

APPENDIX A

ESTIMATES OF SAMPLING ERRORS

Two types of errors affect the estimates from a sample survey: (1) nonsampling errors and (2) sampling errors. Nonsampling errors are the result of errors committed during data collection and data processing, such as failure to locate and interview the correct household, misunderstanding of the questions on the part of either the interviewer or the respondent, and data entry errors. Although numerous efforts were made during the implementation of NFHS-2 to minimize nonsampling errors, they are impossible to avoid and difficult to evaluate statistically.

Sampling errors, on the other hand, can be evaluated statistically. The sample of women selected in NFHS-2 is only one of many samples that could have been selected from the same population, using the same design and expected sample size. Each of these samples would yield results that differ somewhat from the results of the actual sample selected. The sampling error is a measure of the variability among all possible samples. Although the degree of variability is not known exactly, it can be estimated from the survey results.

The sampling error is usually measured by the *standard error* for a particular statistic (for example, a mean or percentage), which is the square root of the variance. The standard error can be used to calculate confidence intervals within which the true value for the population can reasonably be assumed to fall. For example, for any given statistic calculated from a sample survey, the value of that statistic will fall within a range, calculated as the value of the statistic plus or minus two times the standard error of that statistic, in 95 percent of all possible samples of identical size and design.

If the sample of women had been selected as a simple random sample, it would have been possible, for many statistics, to use straightforward formulas for calculating sampling errors. However, the NFHS-2 sample is the result of a multi-stage stratified sample design, and it is therefore necessary to use more complex formulas. The computer software used to calculate sampling errors for NFHS-2 is ISSA (the Integrated System for Survey Analysis). The linear Taylor series approximation method for variance estimation is used for estimates of means, proportions, and ratios. The JACKKNIFE repeated replication method is used with ISSA for variance estimation for more complex statistics such as fertility and mortality rates.

The ISSA package treats any percentage or average as a ratio estimate, $r = y/x$, where y represents the sample value for variable y , and x represents the number of cases in the group or subgroup under consideration. The variance of r is computed using the formula given below, with the standard error being the square root of the variance:

$$var(r) = \frac{1 - f}{x^2} \sum_{h=1}^H \left[\frac{m_h}{m_h - 1} \left(\sum_{i=1}^{m_h} z_{hi}^2 - \frac{z_h^2}{m_h} \right) \right]$$

in which

$$z_{hi} = y_{hi} - rx_{hi}$$

$$z_h = y_h - rx_h$$

where

- h = the stratum that varies from 1 to H,
- m_h = the total number of PSUs selected in the h^{th} stratum,
- y_{hi} = the sum of the values of variable y in PSU i in the h^{th} stratum,
- x_{hi} = the sum of the number of cases in PSU i in the h^{th} stratum,
- f = the overall sampling fraction, which is so small that the program ignores it.

In addition to the standard error, ISSA computes the relative standard error, confidence limits for the estimates, and the design effect (DEFT) for each estimate. The design effect is defined as the ratio of the standard error using the given sample design to the standard error that would result if a simple random sample had been used. A DEFT value of 1.0 indicates that the sample design is as efficient as a simple random sample, while a value greater than 1.0 indicates the increase in the sampling error due to the use of a more complex and less statistically efficient design.

Sampling errors for NFHS-2 are calculated for selected variables considered to be of primary interest. The results in this appendix are presented for the state as a whole and for urban and rural areas separately, except for the variable on salt iodization for which the results are shown separately for large cities, small cities, towns, and rural areas. For each variable, the type of statistic (mean, proportion, ratio, or rate) and the base population are given in Table A.1. Table A.2 presents the value of the statistic (R), its standard error (SE), the relative standard error (SE/R), and the 95 percent confidence limits ($R \pm 2SE$) for each variable. In addition, for all variables except the fertility and mortality rates, the table shows the unweighted number of cases (N), the weighted number of cases (WN), the standard error assuming a simple random sample (SER), and the design effect (DEFT).

