Consistency of Reporting Terminated Pregnancies in DHS Reproductive Calendars

Kerry L.D. MacQuarrie, The DHS Program, Avenir Health
William Winfrey, Avenir Health
Jacqueline Meijer-Irons, Independent consultant

For submission to:
PAA Annual Meeting, Austin, TX, April 10-14, 2019
Session 101: Abortion
Session 123: Innovations in Measurement for Fertility, Family Planning, and Sexual and Reproductive Health

Corresponding author: Kerry L.D. MacQuarrie, The DHS Program, ICF, 530 Gaither Road, Suite 500, Rockville, MD 20850; phone: +1 301-572-0282; email: Kerry.Macquarrie@icf.com

Short Abstract

This study examines the consistency of reporting of terminated pregnancies in DHS reproductive calendars. We examine two measures of consistency: within-survey consistency and between-survey consistency. We find that 70-80% of surveys are inconsistent by either or both of these measures. Reporting of terminated pregnancies deteriorates in earlier periods of the calendar. Africa has a higher proportion of inconsistent surveys (78%), whereas in the other regions about two-thirds of surveys are classified as inconsistent. Consistency appears to vary with type of history (pregnancy or birth), direction of the history, and the inclusion of questions that distinguish between miscarriage and induced abortion. Survey length and mode of data collection appear unrelated to either consistency or inconsistency of reporting terminated pregnancies. Our findings suggest the need for more experimentation with and rigorous assessment of modifications to the calendar and modes of data collection.

1. Introduction

The Demographic and Health Surveys are primary sources of data on numerous fertility-related indicators, reproductive behavior, and experiences, including terminated pregnancies. However, it has been frequently noted in the literature that survey-based estimates of miscarriages, stillbirths, and induced abortion are underestimated, particularly when the survey does not singularly focus on such events (Huntington, Mensch, and Miller 1996; Jones and Forrest 1992; Paxman et al. 1993; Rossier 2003; Toulemon and Leridon 1992). For example, one study using modified narrative calendar method found an induced abortion ratio five times greater than that estimated from a traditional DHS reproductive calendar (Edmeades et al. 2010). A variety of factors—time since the event, grief, trauma, stigma, and social desirability among them—all influence recall and the willingness to report terminated pregnancies in a survey setting.

Nonetheless, research has shown that calendar-type data collection instruments can improve the recording reproductive events over simple direct questions or tabular arrays of data collection (Axinn, Pearce, and Ghimire 1999; Belli et al. 2004; Glasner and Van der Vaart 2009; Luke, Clark, and Zulu 2011; Martyn and Belli 2002). This is particularly the case where calendar instruments make use of memory prompts and anchor events among other timing cues (Axinn, Pearce, and Ghimire 1999; Belli 1998; Edmeades et al. 2010; Glasner and Van der Vaart 2009).

The DHS has employed a reproductive calendar since 1986 to collect data on births, pregnancies, and episodes of contraceptive use. Although numerous studies have evaluated such a calendar with regard to the quality and consistency of data on contraception, perinatal mortality, age displacement, and birth intervals, this is believed to be the first study examining the consistency of calendar data on terminated pregnancies.

1.1. The Birth History

The DHS Program has used two different instruments to measure women's reproductive events: (1) birth or pregnancy histories and (2) the reproductive calendar[1]. The use of either a birth history or a pregnancy history within a single DHS survey is mutually exclusive: a survey administers one or the other but not both. In contrast, the reproductive calendar can accompany a birth or pregnancy history or be omitted altogether.

A birth history inquires about all live births a woman has experienced and, for each, if the birth is multiple/single, the birthdate, sex, current age, name of child, whether the child is alive and if she/he died how old she/he was when she died.

The pregnancy history queries a woman about pregnancies and their result (rather than births). The respondent is asked the following about each pregnancy: single/multiple pregnancy, how the pregnancy ended (live, not live), name, sex, birth month/day, and currently alive or not. If the pregnancy ended without a live birth, the woman is asked about the month and year and the duration of the pregnancy when the pregnancy ended. Occasionally, the respondent is asked to report herself whether the pregnancy ended in a stillbirth, miscarriage, or induced abortion.

Pregnancy or birth histories may be a forward reported history: reporting starts with women's first pregnancy/birth and move to their most recent pregnancy/birth; or a backward reported history: reporting begins with the most recent pregnancy/birth and move to their first pregnancy/birth.

1.2. The Reproductive Calendar

The reproductive calendar is a separate instrument that may accompany either form of tabular history. This calendar includes monthly data for the 60-72 months preceding the survey on respondents' pregnancies, births, miscarriages/terminations, and patterns of contraceptive use. After completing the birth/pregnancy history, the interviewer marks a B in the appropriate cells of the reproductive calendar. The respondent is asked the duration of each pregnancy for each birth and P's are entered as appropriate in the calendar. The final month of pregnancy is the month marked B in the calendar. Therefore the duration of a pregnancy is the number of P's plus one.

