]

[23] Stata 8 can adjust for a single level of clustering but not for two, such as maternal clustering nested within psu clustering. The procedure described here was applied to all primary sampling units within a specific survey and therefore is additional to the psu-level clustering design effect.
[24] Occasionally the design effect is less than 1.0. This can happen for exactly the same reasons that an F ratio in analysis of variance can be less than 1.0.

Although a design effect of 1.5 is substantial and is found in many surveys, it is important to understand that the effect only applies to a rather small proportion of the children in these surveys. For example, in Table 5.1 the first design effect greater than 1.5 is for cough in the Armenia 2000 survey, and is 1.76. But this effect only applies to 422, or about one-fourth, of the children with a “yes” or “no” code for the item. A majority of the children in this survey were in sibships of size one and were not subject to clustering. Then, of the mothers with two or more children in the window, about 40 percent were the *only* such woman in their primary sampling unit. Finally, a few PSUs, particularly those with the fewest mothers and children, had no variation in the outcome (all or none of the children in the psu had a cough in the past two weeks). Much of the clustering at this level is thus indistinguishable from psu-level clustering and its effect on standard errors is corrected with a psu-level adjustment for clustering.

Although the issue of maternal clustering applies to relatively few children, we recommend that DHS consider the possibility of reducing the data collection to two index children. The case for such a limitation would be based primarily on considerations given earlier in this section, but it does appear that the current practice also has some statistical inefficiency because of maternal clustering. More complete checks would be necessary before making a decision, and other considerations are important, but these comparisons suggest that a restriction to the two youngest children in the window would not impair the estimates of child health.

Table 5.3 Design effects from maternal clustering, for three indicators, limited to surveys in which at least one of the design effects was greater than 1.5

	Has card01			Diarrhea01			Cough01		
Survey	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Armenia 2000	1,659	209	1.4	1,654	146	1.2	1,652	422	1.8
Benin 1996	2,329	116	1.6	2,728	209	1.1	2,727	273	1.3
Bolivia 1994	1,909	154	1.7	3,354	382	1.3	3,350	353	1.7
Bolivia 1998	3,256	914	1.6	6,709	2,403	1.2	6,706	2,984	1.3
Cameroon 1998	1,568	131	1.5	2,122	166	1.1	2,071	210	1.6
Central African Republic 1994/95	1,980	166	1.4	2,500	250	1.3	2,501	354	1.5
Cote d'Ivoire 1994	3,051	181	1.6	3,313	114	1.3	2,952	196	1.2
Cote d'Ivoire 1998/99	1,364	241	1.5	1,632	532	1.1	1,631	575	1.3
Dominican Republic 1999	267	19	1.0	576	163	0.9	574	208	1.8
Dominican Republic 2002	5,058	340	1.8	10,807	3,222	1.0	10,764	3,999	1.3
Ghana 1993	1,652	69	0.9	1,858	38	1.6	1,794	20	1.6
India 1998/99	20,749	3,998	1.6	30,896	5,960	1.1	30,904	6,158	1.3
Indonesia 1997	6,302	743	1.7	15,690	1,739	1.2	16,155	3,518	1.4
Kazakhstan 1999	1,119	7	1.5	1,248	197	1.1	1,247	177	1.2
Kenya 1998	2,065	81	1.6	3,232	198	1.3	3,229	336	1.7
Kyrgyz Republic 1997	932	53	1.5	1,051	59	1.4	1,051	49	1.8
Madagascar 1997	2,581	267	1.5	3,295	455	1.2	3,283	506	1.3
Mozambique 1997	3,063	34	1.8	3,713	97	1.0	3,709	152	1.5
Nicaragua 1997/98	6,148	1,125	1.2	7,906	2,722	1.1	7,898	3,638	1.5
Nigeria 1999	1,882	51	1.8	3,119	143	1.3	3,096	253	1.5
Nigeria 2003	3,653	866	1.5	5,068	2,017	1.1	5,054	2,255	1.2
Philippines 2003	3,075	772	1.6	6,825	1,662	1.1	6,818	2,553	1.2
South Africa 1998	3,257	124	1.3	4,516	548	1.2	4,504	637	1.5
Turkey 1993	1,677	251	1.8	3,139	627	1.2	2,988	603	1.4
Vietnam 1997	922	13	1.1	1,722	51	1.6	1,723	113	1.4
Vietnam 2002	924	6	0.7	1,300	44	1.3	1,300	70	1.8
Zimbabwe 1994	1,900	21	1.7	2,218	93	1.1	2,216	147	1.4
Zimbabwe 1999	2,393	215	1.6	3,162	822	1.1	3,138	1,022	1.2

Note: Sampling weights are ignored. Column (1) is the total number of children in the survey for which the indicator was coded 0 or 1; Column (2) is the number of children for whom the issue of maternal clustering was relevant; and Column (3) is the maternal design effect for those children.

5.2 Effect of Collecting More Health Data on Children Who Have Died

This section briefly considers whether it might be possible to expand some of the data collection to include more information about children who have died. It will be limited to 36 of the surveys conducted during 1998-2003 as part of Measure DHS+. The Mozambique 2003 survey is not included.[25]

All of the variables included in this assessment that come from blocks b, m, and v4 are asked with respect to all children, regardless of survival status, with the exception of those that describe foods in the past seven days. All variables in block h, on vaccinations, diarrhea, and fever and cough; hw, on height and weight; and ml, on malaria, are asked only of surviving children. Most of the information that is omitted for children who died is probably related to the probability of survival. We will briefly look into the question of whether there are alternatives to this aspect of the survey design.

Of the questions that are *not* asked about children who died, most are either anthropometric or refer to the past two weeks. For example, height and weight are measured during the household interview and could not be replaced with any kind of reliable recollection that could be attached to a specific age of the child. Block m does include questions about each child's size (m18) and weight (m19) at birth. The question about birth weight produces accepted values for about 47 percent of children who survived and only about 25 percent of children who died. (Birth weight is mostly known for children born in institutional settings; rural and poorer children are less likely to have been weighed but are more likely to have died.) Questions in block h that refer to symptoms of illness in the past two weeks, specifically diarrhea, fever, and cough, are also essentially current status questions, and would not produce reliable responses if they had a more remote reference period.

However, there is one type of health data that is not obtained for children who died, which is not anthropometric and does not have the past two weeks as the reference period—specifically, the vaccination data (h0-h9). Would it be useful to obtain information about childhood vaccinations for all children, including those who have died, in order to see whether there is a relationship between whether a child was vaccinated and whether or not s/he died?

To investigate this question, Table 5.4 lists the average number of children (unweighted) across the 36 surveys in DHS phase 4 who were exposed to selected months of age from 0 to 59 months. For example, each child who was older than month 0 at the time of interview was given one month of exposure to month 0; each child who had died before the interview but died after month 0 was given one month of exposure to month 0; each child who was currently alive but whose age was month 0 was given half a month of exposure to month 0; each child who died before the interview and died during month 0 was given half a month of exposure to month 0.[26] We see that, in the average survey, the proportion of total exposure to successive months of age that comes from children who died drops off very rapidly by age, primarily because of the relatively early timing of most child deaths but also because of the overall thinning out of exposure to later ages in the window. By age 12 months, the average survey has exposure from fewer than 100 children who died, and by age 24 months, fewer than 50. After 24 months, less than 1 percent of all exposure comes from children who died. The table implies that there would be some gain, although it would be modest, in collecting information about events that tend to occur during the first year of life, for potential use in predicting mortality during ages 1-4.