Table A.1 List of selected variables for sampling errors, Kerala, 1999

Variable	Estimate	Base population
Sex ratio	Ratio	<i>De facto</i> household population
Illiterate	Proportion	<i>De facto</i> household population age 6 and above
Have tuberculosis	Rate	1,000 <i>de jure</i> household population
Salt iodized at 15 ppm or more	Proportion	Households
Illiterate	Proportion	Ever-married women age 15–49
High school complete and above	Proportion	Ever-married women age 15–49
Currently married	Proportion	Ever-married women age 15–49
Number of children ever born	Mean	Currently married women age 15–49
Number of living children	Mean	Currently married women age 15–49
Have ever used any method	Proportion	Currently married women age 15–49
Currently using any method	Proportion	Currently married women age 15–49
Currently using any modern method	Proportion	Currently married women age 15–49
Currently using pills	Proportion	Currently married women age 15–49
Currently using IUD	Proportion	Currently married women age 15–49
Currently using condoms	Proportion	Currently married women age 15–49
Currently using female sterilization	Proportion	Currently married women age 15–49
Currently using male sterilization	Proportion	Currently married women age 15–49
Currently using rhythm/safe period	Proportion	Currently married women age 15–49
Using public source for modern method	Proportion	Current users of modern methods
Do not want any more children	Proportion	Currently married women age 15–49
Want to delay birth at least 2 years	Proportion	Currently married women age 15–49
Ideal number of children	Mean	Ever-married women age 15–49
Ideal number of sons	Mean	Ever-married women age 15–49
Ideal number of daughters	Mean	Ever-married women age 15–49
Visited by health/family planning worker	Proportion	Ever-married women age 15–49
Received no antenatal check-up	Proportion	Births in the past 3 years
Received iron and folic acid tablets or syrup	Proportion	Births in the past 3 years
Received medical assistance during delivery	Proportion	Births in the past 3 years
Received postpartum check-up	Proportion	Noninstitutional births in the past 3 years
Had diarrhoea in the past 2 weeks	Proportion	Children under 3 years
Treated with ORS packets	Proportion	Children under 3 with diarrhoea in past 2 weeks
Taken to a health facility/provider for diarrhoea	Proportion	Children under 3 with diarrhoea in past 2 weeks
Showing a vaccination card	Proportion	Children age 12–23 months
Received BCG vaccination	Proportion	Children age 12–23 months
Received DPT vaccination (3 doses)	Proportion	Children age 12–23 months
Received polio vaccination (3 doses)	Proportion	Children age 12–23 months
Received measles vaccination	Proportion	Children age 12–23 months
Fully vaccinated	Proportion	Children age 12–23 months
Received vitamin A	Proportion	Children age 12–35 months
Had reproductive health problem	Proportion	Currently married women age 15–49
Not involved in any decisionmaking	Proportion	Ever-married women age 15–49
Ever beaten or physically mistreated since age 15	Proportion	Ever-married women age 15–49
Not worked in past 12 months	Proportion	Ever-married women age 15–49
Anaemic women	Proportion	Ever-married women age 15–49
Anaemic children	Proportion	Children age 6–35 months
Fertility rates	Rate	All women, population
Mortality rates	Rate	Births, population