[1] The description of these instruments here draws heavily from (Bradley, Winfrey, and Croft 2015).

Next the interviewer enters into the calendar data on any pregnancies that did not end in a live birth from either the pregnancy history or the supplemental questions in surveys with a birth history. For each such pregnancy the interviewer enters a T in the calendar in the cell corresponding to the month in which the pregnancy ended and P's for the months corresponding to the duration of the pregnancy. A T is recorded for all types of pregnancies ending other than in a live birth, whether miscarriage, induced abortion, or stillbirth. Stillbirths are differentiated from miscarriages and induced abortions based on the duration of pregnancy (≥7 months vs. ≤6 months). No differentiation is made in the calendar between miscarriages and induced abortions. Except for the few surveys with a pregnancy history in which respondents nominate this information, we cannot distinguish miscarriages from induced abortions.

2. Methods and Data

In this study, we analyze 162 standard DHS surveys conducted between 1991 and 2017 from 62 countries that administered a reproductive calendar. We analyze a standardized 60-month period of observation in each survey's calendar.

2.1. Termination ratio

Our primary measure is a termination ratio, which in any given month of the calendar is the ratio of terminated pregnancies to live births, expressed per 100 live births.

$$\left(\frac{terminated\ pregnancies}{live\ births}\right) \times 100$$

We aggregate monthly termination ratios into termination ratios for five 12-month intervals in two ways. The first is with regard to the woman's month of interview. We calculate the termination ratio for each 12-month interval (0-11 months, 12-23 months, 24-35 months, 36-47 months, and 48-59 months) from the month of interview during the observation period. The second is with regard to calendar month and year. Thus, for a survey's calendar beginning in January 2010, and June 2015 being the most recent month in the observation period, we calculate the termination ratio for the following 12-month intervals: July 2014–June 2015, July 2013–June 2014, July 2012–June 2013, July 2011–June 2012, and July 2010–June 2011.

2.2. Within-survey consistency

In this study, we first examine within-survey consistency in the termination ratios, in two ways. First, for all 162 surveys, we generate plots to describe termination ratios over the course of the 60-month observation period into five 12-month intervals. Next, we apply a statistical test of the equivalence of termination ratios estimated at the most recent and earliest interval in the calendar.

Ceteris paribus, we would expect that the termination ratio would be constant over the course of a 60-month calendar. A Wald test of equivalence tests whether the relative odds of termination during the 0-11 months preceding the survey is equivalent to that in the 48-59 months preceding the survey. Thus, with a p-value of the Wald test of less than 0.05, we reject the null hypothesis that the termination ratios at the start and the end of the calendar are statistically equivalent and deem the survey "inconsistent."

2.3. Between-survey consistency

Of the 62 countries that have at least one survey with a calendar, more than half—34 countries—have multiple calendars that are consecutive or overlap in time. For these surveys, we supplement the assessment of within-survey consistency by assessing between-survey consistency in termination ratios from one survey to the next. This analysis comprises 121 calendars.

For between-survey analysis, we first use graphical analysis to examine the continuity of termination ratios from one survey to the next. In particular, we examine the termination ratios calculated from two surveys at adjoining (or overlapping) moments of the surveys. We follow the graphical analysis with a formal test of between-survey

consistency. With each pair of consecutive surveys, we calculate the absolute and relative difference in the termination ratios between the earliest observed interval of one survey and the most recent observed interval of the preceding survey. If the relative difference is greater than the absolute value of 20.6%, we reject the null hypothesis that there is no significant difference between the two termination ratios, meaning that between-survey inconsistency is evident. We use a cutoff of 20.6% as the threshold because it is the minimum difference for which a significant difference in termination ratios can be detected.

Fertility patterns, population health status, and health care systems may change over time, influencing the real ratio of pregnancies that end in termination to those ending in live birth. Such secular trends may or may not be detectable in a single 5-year calendar. In the absence of a secular trend, calendars should exhibit both within-survey and between-survey consistency. In the presence of a secular trend, a survey may be marked by within-survey inconsistency (a significantly increasing or decreasing termination ratio within each survey calendar), but should nonetheless exhibit between-survey consistency (adjoining or overlapping moments of the calendars should be statistically similar). In such a case, the change in the termination ratio can be attributed to a secular trend. However, if between-survey inconsistency persists, change in the termination ratio cannot be attributed to a secular trend, but to reporting inconsistency.