[25] Thirty–seven surveys were done during this time interval. The Mozambique 2003 survey was not yet available at the time when this chapter was written, but its omission has no effect on the analysis or conclusions.
[26] Most children who die in month 0 die early in that month, so half a month of exposure is an over-estimate.

Table 5.4 Cumulative exposure to selected months of age, averaged across 36 Measure DHS+ surveys, that were experienced by children who died before the survey, children who survived to the date of the survey, and all children combined, and the percentage of all exposure that was contributed by children who died before the survey

Age in months	Survival status			Percent who died
	Dead	Alive	Total	
0	472.8	6763.2	7236.0	6.53
6	223.3	5937.4	6160.7	3.62
8	123.9	5170.5	5294.3	2.34
12	73.2	4363.7	4436.8	1.65
24	39.7	3666.2	3705.8	1.07
30	18.0	2896.3	2914.3	0.62
36	10.4	2212.8	2223.3	0.47
42	3.6	1610.1	1613.7	0.22
48	1.9	1059.1	1060.9	0.18
54	0.1	482.2	482.3	0.02

We next look into the timing of the vaccinations. Table 5.5 gives the months of age at the 50th percentile (median) and at the 75th, 90th, and 99th percentile of age for those (surviving) children who did get vaccinated before the date of interview.[27] The great majority of vaccinations do occur in the first year of life. Indeed, the majority occur in the first six months.

Table 5.5 The 50th percentile (median), 75th percentile, 90th percentile, and 99th percentile of months of age when child received vaccinations

Type of vaccination	Percentile of the distribution				n
	50th	75th	90th	99th	
BCG	0	1	2	14	45,242
Polio 0	1	2	4	19	102,904
Polio 1	2	3	6	23	101,118
Polio 2	4	5	9	26	92,030
Polio 3	6	8	12	30	81,390
DPT 1	2	3	6	9	100,990
DPT 2	4	5	9	26	91,647
DPT 3	6	8	12	30	81,573
Measles	10	10	19	41	69,636

Source: 36 MEASURE DHS+ surveys, pooled and weighted to give equal weight to every survey. n is the number of children who received the specific vaccination.

[27] This distribution is somewhat affected by censoring. Children who *would* receive a vaccination after the date of interview, but before their fifth birthday, are not included. That is, the distribution over-represents children who receive their vaccinations earlier. We have also calculated this table for children 48-59 months, to reduce the effect of censoring. That table gives somewhat later ages for the 90th and 99th percentiles. The only change for the 50th percentile is that the median age for the measles vaccination increases by one month. The 75th percentile for Polio 1, 2, 3, and DPT 1, 2, 3, increases by one month. The 75th percentile for the measles vaccination increases from 10 months to 15 months.

Other data quality considerations are relevant to this possible extension of the data collection. For example, if a child died, it is less likely that the mother could show a health card for the child. This might not be a serious difficulty, because in many surveys it is almost as common for the mother to say that the card is missing or not to show it as it is for the mother to be able to show the card. It is also possible that the risk of transfers of children outside the age range of eligibility for the health questions would be exacerbated. If the risk of transfers increased for children who had died, then estimates of recent mortality would be biased downwards. The estimates of infant and child mortality are among the most important products of DHS surveys, and should not be compromised.

We recommend that DHS consider extending the vaccination questions to children who have died, but the decision should take account of possible unintended consequences such as age transfers.[28]

[28] DHS-II included some tests of the possible inclusion of vaccination information for children who had died. Negative conclusions at that time should not preclude revisiting the issue.

6 Conclusions

The main focus of this assessment of the DHS health data has been on the incidence of missing data, sometimes including "don't know" responses and responses that were flagged for implausible values. We have also looked for evidence of statistically significant deviations from randomness and possible bias in the overall means and distributions induced by non-randomness.

First and most important, the level of missing data is remarkably low for most of the health variables in most of the DHS surveys. For most surveys and indicators it is below about 2 percent, and for others it is generally below other arbitrary thresholds such as 4 percent. DHS has clearly pursued a strategy of trying to get as high a response level as possible. The question about a health card, at the beginning of the section on vaccinations, epitomizes this strategy. The possible responses are:

0: No card; never had one;
1: Yes, and the card is shown to the interviewer;
2: Yes, but the card is not shown to the interviewer; and
3: The child had a card but no longer has one.

Codes 2 and 3 could have led to a missing response in some plausible questionnaire designs, but DHS found a pattern to the cases that could not give a simple yes or no answer to the question. Similarly, the options provided for each vaccination are designed to cover all possibilities. The coding system when there are many options, for example, regarding types of liquids and foods given to a child, is simple to apply in the field and almost certainly helps to avoid missing responses.

Second, although the probability of missing often has a systematic pattern, this is almost always related, in the expected direction, to characteristics of the respondent rather than to the data collection as such. Thus, missing data are generally more prevalent in rural areas and for less educated women. We have not found any survey in which the number of interviewer visits was significantly related to the probability of missing.[29] In several instances we have tried to trace the missing cases to specific interviewers or field supervisors, rather than subnational regions. With the important exception of the anthropometric measurements, these efforts were unsuccessful.

Third, we have not found any surveys for which an adjustment for differential nonresponse would have made more than a trivial change in the reported means or percentages on the substantive outcomes. Several examples were given in this report, and the net change in the distribution, measured with the index of dissimilarity, was always less than one percentage point.

There is a high level of heaping at final digits 0 and 5 for the height and weight measurements, which can be attributed largely to the fineness of the scales that are used—a tenth of a kilogram for weight and a tenth of a centimeter for height. Under any reasonable dispersion of these cases around the heaped values, there is a negligible effect on means, standard deviations, and percentiles. The raw distributions of weight and height for women and children show some implausible values, at both the low and the high ends. It is possible that less concern for a somewhat artificial level of accuracy, in particular the alleged measurement of height in millimeters, would reduce the incidence of extreme values as well as heaping. Some interviewers, mainly in the Nigeria 1999 survey but elsewhere as well, clearly misplaced the decimal points in their entries, and either multiplied the correct values by 10 or divided them by 10. The BMI for women and the age- and sex-adjusted indices of weight given height for children are probably the most robust indicators of nutritional status.

[29] When all surveys are pooled, producing a huge total sample size, this variable is sometimes statistically significant, but the pseudo-R^2 is well under .01.

A handful of the surveys stand out repeatedly for their high levels of missing, flagged, or otherwise suspicious responses. The Nigeria 1999 survey is the most egregious example, and would serve as a textbook example of a survey with data problems. We understand that the files from this survey are not in general distribution.

It happens that DHS had virtually no control over the 1999 survey. All training and fieldwork were done *without* DHS assistance or input; DHS only became involved in the data processing and report writing. The report on this survey noted (page 6) that Macro International had no involvement in the data collection. DHS cannot easily refuse to provide assistance in the later phases of the work, when asked to do so, but perhaps it should refuse to permit its name from being used when the data are so poor. Certainly, DHS should be very wary of conducting a "limited technical assistance" survey in countries where a highly capable institution is not implementing the survey.