Table A.2 Sampling errors, Kerala, 1999

Variable/ residence	Value (R)	Standard error (SE)	Number of cases		Standard error assuming SRS (SER)	Design effect (DEFT)	Relative standard error (SE/R)	Confidence limits	
			Unweighted (N)	Weighted (WN)				R-2SE	R+2SE
Sex ratio (<i>De facto</i> household population)									
Urban	1078	28.687	2055	1636	26.432	1.085	0.027	1021	1136
Rural	1081	17.896	4908	5336	17.548	1.020	0.017	1045	1117
Total	1080	15.287	6963	6973	14.630	1.045	0.014	1050	1111
Illiterate (<i>De facto</i> household population age 6 and above)									
Urban	0.077	0.009	3871	3081	0.005	1.747	0.115	0.059	0.095
Rural	0.123	0.009	9098	9891	0.004	2.148	0.070	0.106	0.141
Total	0.112	0.007	12969	12972	0.003	2.205	0.064	0.098	0.127
Have tuberculosis (1,000 <i>de jure</i> household population)									
Urban	3.484	1.189	4270	3399	1.058	1.123	0.341	1.107	5.861
Rural	5.807	1.203	10162	11052	1.017	1.184	0.207	3.401	8.214
Total	5.261	0.962	14432	14451	0.796	1.208	0.183	3.338	7.184
Salt iodized at 15 ppm or more (Households)									
Large city	0.608	0.115	102	83	0.049	2.361	0.189	0.378	0.837
Small city	0.602	0.155	83	66	0.054	2.860	0.257	0.293	0.912
Town	0.534	0.045	670	532	0.019	2.349	0.085	0.443	0.624
Rural	0.343	0.028	1979	2153	0.011	2.612	0.081	0.287	0.399
Total	0.393	0.025	2834	2834	0.009	2.674	0.062	0.343	0.442
Illiterate (Ever-married women age 15-49)									
Urban	0.080	0.014	846	667	0.009	1.521	0.177	0.052	0.109
Rural	0.139	0.015	2038	2217	0.008	2.000	0.110	0.109	0.170
Total	0.126	0.012	2884	2884	0.006	2.017	0.099	0.101	0.151
High school complete and above (Ever-married women age 15-49)									
Urban	0.494	0.037	846	667	0.017	2.149	0.075	0.420	0.567
Rural	0.374	0.022	2038	2217	0.011	2.044	0.059	0.330	0.418
Total	0.402	0.019	2884	2884	0.009	2.134	0.049	0.363	0.441
Currently married (Ever-married women age 15-49)									
Urban	0.941	0.008	846	667	0.008	1.018	0.009	0.924	0.957
Rural	0.924	0.006	2038	2217	0.006	1.089	0.007	0.911	0.937
Total	0.928	0.005	2884	2884	0.005	1.096	0.006	0.917	0.938
Number of children ever born (Currently married women age 15-49)									
Urban	1.990	0.050	796	628	0.043	1.173	0.025	1.889	2.091
Rural	2.359	0.053	1882	2048	0.036	1.481	0.022	2.253	2.464
Total	2.272	0.045	2678	2675	0.029	1.568	0.020	2.182	2.362
Number of living children (Currently married women age 15-49)									
Urban	1.911	0.043	796	628	0.039	1.100	0.023	1.824	1.997
Rural	2.249	0.047	1882	2048	0.033	1.430	0.021	2.154	2.344
Total	2.170	0.040	2678	2675	0.027	1.517	0.019	2.089	2.250
Have ever used any method (Currently married women age 15-49)									
Urban	0.794	0.020	796	628	0.014	1.370	0.025	0.755	0.833
Rural	0.763	0.018	1882	2048	0.010	1.861	0.024	0.727	0.800
Total	0.770	0.015	2678	2675	0.008	1.817	0.019	0.741	0.800

Table A.2 Sampling errors, Kerala, 1999 (contd.)