2.4. Patterns across survey characteristics

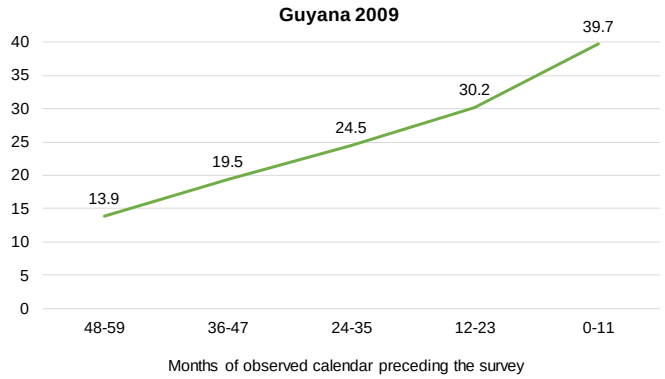
Based on the two tests of within-survey and between-survey consistency, we classify each survey as either "consistent" or "inconsistent." A survey is classified as "inconsistent" if both the within-survey and between-survey tests indicate inconsistency, and it is classified as "consistent" if either the within-survey or between-survey test, or both, suggests consistency. If only one test is available, a survey is classified according to the results of the within-survey test.

In our final analysis, we then disaggregate consistent and inconsistent surveys according to six survey characteristics. The following comparisons are made: (1) birth history compared to pregnancy history; (2) forward or backward reporting of events in the birth/pregnancy histories; (3) inclusion of questions to differentiate types of termination, namely miscarriage and induced abortion; (4) computer-assisted personal interview (CAPI) survey versus pen-and-paper survey; (5) longer versus shorter questionnaire; and (6) phase of the DHS Program.

3. Preliminary results

We find that overall about 80% of surveys (128 of 162 surveys) indicate within-survey inconsistency, while about 70% of surveys (55 of 78 surveys) indicate between-survey inconsistency. Almost without exception, our analysis of within-survey consistency indicates higher termination ratios reported closer to the date of interview. Guyana 2009 is an illustrative example of within-survey inconsistency.

Figure 1 Ratios of terminated pregnancies to live births (per 100)



We identify four patterned groups of surveys regarding secular trends and between-survey consistency: Pattern 1 is inconsistent surveys with no evidence of a secular trend. Pattern 2 is consistent surveys with no evidence of secular

trend. Pattern 3 is consistent surveys with evidence of a secular trend and Pattern 4 is inconsistent surveys with evidence of a secular trend.

The largest group of surveys shows within-survey and between-survey inconsistency with no evidence of a change in the underlying termination ratio over time (Pattern 1). Next in number are surveys showing within-survey and between-survey inconsistency accompanied by a change (usually an increase) in termination ratios over time. However, the magnitude of the trend is insufficient to fully explain the inconsistency observed in reported termination ratios (Pattern 4). Zimbabwe (Figure 2) and Ghana (Figure 3) are illustrative examples of Pattern 1 and Pattern 4, respectively. Just 6 countries exemplify Patterns 2 and 3 with between-survey consistency.

Figure 2 Ratios of terminated pregnancies to live births (per 100) in consecutive DHS calendars, Zimbabwe

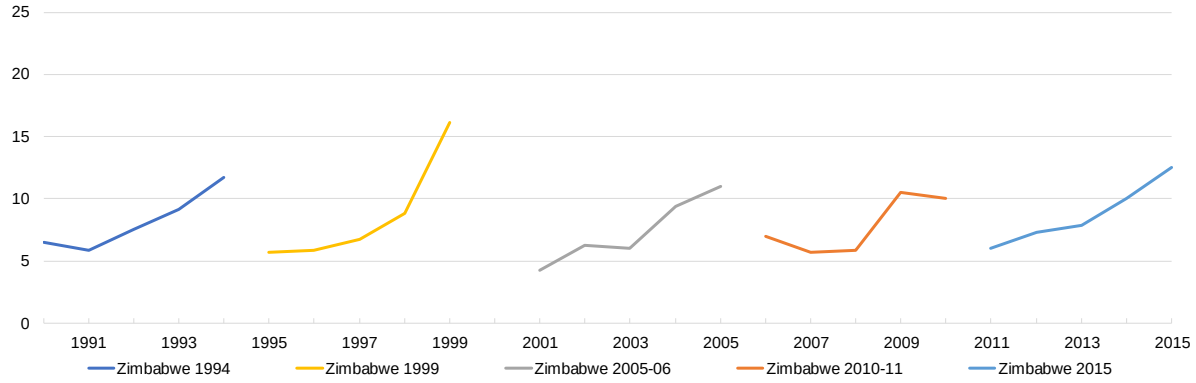
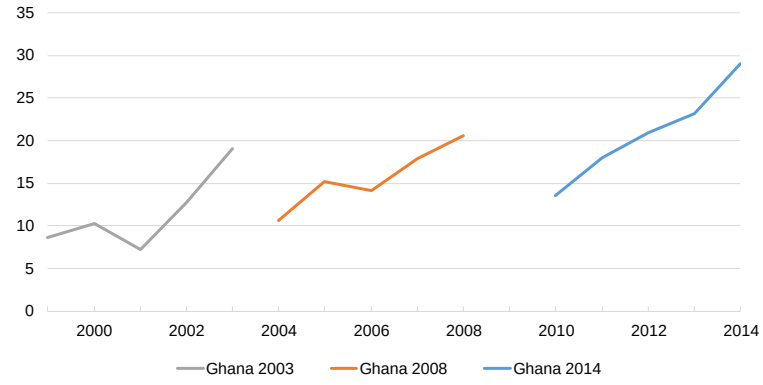