The Nigeria 2003 survey is a remarkable contrast with the 1999 survey. In virtually every respect, it comes out as one of the best sub-Saharan African surveys. DHS was involved in all phases of this survey, and the difference is a clear demonstration of the value of DHS technical assistance. The Nigeria 1999 survey was included in this analysis to demonstrate the importance of careful training and supervision, and to reveal specific vulnerabilities in the types of data collected by DHS.

The Kazakhstan 1999 survey, the Burkina Faso 2003 survey, and a few other surveys had high levels of several indicators. Otherwise, many surveys had only one or two problems, and most surveys did not have serious levels of any indicators

Within the report, a number of recommendations have been made. These are generally rather minor, except for those relating to measurements of height and weight. Otherwise, the most important recommendation would be that DHS consider the possibility of limiting the number of index children to the two youngest children born in the window, as was already done for several surveys. Relatively little new information comes in from a third or fourth (and sometimes a fifth or sixth) child.

Finally, it should be noted that the data processing standards followed by DHS have greatly facilitated this assessment. The standardization of variables and codes; the distinction between not applicable, missing, and flagged codes; and the distribution of the data files themselves in alternative formats have been crucially important.

We know of no comparable program of collection and dissemination of survey data on the health and nutrition of women and children. It is reassuring that the quality of this huge database has proven to be exemplary.

References

Chidambaram, V.C., and T. W. Pullum. 1981. Estimating fertility trends from retrospective birth histories: Sensitivity to imputation of missing dates. *Population Studies 35*: 307-320.

Chidambaram, V.C., and Z A. Sathar. 1984. *Age and date reporting*. WFS Comparative Studies No. 5. Voorburg, Netherlands: International Statistical Institute.

Chidambaram, V.C., J. G. Cleland, and V. Verma. 1980. *Some aspects of WFS data quality: A preliminary assessment*. WFS Comparative Studies No. 16. Voorburg, Netherlands: International Statistical Institute.

Curtis, S. L. 1995. *Assessment of the quality of data used for direct estimation of infant and child mortality in DHS-II surveys*. DHS Occasional Papers No. 3. Calverton, Maryland: Macro International Inc.

Curtis, S. L. and F. Arnold. 1994. *An evaluation of the Pakistan DHS survey bas ed on the reinterview survey*. DHS Occasional Papers No. 1. Calverton, Maryland: Macro International Inc.

Demographic and Health Surveys. 1990. *An assessment of DHS-I data quality*. DHS Methodological Reports No. 1. Columbia, Maryland: Institute for Resource Development/Macro Systems, Inc.

Ewbank, D. C. 1981. *Age misreporting and age-selective underenumeration: Sources, patterns and consequences for demographic analysis*. Committee on Population and Demography, Report No. 4. Washington, D.C.: National Academy Press.

Gage, A J. 1995. *An assessment of the quality of data on age at first union, first birth, and first sexual intercourse for Phase II of the Demographic and Health Surveys Program*. DHS Occasional Papers No. 4. Calverton, Maryland: Macro International Inc.

Goldman, N., A. J. Coale, and M Weinstein. 1979. *The quality of data in the Nepal fertility survey*. WFS Scientific Reports, No. 6. Voorburg, Netherlands: International Statistical Institute.

Goldman, N., S. O. Rutstein, and S. Singh. 1985. *Assessment of the quality of data in 41 WFS surveys: A comparative approach*. WFS Comparative Studies, No. 44. Voorburg, Netherlands: International Statistical Institute.

Hobbs., F. 2004. Age and sex composition. In J. S. Siegel and D.A. Swanson (eds.), *The methods and materials of demography*, second edition. San Diego, California: Elsevier Academic Press, pp. 125-173.

Macro International, Inc. 1993. *An assessment of the quality of health data in DHS-I surveys.* DHS Methodological Reports, No. 2. Calverton, Maryland: Macro International, Inc.

Marckwardt, A.M., and S.O. Rutstein. 1996. *Accuracy of DHS-II demographic data: Gains and losses in comparison with earlier surveys*. DHS Working Papers No. 19. Calverton, Maryland: Macro International Inc.

Mason, K.O., and L.G. Cope. 1987. Sources of age and date-of-birth misreporting in the 1900 U.S. census. *Demography* 24: 563-573.

Nag, M.H., E.G. Stockwell, and L.M. Snavely. 1973. Digit preference and avoidance in the age statistics of some recent African censuses: Some patterns and correlates. *International Statistical Review* 41: 165-174.

Nestel, P., and S. Rutstein. 2002. Defining nutritional status of women in developing countries**.** *Public Health Nutrition* 5: 17-27.

Potter, J.E. 1977. Problems in using birth history analysis to estimate trends in fertility. *Population Studies* 31: 335-364.

Pullum, T.W., N. Ozsever, and T. Harpham. 1984. *An assessment of the machine editing policies of the World Fertility Survey*. WFS Scientific Report Series, No. 54. World Fertility Survey. Voorburg, Netherlands: International Statistical Institute.

Pullum, T.W., T. Harpham, and N. Ozsever. 1986. The machine editing of large sample surveys: The experience of the World Fertility Survey. *International Statistical Review* 54: 311-326.

Pullum, T.W. 1987. Analytic methodology. In J. Cleland and C. Scott (eds.), *The World Fertility Survey: An assessment of its contribution*. New York: Oxford University Press, pp. 644-676.

Pullum, T.W. 2006. *An assessment of age and date reporting in the Demographic and Health Surveys, 1985-2003.* DHS Methodological Reports 5. Calverton, Maryland: Macro International, Inc.

Pullum, T.W. 1988. *An approach to the reconciliation of demographic survey data from the Philippines*. Occasional Paper No. 2. Population Technical Assistance Project. Arlington, Virginia: International Science and Technology Institute.

Pullum, T.W. 1990. Statistical methods to adjust for date and age misreporting to improve estimates of vital rates in Pakistan. *Statistics in Medicine* 10: 191-200.

Pullum, T.W. 2005. A statistical reformulation of demographic methods to assess the quality of age and date reporting, with application to the Demographic and Health Surveys. Paper presented at 2005 Annual Meetings of the Population Association of America, Philadelphia, March 31-April 2.

Pullum, T.W., and S.L. Stokes. 1997. Identifying and adjusting for recall error, with application to fertility surveys. In L. Lyberg and P. Biemer, et al. (eds.), *Survey Measurement and Process Quality*. New York: John Wiley and Sons, pp. 711-732.

Rutstein, S.O., and G.T. Bicego. 1990. Assessment of the quality of data used to ascertain eligibility and age in the Demographic and Health Surveys. In *An assessment of DHS-I data quality*. DHS Methodological Reports, No. 1. Columbia, Maryland: Institute for Resource Development/Macro Systems, Inc.

Siegel, J. and D. Swanson, eds. 2004. *The methods and materials of demography*, second edition. San Diego, California: Elsevier Academic Press.

Singer, E. ed. 2007. *Special issue on survey nonresponse*. Public Opinion Quarterly 71(1). New York: Oxford University Press.

