

APPENDIX C

DATA QUALITY TABLES

APPENDIX C

DATA QUALITY TABLES

The purpose of this Appendix is to provide the data user with a view of the general quality of the ZDHS data. The tables in this appendix refer to possible *non-sampling* errors: digit preference, rounding or heaping on certain ages or dates; omission of events occurring further in the past; deliberate distortion of information by some interviewers in an attempt to lighten their workloads; non-cooperation of the respondent in providing information or refusal to be measured and weighed, etc. A description of the magnitude of such non-sampling errors is provided in the following paragraphs.

The distribution of the de facto household population by single year of age is presented in Table C.1. The data show very little preference to report ages that end in zeros and fives (age “heaping” or digit preference) that is commonly found in countries where ages are not known well. There is some evidence of irregularities in the age distribution. However, it is difficult to find any pattern to these results and they may be due to random errors.

There is some evidence that interviewers “displaced” women age 15 and 49 years and men age 15 and 59 years outside of the eligible age range (15-49) and (15-59) respectively, presumably in order to avoid the need to interview them. For example, the number of women and men age 15 is substantially lower than the number age 16, 14 and 13. For women, the number reported at age 14 (572) is more than that reported at age 13 (485) and 178 more than that at age 15 (394). At the other end of the range, the number of women age 49 is lower than the number age 50. The number of men age 59 is also lower than the number age 60, implying that interviewers assigned an age of 50 (or 51) and 60 (or 61) to women and men respectively, in order to avoid interviewing them. A comparative study of DHS surveys noted some severe displacement out of the eligible age range (Rutstein and Bicego, 1990).

Differential rates of response by age for female and male respondents are also shown in Table C.2. The data do not indicate any strong pattern of response rates by age. The five-year age distribution of respondents is as expected, namely that the percentage of respondents decreases with age.

Information on the completeness of reporting selected important variables is provided in Table C.3. Overall, the percentage of cases with missing information is extraordinarily low. Month of birth was missing for one percent of births that occurred in the 15 years before the survey and remarkably, both month and year were recorded for all the cases. Age at death was missing for an infinitesimal proportion of non-surviving births. Only for the size of child at birth and the anthropometric measurements are there sizeable proportions for which data are missing. While about 10 percent of births in the last 59 months have missing information on size at birth, 8 percent of children under five were not measured. The missing information on the size of the child at birth could be attributed to the child not having been weighed at birth or interviewer negligence. The computer programs have treated the “don’t know” and “not stated” as missing information. The main reason for not measuring children was that the child was not present, either because he/she did not live with the mother or because he/she was not at home. Very few mothers refused to let their children be measured.

According to Table C.4, the information on birth dating is of good quality: both month and year of birth were provided for 98 percent of all births and for 99 percent of surviving children. As expected, information on birth dates is more complete for children who were still living at the time of the survey than for those who had died. Still, both month and year of birth were provided for 97 percent of dead children. Sex ratios are somewhat on the low side; the expected value would be 102 to 103, while those from the ZDHS are often less than 100. This indicates some possible undercounting of male births.

There is very little evidence of transference of births out of 1991 to earlier years. In fact, the ratio of births in 1991 to the average of the two adjoining years is 94 which shows fewer births in 1991 than the average births of the two adjoining years.

Measurement of childhood deaths through retrospective household surveys often suffers from underreporting of deaths, in particular those deaths which occur very early in infancy. If early neonatal deaths are selectively underreported, the result would be an abnormally low ratio of deaths under seven days to all neonatal deaths and an abnormally low ratio of neonatal to infant mortality. Changes in these ratios over time can be examined to detect the hypothesis that underreporting of early infant deaths is more common for births that occurred longer before the survey.