Figure 3 Ratios of terminated pregnancies to live births (per 100) in consecutive DHS calendars, Ghana



Africa has a higher proportion of inconsistent surveys (78%), whereas in each of the other three regions about two-thirds of surveys are classified as inconsistent. The proportion of surveys marked by between-survey inconsistency was substantially lower than that with within-survey inconsistency in Asia. The highest proportion of consistent surveys occurred during DHS-4. Consistency of reporting of terminated pregnancies in the DHS calendar appears to be somewhat more likely with pregnancy histories, backward reported histories, and the inclusion of questions that distinguish miscarriage and induced abortion, though the potential for confounding exists. Surveys with these characteristics are few in number and are clustered in certain regions. It is not possible to disentangle whether more consistent reporting is due to a survey characteristic or a characteristic of those countries which tend to adopt these modifications in their surveys. Survey length and mode of data collection appear unrelated to the consistency or inconsistency of a survey in reporting pregnancy termination.

Our findings, and the study's limitations, suggest more experimentation with modifications to the calendar (including type and direction of histories, mode of data collection, and more), perhaps with survey characteristics being randomly assigned within or across surveys, and accompanied by rigorous assessment.

References

Axinn, W.G., L.D. Pearce, and D. Ghimire. 1999. "Innovations in Life History Calendar Applications." *Social Science Research* 28(3):243-264.

Belli, R.F. 1998. "The Structure of Autobiographical Memory and the Event History Calendar: Potential Improvements in the Quality of Retrospective Reports in Surveys." *Memory* 6(4):383-406.

Belli, R.F., E.H. Lee, F.P. Stafford, and C.-H. Chou. 2004. "Calendar and Question-List Survey Methods: Association between Interviewer Behaviors and Data Quality." *Journal of Official Statistics* 20(2):185.

Bradley, S.E.K., W. Winfrey, and T.N. Croft. 2015. *Contraceptive Use and Perinatal Mortality in the Dhs: An Assessment of the Quality and Consistency of Calendars and Histories*. DHS Methodological Reports No. 17. Rockville, Maryland, USA: ICF International. Available at http://dhsprogram.com/pubs/pdf/MR17/MR17.pdf.

Edmeades, J., L. Nyblade, A. Malhotra, K. MacQuarrie, S. Parasuraman, and S. Walia. 2010. "Methodological Innovation in Studying Abortion in Developing Countries: A ‘‘Narrative’’ Quantitative Survey in Madhya Pradesh, India." *Journal of Mixed Methods Research* 4(3):176-198.

Glasner, T., and W. Van der Vaart. 2009. "Applications of Calendar Instruments in Social Surveys: A Review." *Quality and Quantity* 43(3):333-349.

Huntington, D., B. Mensch, and V.C. Miller. 1996. "Survey Questions for the Measurement of Induced Abortion." *Studies in Family Planning*:155-161.

Jones, E.F., and J.D. Forrest. 1992. "Underreporting of Abortion in Surveys of U.S. Women: 1976 to 1988." *Demography* 29(1):113-126.

Luke, N., S. Clark, and E.M. Zulu. 2011. "The Relationship History Calendar: Improving the Scope and Quality of Data on Youth Sexual Behavior." *Demography* 48(3):1151-1176.

Martyn, K.K., and R.F. Belli. 2002. "Retrospective Data Collection Using Event History Calendars." *Nursing research* 51(4):270-274.

Paxman, J.M., A. Rizo, L. Brown, and J. Benson. 1993. "The Clandestine Epidemic: The Practice of Unsafe Abortion in Latin America." *Studies in family planning*:205-226.

Rossier, C. 2003. "Estimating Induced Abortion Rates: A Review." *Studies in family planning* 34(2):87-102.

Toulemon, L., and H. Leridon. 1992. "Maîtrise De La Fécondité Et Appartenance Sociale: Contraception, Grossesses Accidentelles Et Avortements [Birth Control and Social Group: Contraception, Accidental Pregnancies
and Abortion]." *Population* 47(1):1-45.