SAMPLE DESIGN

A.1 INTRODUCTION

The 2001-2002 Zambia Demographic and Health Survey (ZDHS) is a comprehensive nationally representative population and health survey carried out by the Central Statistical Office in partnership with the Central Board of Health. ORC Macro provided financial and technical assistance for the survey through the USAID-funded MEASURE *DHS+* programme. Additional funding for the ZDHS was received from the Government of Japan, UNFPA, and DANIDA.

A.2 SURVEY OBJECTIVES

The principal objective of the ZDHS is to provide current and reliable data on fertility and family planning behaviour, child mortality, children's nutritional status, the utilization of maternal, child health services, knowledge and prevalence of HIV and syphilis.

The population covered by the 2001-2002 ZDHS is defined as the universe of all women age 15-49 in Zambia and all men age 15-59. A sample of households was selected and all women age 15-49 identified in the households were interviewed. In addition, in a subsample of one-third of all the households selected for the ZDHS, all men 15-59 were eligible to be interviewed if they were either permanent residents or visitors present in the household on the night before the survey.

A.3 SAMPLE DOMAINS

The Zambia DHS collected demographic and health information from a nationally representative sample of women and men age 15-49 and 15-59, respectively. The primary focus of the 2001 Zambia DHS is to provide estimates of key population and health indicators, including fertility and mortality rates, for the country as a whole, and for urban and rural areas separately. Also, the sample was designed to provide estimates of key variables for the nine provinces, namely, 1) Central, 2) Copperbelt, 3) Eastern, 4) Luapula, 5) Lusaka, 6) Northern, 7) North-Western, 8) Southern, and 9) Western. In addition, the sample provides basic information for a total of 12 combined districts (not each separately) that are the special focus of the Zambia Integrated Health Programme (Livingstone, Kalomo, Chibombo, Kabwe Urban, Ndola Urban, Kitwe, Chipata, Lundazi, Chama, Kasama, Samfya, and Mwense).

A.4 SAMPLE FRAME

Zambia is divided into nine provinces. In turn, each province is subdivided in districts, each district into constituencies, and each constituency into wards. In addition to these administrative units, during the 2000 population census, each ward was subdivided into convenient areas called census supervisory areas (CSAs), and in turn each CSA into standard enumeration areas (SEAs). In total Zambia has 72 districts, 150 constituencies, 1,289 wards, about 4,400 CSAs, and about 16,400 SEAs. Preliminary information on the counts of households and population, as well as cartographic materials were available from the 2000 population census for the SEAs. Therefore, the sample frame for this survey was the list of SEAs developed from the 2000 population census.

A.5   STRATIFICATION

In the preliminary census frame, the SEAs were grouped by CSAs, by CSAs within a ward, by wards within a constituency, by constituencies within a district and by districts within a province for purposes of the ZDHS. The SEAs were further stratified separately by urban and rural areas within each province.

A.6   SAMPLE ALLOCATION

The primary sampling unit (PSU), the cluster for the 2001-2002 ZDHS, is defined on the basis of SEAs from the census frame. A minimum requirement of 85 households for the cluster size was imposed in the design. If an SEA did not have 85 households, it was combined with an adjacent SEA; thus, the ZDHS cluster comprised one or more SEAs. The number of clusters in each district was not allocated proportional to the total population due to the need to present estimates by each of the nine provinces. Zambia is a country where two-thirds of the population reside in rural areas, and one-third in urban areas. Table A.1 shows the proportional and the squared root allocations of 320 clusters

Table A.1. Proportional allocation and squared root allocation of 320 clusters by province

Province	Percentage of households (Census 2000)	Allocation for a sample of 320 clusters		
		Proportional	Square root	Adjusted
Central	9.81	31	35	31
Copperbelt	15.32	49	44	44
Eastern	13.57	44	37	40
Luapula	9.04	29	32	30
Lusaka	13.35	43	40	40
Northern	14.80	47	40	43
North-Western	6.10	19	26	30
Southern	9.90	32	35	32
Western	8.10	26	30	30
Total	100.00	320	320	320

The target for the 2001-2002 ZDHS sample was 8,000 completed interviews. Based on the level of non-response found in the 1996 ZDHS, to achieve this target, approximately 8,200 households were selected, with all women age 15-49 being interviewed. The target was to reach a minimum of 750 completed interviews per province. In each province the number of households was distributed proportionately among the urban and rural areas. Table A.2 shows the distribution of about 8,200 households by province.