Stanton, C, N. Abderrahim, and K. Hill. 1997. *DHS maternal mortality lndicators: An assessment of data quality and implications for data use*. DHS Analytical Report No. 4. Calverton, Maryland: Macro International Inc.

United Nations. 1967. *Manual II: Methods of appraisal of quality of basic data for population estimates*. Population Studies, No. 23. New York: United Nations Department of Economic and Social Affairs.

United Nations. 1987. *A comparative evaluation of data quality in thirty-eight World Fertility Surveys*. New York: United Nations Department of International Economic and Social Affairs.

Appendix A Procedure to Estimate Nonrandomness of Missing Values and Bias in Estimates

Many of the health-related variables have the same coding scheme, with the following legal codes:

0: No
1: Yes
8: Don't know (DK)
9: Missing (not stated)

All variables with this coding scheme are treated similarly and referred to generically as "y0189." These variables also may have cases with a "not applicable (NA)" code, which in Stata is coded with a dot ("."). The NA code is not problematic; it reflects the skip patterns in the questionnaire. This analysis would be much more difficult if DHS did not routinely distinguish between NA and "not stated."

"Don't know"(DK) could be treated as a valid response for an attitude variable, for example, but the health questions refer to matters of fact, such as whether a child did or did not receive a particular vaccination, food supplement, etc. With additional probing, it might have been possible to replace a DK response with "yes" or "no." In general, we treat DK and "not stated" as equivalent and pool them. For some health-related variables, there is no code for DK, but the generic description "y0189" includes such variables.

The analysis of a generic y0189 variable proceeds through two steps. In the first step, a binary variable y01_missing is constructed with the following commands (expressed in Stata syntax):

```
generate y01_missing=0
replace y01_missing =1 if y0189==8 | y0189==9
replace y01_missing =. if y0189==.
```

This indicator variable is coded 1 if the variable is missing (coded 8 or 9) and 0 if it is not missing (coded 0 or 1). y0189_missing is then used as the dependent variable in a logit regression with several covariates that are generally related to the probability of missing and do not themselves have missing cases, specifically type of place of residence (v025, age of the child's mother (or of the woman herself) in five-year intervals (v013); level of education of the child's mother or of the woman herself) in four categories (v106), the child's age in single years 0–4 (b8); number of surviving children in the window (hidxmax_surv), coded as 1, 2, and 3 or more, and number of visits by the interviewer (v027) also coded as 1, 2, and 3 or more. These variables are re-labeled x1, x2, x3, x4, x5, and x6, respectively, and are treated as categorical variables (by using "xi:" and "i." in Stata). Using logit regression, y0189_missing is regressed on each of the six predictors in turn. These regressions also use v001 (the id code for primary sampling units or clusters) and v005 (the sampling weight variable) or a modification of v005. For example, the logit regression of the "not stated" indicator on x1 would have this syntax:

```
xi: logit y0189_missing i.x1 [pweight=v005], cluster(v001)
predict phat.
```

The "predict" command used after this regression produces the fitted probability that the response for y0189 will be 8 or 9, calculated for every case that gave any kind of a response to y0189.

The pseudo-R^2 from this logit regression is the proportion of deviance in y0189_missing that is "explained" by the four predictors. This term is placed into the output file and serves as a measure of the non-randomness in the missing response.

In the second step, a binary variable y01 is constructed as follows:

```
generate y01=y0189
replace y01=. if y0189_missing==1
```

This variable is defined exactly the same as y0189 itself, except that cases missing (coded 8 or 9) on y0189 are dropped (assigned code "."). y01 is then used as the dependent variable in a logit regression with the same covariates as above, and including adjustments for the sampling weights (v005) and clusters or primary sampling units (v001) for the specific survey:

```
xi: logit y01 i.x1 i.x2 i.x3 i.x4 i.x5 i.x6 [pweight=v005], cluster(v001)
predict p1hat
gen p0hat=1-p1hat
```

Interactions of the covariates are not included because they are statistically unstable. The "predict" command used after this regression produces the fitted probability, p1hat, that the response for y01 will be 1, calculated for every case that gave a 0 or 1 response to y0189. p0hat will be the fitted probability that y = 0. The means of p0hat and p1hat, calculated with weights, for the non-missing cases will be equal to the *observed* (weighted) proportion of cases with codes 0 and 1, respectively.

In Stata, the "predict" command will extend to cases that are missing on the dependent variable. Here, the missing cases are those for which y0189_missing = 1. If the means of p0hat and p1hat are calculated for those cases, we will obtain the estimated proportion of cases that would have codes 0 and 1, respectively. These proportions will be adjusted for non-randomness in the pattern of non-response. Finally, by pooling the estimates for the missing and non-missing cases, we can evaluate the extent of bias in the observed proportions.

Exactly the same steps can be applied to other variables that have more than two non-missing outcomes, the only real difference being that the logit regression in the second step will be replaced by a multinomial logit, and to interval-level variables with the logit regression in the second step replaced by an ordinary least squares multiple regression.

Finally, we acknowledge that the six variables that are used in this model may not be sufficient to capture all of the selectivity in the missing responses. It is likely that our estimates of differences between the observed and fitted proportions are conservative. It is also possible, although less likely, that some other covariates would *reduce* these estimated differences, if some omitted covariates work in the opposite direction.

Appendix B List of Variables

This appendix lists variables in the core questionnaire, related to maternal and child health and nutrition, that appeared in phase 3 or phase 4 DHS surveys. These are variables with prefixes b, h, hw, m, ml, or v4. Some of the variables listed in this appendix, and included in the analysis, were used in only one or a few surveys, and the exact meaning may have varied from one survey to another. Thus, country-specific category labels given here may not apply to all surveys that used the variable.

The content of each group of variables reflects the evolution of the core questionnaire, particularly an expansion outwards from the "v4" block of variables, which refers only to the most recent birth, and into the "m" block, which refers to all children born in the window.

This list omits the variables bidx, hidx, etc., which identify the sequence of the child, but includes some other variables that are flags, etc.

Prefix b, for births:

Included in both phase 3 and phase 4 DHS surveys:

bord	"Birth order number"
b0	"Child is twin"
b1	"Month of birth"
b2	"Year of birth"
b3	"Date of birth (CMC)"
b4	"Sex of child"
b5	"Child is alive"
b6	"Age at death"
b7	"Age at death (months-imputed)"
b8	"Current age of child"
b9	"Who child lives with"
b10	"Completeness of information"
b11	"Preceding birth interval"
b12	"Succeeding birth interval"
b13	"Flag for age at death"
b14	"Birth interval >= 4 years"
b15	"Live birth between births"

Prefix b variables included in phase 4 but not phase 3 DHS surveys:

b16	"Child's line number in household"

Prefix b variables included in phase 3 but not phase 4 DHS surveys: none

Prefix h, vaccinations and child illness:

Included in both phase 3 and phase 4 DHS surveys:

h0	"Polio 0"
h0d	"Polio 0 day"
h0m	"POLIO 0 month"
h0y	"POLIO 0 year"
h1	"Has health card"
h2	"Received BCG"
h2d	"BCG day"
h2m	"BCG month"
h2y	"BCG year"
h3	"Received DPT 1"
h3d	"DPT 1 day"
h3m	"DPT 1 month"
h3y	"DPT 1 year"
h4	"Received POLIO 1"
h4d	"POLIO 1 day"
h4m	"POLIO 1 month"
h4y	"POLIO 1 year"
h5	"Received DPT 2"
h5d	"DPT 2 day"
h5m	"DPT 2 month"
h5y	"DPT 2 year"
h6	"Received POLIO 2"
h6d	"POLIO 2 day"
h6m	"POLIO 2 month"
h6y	"POLIO 2 year"
h7	"Received DPT 3"
h7d	"DPT 3 day"
h7m	"DPT 3 month"
h7y	"DPT 3 year"
h8	"Received POLIO 3"
h8d	"POLIO 3 day"
h8m	"POLIO 3 month"
h8y	"POLIO 3 year"
h9	"Received MEASLES"
h9d	"MEASLES day"
h9m	"MEASLES month"
h9y	"MEASLES year"
h10	"Ever had vaccination"
h11	"Had diarrhea recently"
h11b	"Blood in the stools"
h11c	"Bowel movements in worst"
h12a	"Diarrhea: government hosp."
h12b	"Diarrhea: govt health center"
h12c	"Diarrhea: govt maternity dispensary"
h12d	"Diarrhea: mobile clinic"
h12e	"Diarrhea: comm.health worker"
h12f	"Diarrhea: CSPS"

h12g	“Diarrhea: SMI”
h12h	“Diarrhea: Community pharmacy depot”
h12i	“Diarrhea: other public”
h12j	“Diarrhea: private hosp/clinic”
h12k	“Diarrhea: private pharmacy”
h12l	“Diarrhea: private doctor”
h12m	“Diarrhea: private mobile clinic”
h12n	“Diarrhea: comm.health worker”
h12o	“Diarrhea: Nurse’s office”
h12p	“Diarrhea: religious dispensary”
h12q	“Diarrhea: CS med.priv sec”
h12r	“Diarrhea: other med.priv”
h12s	“Diarrhea: shop”
h12t	“Diarrhea: traditional practitioner”
h12u	“Diarrhea: Family, friends”
h12v	“Diarrhea: family/ friends”
h12w	“Diarrhea: CS oth.priv sector”
h12x	“Diarrhea: Other”
h12y	“Diarrhea: no treatment”
h12z	“Diarrhea: medical treatment”
h13	“Given oral rehydration”
h13a	“Days given ORS”
h14	“Given recommend. home solution”
h14a	“Days given fluid”
h15	“Given other pills or syrup”
h15a	“Given antibiotics CS”
h15b	“Given an injection”
h15c	“Given an intravenous (IV)”
h15d	“Given home remedy, herbal med.”
h15e	“Given Ersefluril/ typhomicine”
h15f	“Given anti diarrheal medicine”
h15g	“Given CS other treatment”
h15h	“Given CS other treatment”
h16	“Increase or decrease fluids”
h18	“Increase or decrease in fluids”
h18a	“Breastfeeding change”
h20	“Given other treatment”
h21	“Received any treatment”
h21a	“Given no treatment”
h22	“Had fever in last two weeks”
h31	“Had cough in last two weeks”
h31b	“Short, rapid breaths”
h32a	“Fever/cough: government hosp.”
h32b	“Fever/cough: govt health cntr”
h32c	“Fever/cough: govt maternity dispensary”
h32d	“Fever/cough: mobile clinic”
h32e	“Fever/cough: comm.health”
h32f	“Fever/cough: CSPS”
h32g	“Fever/cough: SMI”
h32h	“Fever/cough: Community pharmacy depot”
h32i	“Fever/cough: oth public sector”

h32j	"Fever/cough: private hospital"
h32k	"Fever/cough: private pharmacy"
h32l	"Fever/cough: private doctor"
h32m	"Fever/cough: private mobile"
h32n	"Fever/cough: comm.health"
h32o	"Fever/cough: nurse's office"
h32p	"Fever/cough: religious dispensary"
h32q	"Fever/cough: CS med.priv"
h32r	"Fever/cough: oth med.priv"
h32s	"Fever/cough: shop"
h32t	"Fever/cough: traditional pract"
h32u	"Fever/cough: family/ friends"
h32v	"Fever/cough: CS oth.private"
h32w	"Fever/cough: CS oth.private"
h32x	"Fever/cough: Other"
h32y	"Fever/cough: no treatment"
h32z	"Fever/cough: medical treatment"

Prefix h variables included in phase 4 but not phase 3 DHS surveys:

h33	"Received Vitamin A"
h33d	"Vitamin A Day"
h33m	"Vitamin A month"
h33y	"Vitamin A year"
h34	"Vitamin A in last 6 months"
h35	"Any vaccinations in last 6 months "
h36a	"Vaccinated during Campaign"
h36b	"Vaccinated during Campaign"
h36c	"Vaccinated during Campaign"
h36d	"Vaccinated during Campaign"
h36e	"Vaccinated during Campaign"
h36f	"Vaccinated during Campaign"
h37a	"Fansidar taken for fever"
h37b	"Chloroquine taken for fever"
h37c	"Aspirin taken for fever"
h37d	"Ibuprofen/acetaminophen taken"
h37e	"CS taken"
h37f	"CS taken"
h37g	"CS taken"
h37h	"CS taken"
h37x	"Other taken for fever"
h37y	"Nothing taken for fever"
h37z	"Don't know if or what was"
h38	"Had diarrhea in last 2 weeks: Amount offered to drink"
h39	"Had diarrhea in last 2 weeks: Amount offered to eat"

Prefix h variables included in phase 3 but not phase 4 DHS surveys:

h11a	"How long diarrhea lasted"
h31a	"How long cough lasted"
h33	"Given antibiotics"
h33a	"Given antimalarials"
h34	"Given cough syrup"
h35	"Given other pills or syrup"
h35a	"Given unknown pills or syrup"
h36	"Given injection"
h36a	"Given home remedy, herbal"
h37a	"Given CS other treatment"
h37b	"Given CS other treatment"
h37c	"Given CS other treatment"
h37d	"Given CS other treatment"
h37	"Other treatment for cough"
h38a	"Given no treatment for cough"
h38	"Given any treatment for cough"

Prefix hw, child height and weight:

Included in both phase 3 and phase 4 DHS surveys:

hw1	"Age in months"
hw2	"Weight in kilograms (1 dec.)"
hw3	"Height in centimeters (1 dec.)"
hw4	"Ht/A Percentile"
hw5	"Ht/A Standard deviations"
hw6	"Ht/A Percent of ref. median"
hw7	"Wt/A Percentile"
hw8	"Wt/A Standard deviations"
hw9	"Wt/A Percent of ref. median"
hw10	"Wt/Ht Percentile"
hw11	"Wt/Ht Standard deviations"
hw12	"Wt/Ht Percent of ref. median"
hw13	"Reason not measured"
hw14	"BCG scar on arm or shoulder"
hw15	"Height: lying or standing"
hw16	"Day of birth of child"
hw17	"Date measured (day)"
hw18	"Date measured (month)"
hw19	"Date measured (year)"
hw20	"Arm circumference (cms)"
hw21	"Arm circ/A Percentile"
hw22	"Arm circ/A Standard deviations "
hw23	"Arm c/A Percent of ref. median"
hw24	"Arm circ/Ht Percentile"
hw25	"Arm cir/Ht Standard deviations"
hw26	"Arm cir/Ht Percent ref. median"

