ANALYSIS OF RESPONSE BIAS AND ADJUSTMENT OF HIV PREVALENCE *Appendix G*

BACKGROUND

The 2004 MDHS separately administered the individual interview and collected blood samples for subsequent HIV testing. Respondents had the opportunity to accept or decline to participate in the survey as well as to accept or decline to provide a blood sample. The overall rate of survey participation was 96 percent for women and 86 percent for men (Table 1.2). As discussed in Chapter 12, 70 percent of women consented to the HIV test, as did 63 percent of the men (Table 12.1).

In any survey such as the 2004 MDHS, nonparticipation, and the potential for an associated bias, is a concern. Those who did not participate in the HIV test, for example, may be different in their characteristics or behaviour from those who consented to provide a blood sample. Thus, it has become standard procedure to conduct an analysis of those who are not tested in order to look for potential biases.

Because the characteristics of women and men who were interviewed in the survey but not tested are known, statistical analysis and modeling can be used to provide an estimate of their HIV status, using the relationships found for those who were both interviewed and tested. For (those) eligible respondents who were neither interviewed nor tested, basic information from the household questionnaire can be used in a simpler model to predict their HIV status.

The worldwide Demographic and Health Surveys programme has conducted such analyses for each of the DHS surveys in which HIV data were collected. A description of the general method and the results for several countries can be found in a review paper (Mishra et al., 2006a).

A review of the 2004 MDHS results reveals an overall survey participation rate and a HIV testing rate that fall within the range of experience in other DHS survey countries. However, in Malawi, an additional concern is the low HIV testing rate (39 percent for women, 38 percent for men) and the resulting anomalous HIV prevalence results for Lilongwe District (Tables 12.2.1, 12.2.2, and 12.4).

Because of this, the standard adjustment approach was modified for Malawi. For all of Malawi with the exception of Lilongwe, the standard approach is used, combining the empirical results from those interviewed and tested with the estimated HIV results from the separate models for those interviewed and not tested and for those not interviewed and not tested (see Table G.1). For the subsample of Lilongwe District, the "complete model" was used to estimate the HIV status of respondents with completed interview, and the simple model was used to estimate the HIV status for those who were not interviewed and not tested (see Table G.2). The overall adjusted HIV figure for Malawi is obtained by combining the various component parts.

The following sections provide details of the methodology and adjusted HIV prevalence figures by selected characteristics.

ANALYSIS

In the 2004 MDHS, HIV test results were linked anonymously to all of the respondent's characteristics collected in the questionnaires after scrambling the household and cluster identification codes. To estimate the extent of nonresponse bias and its potential impact on the observed HIV rates, all eligible respondents are divided into three groups: (1) interviewed and tested; (2) not interviewed, not tested; and (3) interviewed, not tested.

To evaluate the effect of nonresponse bias on the survey estimates, HIV prevalence is predicted for the not-tested groups (Groups 2 and 3) based on multivariate models of HIV for those who were tested, using a common set of predictor variables. A logistic regression model is used to calculate predicted HIV prevalence separately for the "not-interviewed, not-tested" and "interviewed, not-tested" groups. Predictions for the "not-interviewed, not-tested" group are based on a limited set of variables (only from the household questionnaire), but predictions for the "interviewed, not-tested" group account for selected individual sociodemographic and behavioural characteristics of the respondents, as collected in the survey. As noted above, for the Lilongwe subsample, all interviewed respondents were treated as if they were in Group 3.

Variables for predicting prevalence in the "not-interviewed, not-tested" group include, age, education, wealth quintile, residence, and geographic region. Additional variables in the "complete model" for predicting prevalence in the "interviewed, not-tested" group include: marital union, current work status, media exposure, religion, STI or STI symptoms in past 12 months, cigarette smoking/tobacco use, age at first sex, number of sex partners in past 12 months, higher-risk sex in past 12 months, condom use at last sex in past 12 months, and willingness to care for a family member with AIDS. In models for the total sample (males and females combined), an additional sex variable was included.

Data processing for the 2004 MDHS was done using CSPro, a software package developed by the MEASURE DHS programme and the U.S. Census Bureau. Multivariate analyses used STATA version 8.0. All analyses are carried out separately for males and females. Adjusted HIV prevalence is calculated as a weighted average of observed prevalence among those who were tested and predicted prevalence in the two groups of not-tested respondents. Sampling weights were applied in accordance with standard DHS procedures. HIV sampling weights were used for tested individuals (Group 1), individual sampling weights were applied to persons who were "interviewed, not tested" (Group 3), and household sampling weights were used for the "not interviewed, not tested" (Group 2), respectively. Further details on the analysis are included in Mishra et al. (2006b).

RESULTS

Table G.1 shows the observed and adjusted HIV prevalence for women, men, and total samples for Malawi excluding Lilongwe, for Lilongwe, and for Malawi total. Overall, the combined adjustment for nonresponse and not tested raises the HIV prevalence by about 0.8 percentage points above the observed level (12.7 percent compared with 11.8 percent). For women, the adjustment adds about a full percentage point onto the observed prevalence, raising the figure from 13.3 percent to 14.4 percent. For men, the effect of adjustment is smaller, adding only about 0.6 percentage points (10.2 and 10.8, respectively).