Table C.5 shows the distribution of deaths under one month of age by age at death in days, while Table C.6 shows the distribution of deaths under two years of age by age at death in months. The data suggest that early infant deaths have *not* been severely underreported in the ZDHS, since the percentage of neonatal deaths occurring in the first 6 days (next-to-last row in Table C.5) and the percentage of infant deaths occurring during the neonatal period (next-to-last row in Table C.6) are reasonable. The former proportions increase over time, implying that some early infant deaths were not reported in the earlier periods; however, much, if not all, of this pattern can be attributed to heaping on 7 days at death, which is more severe for the earlier periods.

Table C.1 Household age distribution

Single-year age distribution of the de facto household population by sex (weighted), Zambia 1996

Age	Males		Females		Age	Males		Females	
	Number	Percent	Number	Percent		Number	Percent	Number	Percent
<1	741	4.0	712	3.7	37	157	0.8	132	0.7
1	658	3.5	719	3.7	38	192	1.0	184	0.9
2	654	3.5	678	3.5	39	143	0.8	157	0.8
3	607	3.3	665	3.4	40	155	0.8	148	0.8
4	627	3.4	618	3.2	41	88	0.5	104	0.5
5	560	3.0	590	3.0	42	123	0.7	135	0.7
6	661	3.6	683	3.5	43	118	0.6	114	0.6
7	574	3.1	544	2.8	44	100	0.5	108	0.6
8	589	3.2	544	2.8	45	76	0.4	81	0.4
9	559	3.0	574	3.0	46	104	0.6	99	0.5
10	584	3.1	630	3.2	47	77	0.4	110	0.6
11	483	2.6	489	2.5	48	92	0.5	122	0.6
12	562	3.0	585	3.0	49	89	0.5	108	0.6
13	495	2.7	485	2.5	50	72	0.4	121	0.6
14	577	3.1	572	2.9	51	64	0.3	101	0.5
15	409	2.2	394	2.0	52	70	0.4	114	0.6
16	435	2.3	470	2.4	53	58	0.3	94	0.5
17	400	2.2	400	2.1	54	105	0.6	121	0.6
18	420	2.3	461	2.4	55	56	0.3	72	0.4
19	381	2.1	405	2.1	56	72	0.4	117	0.6
20	422	2.3	455	2.3	57	57	0.3	78	0.4
21	331	1.8	373	1.9	58	59	0.3	109	0.6
22	340	1.8	353	1.8	59	63	0.3	63	0.3
23	310	1.7	387	2.0	60	77	0.4	75	0.4
24	337	1.8	399	2.1	61	70	0.4	58	0.3
25	244	1.3	275	1.4	62	47	0.3	57	0.3
26	301	1.6	308	1.6	63	54	0.3	56	0.3
27	263	1.4	273	1.4	64	97	0.5	66	0.3
28	336	1.8	334	1.7	65	57	0.3	42	0.2
29	185	1.0	185	1.0	66	45	0.2	56	0.3
30	223	1.2	247	1.3	67	36	0.2	41	0.2
31	174	0.9	246	1.3	68	56	0.3	59	0.3
32	287	1.5	274	1.4	69	46	0.2	37	0.2
33	189	1.0	182	0.9	70+	327	1.8	217	1.1
34	199	1.1	207	1.1	Don't know/ missing	1	0.0	1	0.0
35	162	0.9	154	0.8					
36	192	1.0	185	1.0					
					Total	18,575	100.0	19,407	100.0

Note: The de facto population includes all residents and nonresidents who slept in the household the night before the interview.

Table C.2 Age distribution of eligible and interviewed women and men

Percent distribution of the de facto household population of women age 10-54 and men age 10-69, five-year age distribution of interviewed women age 15-49 and men age 15-59, and the percentage of eligible women and men who were interviewed (weighted) by five-year age groups, Zambia 1996