Prefix hw variables included in phase 4 but not phase 3 DHS surveys:

hw51	"Line no. of parent/caretaker"
hw52	"Read consent statement"
hw53	"Hemoglobin level (g/dl - 1 decimal)"
hw55	"Result of measuring (Hemoglobin)"
hw56	"Hemoglobin level adjusted by altitud (g/dl - 1 decimal)"
hw57	"Anemia level"
hw58	"Agrees to referral"

Prefix hw variables included in phase 3 but not phase 4 DHS surveys: None

Prefix m, antenatal care, delivery, child nutrition:

Included in both phase 3 and phase 4 DHS surveys:

m1	"Tetanus injections bef. birth"
m2a	"Prenatal: doctor"
m2b	"Prenatal: nurse/midwife"
m2c	"Prenatal: auxiliary midwife"
m2d	"Prenatal: Nurse"
m2e	"Prenatal: Midwife"
m2f	"Prenatal: trained birth attendant"
m2g	"Prenatal: trad.birth attendant"
m2h	"Prenatal: relative"
m2i	"Prenatal: CS other person"
m2j	"Prenatal: CS other person"
m2k	"Prenatal: other resp (uncoded)"
m2l	"Prenatal: CS other"
m2m	"Prenatal: CS other"
m2n	"Prenatal: no one"
m3a	"Assistance: doctor"
m3b	"Assistance: nurse/midwife"
m3c	"Assistance: auxiliary midwife"
m3d	"Assistance: nurse"
m3e	"Assistance: midwife"
m3f	"Assistance: trained birth att."
m3g	"Assistance: trad.birth attend."
m3h	"Assistance: relative, friend"
m3i	"Assistance: CS other person"
m3j	"Assistance: CS other person"
m3k	"Assistance: other resp (uncoded)"
m3l	"Assistance: CS other"
m3m	"Assistance: CS other"
m3n	"Assistance: no one"
m4	"Duration of breastfeeding"
m5	"Months of breastfeeding"
m6	"Duration of amenorrhea"
m7	"Months of amenorrhea"
m8	"Duration of abstinence"
m9	"Months of abstinence"

m10	"Time wanted pregnancy"
m11	"Time would have waited"
m13	"Timing of 1st antenatal check"
m14	"Antenatal visits for pregnancy"
m15	"Place of delivery"
m17	"Delivery by caesarian section"
m18	"Size of child at birth"
m19	"Birth weight (kilos - 3 dec.)"
m19a	"Weight at birth recall"
m21	"Reason stopped breastfeeding"
m27	"Flag for breastfeeding"
m28	"Flag for amenorrhea"
m29	"Flag for abstinence"
m30	"At birth - prolonged labor"
m31	"At birth - excessive bleeding"
m32	"At birth - high fever/discharge"
m33	"At birth - convulsions"
m34	"When child put to breast"
m35	"Times breastfed during night"
m36	"Times breastfed during day"
m37a	"Gave child plain water"
m37b	"Gave child sugar water"
m37c	"Gave child juice"
m37d	"Gave child herbal tea"
m37e	"Gave child powder/tinned milk"
m37f	"Gave child baby formula"
m37g	"Gave child fresh milk"
m37h	"CS other liquid"
m37i	"CS other liquid"
m37j	"CS other liquid"
m37k	"CS other liquid"
m37l	"Gave child other liquid"
m37m	"Boullie"
m37n	"Baby food"
m37o	"Family table food"
m37p	"Other solid, semi-solid food"
m37q	"Food made from local grains"
m37r	"Food made from local tubers"
m37s	"Gave child eggs, fish, po"
m37t	"Gave child meat"
m38	"Drank from bottle with nipple"
m39	"Times ate other food yesterday"
m40a	"Last 7 days - plain water"
m40b	"Last 7 days - milk"
m40c	"Last 7 days - other liquids"
m40d	"Last 7 days - local grains"
m40e	"Last 7 days - local tubers"
m40f	"Last 7 days - egg/fish/po"
m40g	"Last 7 days - meat"
m40h	"Last 7 days - other solid food"
m40i	"Last 7 days - CS"

m40j	“Last 7 days - CS”
m40k	“Last 7 days - CS”
m40l	“Last 7 days - CS”
m40m	“Last 7 days - CS”
m40n	“Last 7 days - CS”
m40o	“Last 7 days - CS”

Prefix m variables included in phase 4 but not phase 3 DHS surveys:

m37u	“Times gave child other fruits/vegetables”
m37v	“Times gave child meat, poultry, eggs”
m37w	“Times gave child legumes (lentils, beans, peanuts)”
m37x	“Times gave child cheese/yogurt”
m37y	“Times gave child foods ma “
m37z	“Times gave child bread, food made from flour CS”
m37xx	“Times gave child candies, sweets CS”
m37xy	“Times gave child (shell)fish, other seafood CS”
m37xz	“Times gave child country”
m40p	“Last 7 days - other solid”
m40q	“Last 7 days - food made from local grain”
m40r	“Last 7 days - local roots”
m40s	“Last 7 days - eggs, fish”
m40t	“Last 7 days - meat”
m40u	“Last 7 days - other fruits/vegetables”
m40v	“Last 7 days - meat, poultry, eggs”
m40w	“Last 7 days - legumes (lentils, beans, peanuts)”
m40x	“Last 7 days - cheese/yogurt”
m40y	“Last 7 days - oil, fat, b”
m40z	“Last 7 days - Bread, food made from flour”
m40xx	“Last 7 days - candies, sweets CS”
m40xy	“Last 7 days - (shell)fish, other seafood CS”
m40xz	“Last 7 days - CS”
m41	“Months pregnant for last antenatal visit”
m42a	“During pregnancy - weighed”
m42b	“During pregnancy - height measured”
m42c	“During pregnancy - blood pressure taken”
m42d	“During pregnancy - urine sample taken”
m42e	“During pregnancy - blood sample taken”
m43	“Told about pregnancy complications”
m44	“Told where to go for pregnancy complications”
m45	“During pregancy, given or bought iron tablets/syrup”
m46	“Days tablets or syrup taken”
m47	“During pregnancy, had difficulty with vision during day”
m48	“During pregnancy, had night blindness”
m49a	“During pregnancy - took F “
m49b	“During pregnancy - took C “
m49c	“During pregnancy - took U “
m49d	“During pregnancy - took c “
m49e	“During pregnancy - took c “
m49f	“During pregnancy - took c “
m49g	“During pregnancy - took c “