Table A.2. Expected number of selected households to reach the target of completed interviews by province

Province	Expected number of completed interviews 2001-2002 ZDHS	Completed interviews 1996 ZDHS	Selected households 1996 ZDHS	Expected households selected 2001-2002 ZDHS
Central	775	748	861	892
Copperbelt	1,100	1,129	845	823
Eastern	1,000	1,118	1,154	1,032
Luapula	750	896	866	725
Lusaka	1,000	1,074	867	807
Northern	1,075	783	941	1,292
North-Western	750	567	861	1,139
Southern	800	846	711	672
Western	750	860	910	794
Total	8,000			8,176

The urban-rural distribution was also considered in distributing the sample. The selected households were distributed in 320 clusters in Zambia, 100 clusters in the urban areas, and 220 clusters in the rural areas. Table A.3 shows the distribution of urban and rural clusters selected for the 2001-2002 ZDHS.

Table A.3 Final allocation of women 15-49 with completed interviews and clusters in each province by urban and rural residence

Province	Expected number of completed interviews	Number of clusters		
		Urban	Rural	Total
Central	775	8	23	31
Copperbelt	1,100	32	12	44
Eastern	1,000	3	37	40
Luapula	750	5	25	30
Lusaka	1,000	33	7	40
Northern	1,075	5	38	43
North-Western	750	3	27	30
Southern	800	8	24	32
Western	750	3	27	30
Total	8,000	100	220	320

Under this final allocation, the 12 combined districts of the Zambia Integrated Health Programme have 77 selected clusters, 36 in urban areas and 41 in rural areas.

A.7   SAMPLE SELECTION

The 2001-2002 ZDHS sample was selected using a stratified two-stage cluster design consisting of 320 clusters, 100 in urban and 220 in rural areas. Once the number of households was allocated to each combination of province by urban and rural areas, the number of clusters was calculated based on an average sample take of 25 completed interviews among women 15-49 years. In each urban or rural area in a given province, clusters were selected systematically with probability proportional to the number of households in each cluster. The selection was done using the following formula:

$$P_{1i} = (a * M_i) / (\Sigma M_i)$$

where

a is the number of clusters to be selected in the given combination of province by residence area,

M_i is the number of households of the ith clusters reported in the 2000 summary census information,

ΣM_i is the number of households in the urban (or rural) area in the province according to the 2000 summary census information.

In each selected cluster, a complete household listing operation was carried out and households were selected to achieve a self-weighted sampling fraction in each province. However, since the 2001-2002 ZDHS sample is unbalanced among provinces, a final weighing adjustment procedure is required to provide estimates at every other domain of study.

In a given province, if the overall sampling fraction (f) has been calculated, and if c_i is the number of households selected out of the total households (L_i)—found in the 2001 listing process—for the ith cluster, then the self-weighting condition can be expressed as

$$f = P_{1i} * (c_i / L_i)$$

The final number of households in the ith cluster could be calculated as

$$c_i = (f * L_i) / P_{1i}$$

and the household selection interval for the ith cluster is given as

$$I_i = L_i / c_i$$

$$I_i = P_{1i} / f$$

A.8   RESPONSE RATES

Information on the household and individual interviews for women and men is presented in Tables A.4 and A.5. A total of 8,050 potential households were selected for the 2001-2002 ZDHS, of which 7,261 were actual households. Household interviews were completed for 98.2 percent of the actual households. A total of 7,944 eligible women were found in these households, and 96.4 percent of the women were successfully interviewed. The overall response rate for women was 94.6 percent.