Age	Household population		Persons interviewed		Percent interviewed (weighted)
	Number	Percent	Number	Percent	
WOMEN					
10-14	2,761	NA	NA	NA	NA
15-19	2,129	24.9	2,043	24.7	95.9
20-24	1,967	23.0	1,902	23.0	96.7
25-29	1,375	16.1	1,333	16.1	96.9
30-34	1,156	13.5	1,127	13.6	97.5
35-39	811	9.5	781	9.4	96.2
40-44	609	7.1	583	7.0	95.7
45-49	520	6.1	506	6.1	97.3
50-54	549	NA	NA	NA	NA
15-49	8,568	100.0	8,274	100.0	96.6
MEN					
10-14	721	NA	NA	NA	NA
15-19	522	24.3	479	24.8	91.8
20-24	459	21.3	417	21.6	90.8
25-29	316	14.7	274	14.2	86.6
30-34	267	12.4	235	12.2	87.9
35-39	208	9.7	189	9.8	91.1
40-44	147	6.8	130	6.7	88.1
45-49	95	4.4	84	4.3	87.9
50-54	77	3.6	67	3.5	87.8
55-59	60	2.8	56	2.9	93.5
60-64	107	NA	NA	NA	NA
65+	79	NA	NA	NA	NA
15-59	2,152	100.0	1,932	100.0	89.8

Note: The de facto population includes all residents and nonresidents who slept in the household the night before the interview.

NA = Not applicable

Table C.3. Completeness of reporting

Percentage of observations missing information for selected demographic and health questions (weighted), Zambia 1996

Subject	Reference group	Percentage missing information	Number of cases
Birth date	Births in last 15 years		
Month only		1.02	17,535
Month and year		0.00	17,535
Age at death	Deaths to births in last 15 years	0.19	3,076
Age/date at first union ¹	Ever-married women	0.49	5,989
Respondent's education	All women	0.02	8,021
Child's size at birth	Births in last 35 months	9.53	3,396
Anthropometry²	Living children age 0-35 months		
Height missing		7.94	6,109
Weight missing		7.35	6,109
Height or weight missing		8.05	6,109
Diarrhoea in last 2 weeks	Living children age 0-35 months	3.55	6,109

¹ Both year and age missing

² Child not measured

Table C.4 Births by calendar years

Distribution of births by calendar years for living (L), dead (D), and all (T) children, according to reporting completeness, sex ratio at birth, and ratio of births by calendar year, Zambia 1996

Year	Number of births			Percentage with complete birth date ¹			Sex ratio at birth ²			Calendar ratio ³			Male			Female		
	L	D	T	L	D	T	L	D	T	L	D	T	L	D	T	L	D	T
96	1,242	241	1,483	99.5	97.7	99.2	93.5	93.2	93.5	NA	NA	NA	600	116	716	642	125	766
95	1,172	247	1,419	99.7	97.5	99.3	84.9	110.4	88.9	99.2	96.0	98.6	538	130	668	634	117	751
94	1,121	274	1,395	99.3	98.2	99.1	101.7	124.7	105.8	105.6	123.5	108.7	565	152	717	556	122	678
93	951	197	1,148	99.7	96.0	99.0	89.1	87.2	88.8	85.6	66.6	81.6	448	92	540	503	105	608
92	1,100	317	1,417	99.1	98.3	98.9	92.8	110.3	96.5	117.3	144.5	122.5	530	166	696	571	151	721
91	925	242	1,167	99.2	95.0	98.3	92.7	124.2	98.5	94.4	92.8	94.0	445	134	579	480	108	588
90	861	204	1,065	98.6	97.5	98.4	101.5	77.9	96.5	96.3	94.7	96.0	434	89	523	427	115	542
89	863	189	1,052	98.8	98.8	98.8	97.4	122.2	101.4	102.2	88.8	99.5	426	104	530	437	85	522
88	827	222	1,049	99.0	97.7	98.8	104.3	118.8	107.2	100.1	116.5	103.2	422	120	543	405	101	506
87	789	192	980	98.6	97.3	98.3	94.9	122.7	99.8	NA	NA	NA	384	106	490	405	86	491
92-96	5,586	1,276	6,862	99.5	97.6	99.1	92.3	105.8	94.7	NA	NA	NA	2,681	656	3,337	2,905	620	3,525
87-91	4,265	1,048	5,313	98.9	97.2	98.5	98.0	111.8	100.6	NA	NA	NA	2,111	553	2,664	2,154	495	2,649
82-86	3,299	744	4,043	98.9	95.9	98.3	96.7	101.2	97.5	NA	NA	NA	1,622	374	1,996	1,677	370	2,047
77-81	2,286	567	2,853	98.2	94.8	97.5	103.6	119.8	106.6	NA	NA	NA	1,163	309	1,472	1,123	258	1,381
< 77	1,940	704	2,644	98.3	95.7	97.6	112.9	107.2	111.4	NA	NA	NA	1,029	364	1,393	911	340	1,251
All	17,375	4,339	21,715	98.9	96.5	98.4	98.1	108.4	100.1	NA	NA	NA	8,606	2,257	10,863	8,769	2,083	10,852