m49x “During pregnancy - took o “
m49z “During pregnancy - took n “
m50 “After birth, health professional checked health”
m51 “Checkup after deliver timing”
m52 “After birth, health checked”
m53 “Place for checkup”
m54 “Received Vitamin A dose”
m55a “First 3 days, given milk other than breast milk”
m55b “First 3 days, given plain water”
m55c “First 3 days, given sugar/glucose water”
m55d “First 3 days, given gripe water”
m55e “First 3 days, given sugar/salt/water solution”
m55f “First 3 days, given fruit juice”
m55g “First 3 days, given infant formula”
m55h “First 3 days, given tea/infusions”
m55i “First 3 days, given honey”
m55j “First 3 days, given count”
m55k “First 3 days, given count”
m55l “First 3 days, given count”
m55m “First 3 days, given count”
m55n “First 3 days, given count”
m55x “First 3 days, given other”
m55z “First 3 days, given nothing”
m56 “Sugar added to any foods”
m57a “Antenatal care: your home”
m57b “Antenatal care: other home”
m57c “Antenatal care: CS home”
m57d “Antenatal care: CS home”
m57e “Antenatal care: govt. hospital”
m57f “Antenatal care: govt. health center”
m57g “Antenatal care: govt. health post”
m57h “Antenatal care: public mobile clinic”
m57i “Antenatal care: CS public”
m57j “Antenatal care: CS public”
m57k “Antenatal care: CS public”
m57l “Antenatal care: other public”
m57m “Antenatal care: pvt. hospital/clinic”
m57n “Antenatal care: pvt. mobile clinic”
m57o “Antenatal care: CS pvt.”
m57p “Antenatal care: CS pvt.”
m57q “Antenatal care: CS pvt.”
m57r “Antenatal care: other private”
m57s “Antenatal care: CS other”
m57t “Antenatal care: CS other”
m57u “Antenatal care: CS other”
m57v “Antenatal care: CS other”
m57x “Antenatal care: other”
m58 “Information about AIDS given at antenatal visit”
m59 “Child registered at birth”

Prefix m variables included in phase 3 but not phase 4 DHS surveys:

m12 “Antenatal card for pregnancy”
m16 “Premature birth”
m20 “Reason did not breastfeed - “
m22 “Child given other food”
m23 “Age for formula or other milk”
m24 “Age for plain water”
m25 “Age for other liquids”
m26 “Age for solid or mushy food”
m444i1 “Month of IMOVAX 1”
m444i2 “Month of IMOVAX 2”
m444yf “Month of Yellow fever”

Prefix ml, malaria:

Included in both phase 3 and phase 4 DHS surveys: none

Prefix ml variables included in phase 4 but not phase 3 DHS surveys:

ml0 “Type of bednet(s) child slept under last night”
ml1 “Times took Fansidar during fever”
ml2 “Type of visit at source for antimalarial during pregnancy”
ml11 “Child has fever now”
ml12 “Child has had convulsions in last 2 weeks”
ml13a “Fansidar taken for fever/convulsion”
ml13b “Chloroquine taken for fever/convulsion”
ml13c “Amodiaquine taken for fever”
ml13d “Quinine taken for fever/convulsion”
ml13e “Aspirin taken for fever/convulsions”
ml13f “Panadol taken for fever/convulsion”
ml13g “Ibuprofen/Acetaminophen taken for fever”
ml13h “Herbs, traditional medicine”
ml13i “Seprin taken for fever/conv”
ml13j “Cafenol taken for fever/con”
ml13k “Penicillin taken for fever/”
ml13l “Taken for fever/convulsion:”
ml13m “Taken for fever/convulsion:”
ml13x “Other taken for fever/convulsion”
ml13y “Nothing taken for fever/convulsion”
ml13z “Don’t know if or what was taken for fever/convulsion”
ml14a “Injection for fever/convulsion”
ml14b “Suppository for fever/convulsion”
ml14y “No suppository or injection for fever/convulsion”
ml14z “Don’t know if suppository or injection for fever/convulsion”
ml15a “When started Fansidar”
ml15b “Days child took Fansidar”
ml15c “First source for Fansidar”
ml16a “When started Chloroquine”
ml16b “Days child took Chloroquine”
ml16c “First source for Chloroquine”

ml17a “When started Amodiaquine”
ml17b “Days child took Amodiaquine”
ml17c “First source for Amodiaquine”
ml18a “When started Quinine”
ml18b “Days child took Quinine”
ml18c “First source for Quinine”
ml19a “For fever/conv: Consulted traditional healer”
ml19b “For fever/conv: Gave tepid sponging”
ml19c “For fever/conv: Gave herbs”
ml19d “For fever/conv: Gave medici”
ml19e “For fever/conv: Gave medici”
ml19f “For fever/conv: Taken to go”
ml19x “For fever/conv: Other”
ml19y “For fever/conv: Gave nothing”
ml19z “For fever/conv: Don’t know if something else was done”
ml101 “Type of bednet(s) slept under last night”

Prefix ml variables included in phase 3 but not phase 4 DHS surveys: none

Prefix v4, maternal height and weight, etc.:

Included in both phase 3 and phase 4 DHS surveys:

v401 “Last birth ceasarean section”
v404 “Currently breastfeeding”
v405 “Currently amenorrheic”
v406 “Currently abstaining”
v407 “Times breastfed during night”
v408 “Times breastfed during day”
v409 “Gave child plain water”
v409a “Gave child sugar water”
v410 “Gave child juice”
v410a “Gave child herbal tea”
v411 “Gave child powder/tinned milk”
v411a “Gave child baby formula”
v412 “Gave child fresh milk”
v413 “Gave child other liquid”
v413a “CS other liquid”
v413b “CS other liquid”
v413c “CS other liquid”
v413d “CS other liquid”
v414a “Boullie”
v414b “Baby food”
v414c “Family table food”
v414d “Other solid, semi-solid foods”
v414e “Food made from local grains”
v414f “Food made from local tubers”
v414g “Gave child eggs, fish, po”
v414h “Gave child meat”
v415 “Drank from bottle with nipple”
v416 “Heard of oral rehydration”

v417	"Entries in maternity table"
v418	"Entries in health table"
v419	"Entries in height/weight table"
v420	"Measurer's code"
v421	"Assistant measurer's code"
v426	"When child put to breast"
v436	"Arm circumference (cms-1d)"
v437	"Respondent's weight (kilos-1d)"
v438	"Respondent's height (cms-1d)"
v439	"Ht/A Percentile (resp.)"
v440	"Ht/A Standard deviations (resp)"
v441	"Ht/A Percent ref. median (resp)"
v442	"Wt/Ht Percent ref. median (DHS)"
v443	"Wt/Ht Percent ref. median (Fogarty)"
v444	"Wt/Ht Percent ref. median (WHO)"
v444a	"Wt/Ht Std deviations(resp) DHS"
v445	"Body mass index for respondent"
v446	"Rohrer's index for respondent"
v447	"Result of measurement of resp"
v448	"Drinking pattern with diarrhea"
v449	"Eating pattern with diarrhea"
v450a	"Diarrhea: repeat watery stool"
v450b	"Diarrhea: Any watery stool"
v450c	"Diarrhea: Repeated vomiting"
v450d	"Diarrhea: Any vomiting"
v450e	"Diarrhea: Blood in stools"
v450f	"Diarrhea: Fever"
v450g	"Diarrhea: Marked thirst"
v450h	"Diarrhea: Not eating/drinking"
v450i	"Diarrhea: Getting sicker"
v450j	"Diarrhea: Not getting better"
v450k	"Diarrhea: Country specific"
v450l	"Diarrhea: Country specific"
v450m	"Diarrhea: Country specific"
v450x	"Diarrhea: Other responses"
v450z	"Diarrhea: Does not know"
v451a	"Cough: Fast breathing"
v451b	"Cough: Difficult breathing"
v451c	"Cough: Noisy breathing"
v451d	"Cough: Fever"
v451e	"Cough: Unable to drink"
v451f	"Cough: Not eating/drinking"
v451g	"Cough: Getting sicker"
v451h	"Cough: Not getting better"
v451i	"Cough: CS"
v451j	"Cough: CS"
v451k	"Cough: CS"
v451x	"Cough: Other responses"
v451z	"Knows no sign of illness"