Similarly, a total of 2,658 potential households were selected for the men's survey, of which 2,408 were actual households. Household interviews were completed for 98.3 percent of the actual households. A total of 2,418 eligible men were found in these households and interviews were completed with 88.7 percent of the men. The overall response rate for men was 87.2 percent.

There is no difference by urban-rural residence in the overall response rate for eligible women; however rural men are more likely than urban men to have completed the interview (89.4 percent and 83.0 percent, respectively). The overall response rate among women by province is relatively high and ranges from 90.1 percent for Western province to 97.3 for Luapula province. The overall response rate for men ranges from 78.1 for Western province to 96.9 for Luapula province.

Table A.4 Sample implementation: women's sample

Percent distribution of households and eligible women by results of the household and individual interviews, and household, eligible women and overall response rates, according to urban-rural residence and province, Zambia 2001-2002

| | Residence | | Province | | | | | | | | | |
| | Urban | Rural | Central | Copper-belt | Eastern | Luapula | Lusaka | Northern | North-Western | Southern | Western | Total |
Result												
Selected households												
Completed (C)	92.9	86.9	89.0	92.9	90.1	87.9	95.4	88.0	84.1	88.3	82.2	88.5
HH present but no competent												
respondent at home (HP)	1.3	0.7	0.7	0.8	0.7	0.1	1.5	0.2	0.9	0.6	2.5	0.9
Refused (R)	0.5	0.2	0.2	0.6	0.0	0.0	0.5	0.4	0.4	0.1	0.1	0.3
Dwelling not found (DNF)	0.3	0.6	0.4	0.1	0.3	0.8	0.0	0.9	0.0	1.4	1.0	0.5
Household absent (HA)	0.5	2.2	0.8	0.4	1.3	1.1	0.5	2.1	4.1	0.7	3.7	1.8
Dwelling vacant/address not a												
dwelling (DV)	3.3	6.3	6.8	3.9	5.2	4.9	2.0	5.9	5.8	7.7	6.9	5.5
Dwelling destroy (DD)	1.2	3.0	2.0	1.1	2.3	5.2	0.0	2.4	4.7	1.0	3.0	2.5
Other (O)	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.1	0.5	0.1
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100
Number of sampled households	2,167	5,883	848	793	1,070	717	789	1,273	1,074	698	788	8,050
Household response rate (HRR)	97.8	98.3	98.6	98.4	98.9	98.9	97.9	98.3	98.5	97.6	95.7	98.2
Eligible women												
Completed (EWC)	96.3	96.5	97.1	95.1	94.4	98.4	96.9	97.6	96.4	97.8	94.1	96.4
Not at home (EWNH)	2.3	2.3	1.4	3.2	4.3	0.6	1.7	1.2	2.2	1.0	4.9	2.3
Postponed (EWP)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0
Refused (EWR)	0.8	0.4	0.3	0.7	0.5	0.3	0.9	0.4	0.8	0.3	0.1	0.5
Partly completed (EWPC)	0.1	0.1	0.1	0.0	0.1	0.0	0.1	0.1	0.2	0.0	0.1	0.1
Incapacitated (EWI)	0.6	0.7	1.1	0.9	0.6	0.6	0.4	0.7	0.4	0.6	0.4	0.7
Other (EWO)	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.3	0.3	0.1
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Number of women	2,650	5,294	918	987	947	636	925	1,200	914	723	694	7,944
Eligible woman response												
rate (EWRR)	96.3	96.5	97.1	95.1	94.4	98.4	96.9	97.6	96.4	97.8	94.1	96.4
Overall response rate (ORR)	94.1	94.8	95.7	93.6	93.3	97.3	94.8	96.0	94.9	95.5	90.1	94.6