NA = Not applicable

¹ Both year and month of birth given

² $(B_m/B_f) \times 100$, where B_m and B_f are the numbers of male and female births, respectively

³ $[2B_x / (B_{x-1} + B_{x+1})] \times 100$, where B_x is the number of births in calendar year x

Table C.5 Reporting of age at death in days

Distribution of reported deaths under one month of age by age at death in days and the percentage of neonatal deaths reported to occur at ages 0-6 days, for five-year periods preceding the survey, Zambia 1996

Age at death (in days)	Number of years preceding the survey				Total 0-19
	0-4	5-9	10-14	15-19	
<1	54	43	42	21	159
1	38	33	15	14	100
2	23	18	10	17	69
3	22	16	17	5	60
4	7	11	7	4	28
5	8	6	5	1	20
6	5	3	1	1	11
7	29	33	30	27	119
8	2	1	3	0	6
9	2	1	1	3	6
10	6	2	0	2	9
11	1	0	0	0	1
12	2	3	1	0	6
13	0	0	1	0	1
14	36	29	21	15	102
15	0	0	1	0	1
16	0	1	0	0	1
17	0	0	1	0	1
18	0	2	0	0	2
19	1	0	0	0	1
20	0	3	1	1	5
21	10	11	11	8	40
23	1	0	0	0	1
25	0	1	0	0	1
26	0	1	0	0	1
28	1	3	0	1	4
30	2	0	5	1	9
Total 0-30	249	219	174	122	765
Percent neonatal ¹ early	62.8	59.1	56.4	51.8	58.5

¹ (0-6 days/0-30 days) * 100

Table C.6 Reporting of age at death in months

Distribution of reported deaths under two years of age by age at death in months and the percentage of infant deaths reported to occur at ages under one month, for five-year periods preceding the survey, Zambia 1996

Age at death (in months)	Number of years preceding the survey				Total 0-19
	0-4	5-9	10-14	15-19	
<1 ^a	251	219	174	123	767
1	52	37	32	21	141
2	39	40	31	18	128
3	65	34	20	21	140
4	51	47	27	12	136
5	33	28	18	14	93
6	38	52	39	29	158
7	40	33	13	12	97
8	45	36	40	17	138
9	43	45	20	16	125
10	24	31	7	11	72
11	33	24	10	4	71
12	41	46	35	21	143
13	12	9	16	3	40
14	16	13	9	6	44
15	17	12	6	5	41
16	16	13	8	4	41
17	8	8	5	2	23
18	22	33	17	19	91
19	11	11	11	2	36
20	14	14	11	4	43
21	9	6	6	3	25
22	8	6	4	0	18
23	3	5	5	1	14
24+	4	3	1	1	10
1 year	50	59	32	38	179
Total 0-11	713	626	430	296	2,065
Percent neonatal ^b	35.2	35.0	40.5	41.4	37.1

^a Includes deaths under 1 month reported in days
^b (Under 1 month/under 1 year) * 100