Prefix v4 variables included in phase 4 but not phase 3 DHS surveys:

v452a	"Under age 18 (HH report)"
v452b	"Line no. of parent/careta "
v452c	"Read consent statement"
v453	"Hemoglobin level (g/dl - 1 decimal)"
v454	"Currently pregnant"
v455	"Result of measuring (Hemoglobin)"
v456	"Hemoglobin level adjusted by altitude (g/dl - 1 decimal)"
v457	"Anemia level"
v458	"Agrees to referral"
v459	"Have bednet for sleeping"
v460	"Children under 5 slept under bednet"
v461	"Respondent slept under bednet"
v462	"Washed hands before preparing meals"
v463a	"Smokes cigarettes"
v463b	"Smokes pipe"
v463c	"Smokes other tobacco"
v463d	"Smokes CS"
v463e	"Smokes CS"
v463f	"Smokes CS"
v463g	"Smokes CS"
v463z	"Smokes nothing"
v464	"Number of cigarettes in l "
v465	"Disposal of youngest child's stools when not using toilet"
v466	"When child is seriously ill, can decide whether med tx sought"
v467a	"Getting medical help for self: know where to go"
v467b	"Getting medical help for self: getting permission to go"
v467c	"Getting medical help for self: getting money needed for tx"
v467d	"Getting medical help for self: distance to health facility"
v467e	"Getting medical help for self: having to take transport"
v467f	"Getting medical help for self: not wanting to go alone"
v467g	"Getting medical help for self: concern no female health prov"
v468	"Columns used for Last Birth Only variables"
v469a	"Times gave child plain water"
v469b	"Times gave child sugar water"
v469c	"Times gave child fruit juice"
v469d	"Times gave child herbal tea"
v469e	"Times gave child powdered "
v469f	"Times gave child commercially produced baby formula"
v469g	"Times gave child fresh milk"
v469h	"Times given tinned, powdered or fresh animal milk"
v469i	"Times given CS"
v469j	"Times given CS"
v469k	"Times given CS"
v469l	"Times gave child other liquid"
v469m	"Times given pumpkin, carrots, red/yel yams, red sweet pot."
v469n	"Times given any green leafy vegetables"
v469o	"Times given mango, papaya "
v469p	"Times given other solid foods"
v469q	"Times given food made from local grain"

v469r “Times given local roots/tubers”
v469s “Times gave child eggs, fish”
v469t “Times gave child meat”
v469u “Times gave child other fruits/vegetables”
v469v “Times gave child meat, poultry, eggs”
v469w “Times gave child legumes (lentils, beans, peanuts)”
v469x “Times gave child cheese/yogurt”
v469y “Times gave child oil, fat”
v469z “Times gave child bread, food made from flour CS”
v469xx “Times gave child candies, sweets CS”
v469xy “Times gave child (shell)fish, other seafood CS”
v469xz “Times gave child country sp”
v470a “Last 7 days - plain water”
v470b “Last 7 days - sugar water”
v470c “Last 7 days - fruit juice”
v470d “Last 7 days - herbal tea”
v470e “Last 7 days - powdered/tinned”
v470f “Last 7 days - commercially produced baby formula”
v470g “Last 7 days - fresh milk”
v470h “Last 7 days - tinned, powdered or fresh animal milk”
v470i “Last 7 days - CS”
v470j “Last 7 days - CS”
v470k “Last 7 days - CS”
v470l “Last 7 days - other liquid”
v470m “Last 7 days - pumpkin, carrots, red/yel yams, red sweet potato”
v470n “Last 7 days - any green leafy vegetables”
v470o “Last 7 days - mango, papaya”
v470p “Last 7 days - other solid “
v470q “Last 7 days - food made from local grain”
v470r “Last 7 days - local roots”
v470s “Last 7 days - eggs, fish”
v470t “Last 7 days - meat”
v470u “Last 7 days - other fruits/vegetables”
v470v “Last 7 days - meat, poultry, eggs”
v470w “Last 7 days - legumes (lentils, beans, peanuts)”
v470x “Last 7 days - cheese/yogurt”
v470y “Last 7 days - oil, fat, b”
v470z “Last 7 days - Bread, food made from flour”
v470xx “Last 7 days - candies, sweets CS”
v470xy “Last 7 days - (shell)fish, other seafood CS”
v470xz “Last 7 days - CS”

Prefix v4 variables included in phase 3 but not phase 4 DHS surveys:

v414 “Gave child solid or mushy food”
v422 “Ever prepared ORS solution”
v423 “Quantity of water for ORS solution”
v423a “ORS source: government hosp.”
v424b “ORS source: govt health center”
v424c “ORS source: govt health post”
v424d “ORS source: mobile clinic”

v424e “ORS source: comm.health worker”
v424f “ORS source: CS public sector”
v424g “ORS source: CS public sector”
v424h “ORS source: CS public sector”
v424i “ORS source: CS public sector”
v424j “ORS source: private hosp/clinic.”
v424k “ORS source: private pharmacy”
v424l “ORS source: private doctor”
v424m “ORS source: private mobile clinic”
v424n “ORS source: comm.health worker”
v424w “ORS source: IPPF Center (ASBEF)”
v424o “ORS source: Local Nurse”
v424p “ORS source: CS med.priv sector”
v424q “ORS source: CS med.priv sector”
v424r “ORS source: CS med.priv sector”
v424s “ORS source: shop”
v424t “ORS source: traditional practitioner”
v424u “ORS source: Relatives”
v424v “ORS source: Mosque, church”
v424x “ORS source: Other”
v424y “ORS source: Unknown”
v425 “Home fluid preparation teacher”
v427 “Duration breastfeeding preparation”
v428 “Months breastfeeding preparation”
v429 “Flag for breastfeeding preparation”
v430 “Duration of amenorrhea preparation”
v431 “Months of amenorrhea pre 8?”
v432 “Flag for amenorrhea pre 8? “
v433 “Duration of abstinence pre “
v434 “Months of abstinence pre 8 “
v435 “Flag for abstinence pre 8?

DHS Methodological Reports Series

1. Institute for Resource Development. 1998. **An Assessment of DHS-I Data Quality.**

2. Macro International Inc. 1993. **An Assessment of the Quality of Health Data in DHS-I Surveys.**

3. Aliaga, Alfredo and Pradip K. Muhuri. 1994. **Methods of Estimating Contraceptive Prevalence Rates for Small Areas: Applications in the Dominican Republic and Kenya.**

4. Landers, Alynne and Melissa McNiff. 1994. **Comparability of Questionnaires.**

5. Pullum, Thomas W. 2006. **An Assessment of Age and Date Reporting in the DHS Surveys, 1985-2003.**

6. Pullum, Thomas W. 2008. **An Assessment of the Quality of Data on Health and Nutrition in the DHS Surveys, 1993-2003.**