Part of the rationale for adjustment was to correct the figures for Lilongwe, where a low HIV testing rate and implausibly low HIV results for those who were tested was observed, especially

among women. The adjusted HIV results for Lilongwe have indeed raised the prevalence by a considerable degree. For women, the observed HIV prevalence of 1.6 percent is raised to 11.5 percent by the adjustment. For men, the observed rate of 5.5 percent is increased to 9.2 percent. The resulting adjusted figures for Lilongwe are more in line with the expected HIV levels based on the ANC Sentinel Surveillance results from sites across Malawi. In addition, the adjustment process makes the results for women and men in Lilongwe consistent with the patterns by sex observed in other districts and regions in Malawi.

An examination of the 95% confidence levels for observed and adjusted figures shown in Table G.1 indicates that, with the exception of the results for Lilongwe, the differences between the observed and adjusted figures are not statistically significant.

Table G.1 Observed and adjusted HIV prevalence among women and men age 15-49, Malawi 2004

	Observed			Adjusted		
	Prevalence	95% confidence interval		Prevalence	95% confidence interval	
Geographic area	(R)	R-2SE	R+2SE	(R)	R-2SE	R+2SE
WOMEN						
Malawi, excluding Lilongwe	15.1	13.8	16.4	14.8	13.8	15.8
Lilongwe	1.6	0.0	4.2	11.5	10.0	13.1
Malawi total	13.3	12.1	14.6	14.4	13.5	15.3
MEN						
Malawi, excluding Lilongwe	11.1	9.8	12.4	11.2	10.2	12.1
Lilongwe	5.5	0.9	10.1	9.2	7.8	10.6
Malawi total	10.2	9.0	11.5	10.8	10.0	11.7
TOTAL						
Malawi, excluding Lilongwe	13. 2	12.3	14.2	13.1	12.4	13.7
Lilongwe	3.7	1.0	6.4	10.3	9.3	11.3
Malawi total	11.8	11.0	12.7	12.7	12.0	13.3

Table G.2 compares observed and adjusted HIV prevalence for women and men according to a number of respondent and household characteristics. We note that the differences between observed and adjusted figures are relatively small, for the most part. Observed patterns tend to be maintained despite the adjustment. Perhaps the most notable change evident in this table is in the data specific to the Central region, which includes Lilongwe district. For example, HIV prevalence for women in the Central region increases from 6.6 percent to 9.7 percent following the adjustment.

Table G.2. Observed and adjusted HIV prevalence among women and men age 15-49 by selected background characteristics, Malawi 2004

Background characteristic	Women		Men		Total	
	Observed	Adjusted	Observed	Adjusted	Observed	Adjusted
Age						
15-19	3.7	3.8	0.4	0.4	2.1	2.2
20-24	13.2	14.9	3.9	4.5	9.5	10.6
25-29	15.5	16.5	9.8	12.3	12.6	14.1
30-34	18.1	21.3	20.4	18.4	19.2	20.0
35-39	17.0	18.6	18.4	22.9	17.7	20.6
40-44	17.9	18.8	16.5	18.7	17.2	18.9
45-49	13.3	14.7	9.5	12.5	11.6	13.3
Residence						
Urban	18.0	20.9	16.3	16.0	17.1	18.3
Rural	12.5	13.0	8.8	9.4	10.8	11.3
Region						
Northern	10.4	10.4	5.4	5.2	8.1	8.0
Central	6.6	9.7	6.4	7.5	6.5	8.7
Southern	19.8	19.6	15.1	15.1	17.6	17.3
Education						
No education	13.6	14.5	9.2	10.8	12.3	13.2
Primary 1-4	12.3	13.0	6.5	7.4	9.7	10.4
Primary 5-8	13.0	13.8	10.8	11.4	11.8	12.4
Secondary +	15.7	18.0	14.6	14.2	15.1	16.2
Wealth quintile						
Lowest	10.9	11.5	4.4	5.4	8.3	9.0
Second	10.3	10.0	4.6	4.1	7.6	7.1
Middle	12.7	13.1	12.1	11.9	12.4	12.4
Fourth	14.6	15.9	11.7	13.5	13.2	14.6
Highest	18.0	20.6	14.9	15.4	16.4	18.0
Total	13.3	14.4	10.2	10.8	11.8	12.7

DISCUSSION

Minimizing nonresponse is a major challenge for all population-based surveys. The main reasons for nonresponse are refusal and absence. The analysis of nonresponse in Malawi shows consistency with data from five other DHS survey countries with linked HIV data (Burkina Faso, Cameroon, Ghana, Kenya, and Tanzania) and indicates that nonresponse does not significantly bias national HIV estimates from population-based surveys (Mishra et al., 2006a). The overall effect of nonresponse on observed national HIV prevalence estimates is small.

It is important to recognize that these adjustments only partially address nonresponse bias. The estimates can be adjusted only to the extent that the sociodemographic and behavioural characteristics included in the analysis are correlated with the risk of HIV infection in each country. Another limitation is that the adjustments for the "not-interviewed, not-tested" respondents (mostly absentees) are based on limited information. Absence from the household may itself be associated with increased risk of HIV infection. From the available data, it is not possible to appropriately adjust for bias due to absence.