[1] Using the number of households falling into specific response categories, the household response rate (HRR) is calculated as:

$$\frac{100 \times C}{C + HP + R + DNF}$$

[2] Using the number of eligible women falling into specific response categories, the eligible woman response rate (EWRR) is calculated as:

$$\frac{100 \times EWC}{EWC + EWNH + EWR + EWPC + EWI + EWO}$$

Table A.5 Sample implementation: men's sample

Percent distribution of households and eligible men by results of the household and individual interviews, and household, eligible men and overall response rates, according to urban-rural residence and province, Zambia 2001-2002

| | Residence | | Province | | | | | | | | | |
| | Urban | Rural | Central | Copper-belt | Eastern | Luapula | Lusaka | Northern | North-Western | Southern | Western | Total |
Result												
Selected households												
Completed (C)	92.9	87.7	91.2	92.7	91.5	88.7	96.5	88.4	83.1	89.5	81.9	89.1
HH present but no competent												
respondent at home (HP)	1.3	0.7	0.7	1.1	0.6	0.0	0.8	0.5	1.1	0.9	2.3	0.9
Refused (R)	0.1	0.3	0.4	0.4	0.0	0.0	0.0	0.2	0.6	0.0	0.4	0.2
Dwelling not found (DNF)	0.4	0.4	0.4	0.0	0.0	0.4	0.0	0.7	0.0	1.3	1.2	0.4
Household absent (HA)	0.6	2.5	0.7	0.4	0.6	1.3	0.0	3.3	5.6	0.0	3.8	2.0
Dwelling vacant/address not a												
dwelling (DV)	3.6	5.9	5.3	3.4	6.0	5.5	2.7	5.0	5.4	7.4	6.9	5.3
Dwelling destroy (DD)	1.0	2.4	1.4	1.5	1.4	4.2	0.0	1.9	4.2	0.4	2.7	2.0
Other (O)	0.1	0.2	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.4	0.8	0.2
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Number of sampled households	720	1,938	283	261	352	238	259	421	355	229	260	2,658
Household response rate (HRR)	98.1	98.4	98.5	98.4	99.4	99.5	99.2	98.4	98.0	97.6	95.5	98.3
Eligible men												
Completed (EMC)	84.6	90.8	93.9	85.5	85.3	97.3	84.5	89.1	91.5	91.1	81.8	88.7
Not at home (EMNH)	12.0	5.9	4.1	11.3	13.3	0.5	11.3	5.7	6.5	3.1	14.1	7.9
Postponed (EMP)	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.1
Refused (EMR)	2.1	1.1	0.0	2.3	0.0	1.1	2.8	1.9	0.8	2.7	1.6	1.4
Partly completed (EMPC)	0.1	0.1	0.0	0.0	0.0	0.5	0.4	0.0	0.0	0.0	0.0	0.1
Incapacitated (EMI)	0.7	0.9	1.7	1.0	1.0	0.0	0.7	0.8	0.4	0.4	1.0	0.8
Other (EMO)	0.4	1.2	0.3	0.0	0.3	0.5	0.4	2.5	0.8	2.7	0.5	0.9
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Number of men	814	1,604	294	311	300	186	283	367	260	225	192	2,418
Eligible man response rate (EMRR)	84.6	90.8	93.9	85.5	85.3	97.3	84.5	89.1	91.5	91.1	81.8	88.7
Overall response rate (ORR)	83.0	89.4	92.4	84.1	84.8	96.9	83.8	87.7	89.7	88.9	78.1	87.2

[1] Using the number of households falling into specific response categories, the household response rate (HRR) is calculated as:

$$\frac{100 \times C}{C + HP + R + DNF}$$

[2] Using the number of eligible women falling into specific response categories, the eligible woman response rate (EWRR) is calculated as:

$$\frac{100 \times EWC}{EWC + EWNH + EWR + EWPC + EWI + EWO}$$