Variable/ residence	Value (R)	Standard error (SE)	Number of cases		Standard error assuming SRS (SER)	Design effect (DEFT)	Relative standard error (SE/R)	Confidence limits	
			Unweighted (N)	Weighted (WN)				R-2SE	R+2SE
Currently using any method (Currently married women age 15–49)									
Urban	0.655	0.022	796	628	0.017	1.286	0.033	0.611	0.698
Rural	0.632	0.019	1882	2048	0.011	1.713	0.030	0.594	0.670
Total	0.637	0.015	2678	2675	0.009	1.664	0.024	0.606	0.668
Currently using any modern method (Currently married women age 15–49)									
Urban	0.574	0.020	796	628	0.018	1.137	0.035	0.534	0.614
Rural	0.557	0.019	1882	2048	0.011	1.700	0.035	0.518	0.596
Total	0.561	0.016	2678	2675	0.010	1.628	0.028	0.530	0.592
Currently using pills (Currently married women age 15–49)									
Urban	0.002	0.002	796	628	0.002	0.986	0.699	0.000	0.006
Rural	0.005	0.001	1882	2048	0.002	0.932	0.308	0.002	0.008
Total	0.004	0.001	2678	2675	0.001	0.962	0.284	0.002	0.007
Currently using IUD (Currently married women age 15–49)									
Urban	0.016	0.005	796	628	0.005	1.116	0.306	0.006	0.026
Rural	0.016	0.003	1882	2048	0.003	1.155	0.210	0.009	0.023
Total	0.016	0.003	2678	2675	0.002	1.154	0.175	0.010	0.022
Currently using condoms (Currently married women age 15–49)									
Urban	0.044	0.008	796	628	0.007	1.058	0.175	0.029	0.059
Rural	0.026	0.004	1882	2048	0.004	1.087	0.152	0.018	0.035
Total	0.031	0.004	2678	2675	0.003	1.083	0.118	0.023	0.038
Currently using female sterilization (Currently married women age 15–49)									
Urban	0.485	0.020	796	628	0.018	1.142	0.042	0.444	0.525
Rural	0.486	0.019	1882	2048	0.012	1.646	0.039	0.448	0.524
Total	0.485	0.015	2678	2675	0.010	1.578	0.031	0.455	0.516
Currently using male sterilization (Currently married women age 15–49)									
Urban	0.027	0.010	796	628	0.006	1.670	0.359	0.007	0.046
Rural	0.024	0.005	1882	2048	0.004	1.379	0.201	0.015	0.034
Total	0.025	0.004	2678	2675	0.003	1.444	0.175	0.016	0.034
Currently using rhythm/safe period (Currently married women age 15–49)									
Urban	0.038	0.009	796	628	0.007	1.264	0.226	0.021	0.055
Rural	0.038	0.006	1882	2048	0.004	1.250	0.144	0.027	0.049
Total	0.038	0.005	2678	2675	0.004	1.259	0.122	0.029	0.047
Using public source for modern method (Current users of modern methods)									
Urban	0.634	0.038	457	360	0.023	1.702	0.061	0.557	0.711
Rural	0.674	0.027	1049	1141	0.014	1.895	0.041	0.619	0.728
Total	0.664	0.023	1506	1501	0.012	1.864	0.034	0.619	0.710
Do not want any more children (Currently married women age 15–49)									
Urban	0.201	0.019	796	628	0.014	1.347	0.095	0.162	0.239
Rural	0.162	0.011	1882	2048	0.008	1.286	0.067	0.140	0.184
Total	0.171	0.010	2678	2675	0.007	1.316	0.056	0.152	0.190
Want to delay birth at least two years (Currently married women age 15–49)									
Urban	0.103	0.011	796	628	0.011	1.033	0.108	0.081	0.126
Rural	0.125	0.008	1882	2048	0.008	1.114	0.068	0.108	0.142
Total	0.120	0.007	2678	2675	0.006	1.121	0.059	0.106	0.134

Table A.2 Sampling errors, Kerala, 1999 (contd.)

Variable/ residence	Value (R)	Standard error (SE)	Number of cases		Standard error assuming SRS (SER)	Design effect (DEFT)	Relative standard error (SE/R)	Confidence limits	
			Unweighted (N)	Weighted (WN)				R-2SE	R+2SE
Ideal number of children (Ever-married women age 15–49)									
Urban	2.279	0.049	765	603	0.032	1.556	0.022	2.180	2.377
Rural	2.600	0.070	1801	1958	0.026	2.706	0.027	2.461	2.740
Total	2.525	0.056	2566	2561	0.021	2.696	0.022	2.412	2.637
Ideal number of sons (Ever-married women age 15–49)									
Urban	0.916	0.038	756	596	0.025	1.515	0.041	0.840	0.992
Rural	0.967	0.029	1781	1936	0.019	1.555	0.030	0.908	1.025
Total	0.955	0.024	2537	2532	0.015	1.581	0.025	0.907	1.003
Ideal number of daughters (Ever-married women age 15–49)									
Urban	0.791	0.028	756	596	0.021	1.331	0.035	0.736	0.847
Rural	0.849	0.023	1781	1936	0.016	1.426	0.027	0.803	0.895
Total	0.836	0.019	2537	2532	0.013	1.448	0.023	0.798	0.873
Visited by health/family planning worker (Ever-married women age 15–49)									
Urban	0.157	0.026	846	667	0.013	2.055	0.164	0.106	0.209
Rural	0.187	0.015	2038	2217	0.009	1.680	0.078	0.158	0.216
Total	0.180	0.013	2884	2884	0.007	1.761	0.070	0.155	0.205
Received no antenatal check-up (Births in the past 3 years)									
Urban	0.000	0.000	170	134	0.000	NC	NC	0.000	0.000
Rural	0.004	0.003	527	572	0.003	0.984	0.699	0.000	0.009
Total	0.003	0.002	697	707	0.002	1.027	0.701	0.000	0.007
Received iron and folic acid tablets or syrup (Births in the past 3 years)									
Urban	0.965	0.013	170	134	0.014	0.933	0.014	0.938	0.991
Rural	0.949	0.011	527	572	0.010	1.123	0.011	0.927	0.970
Total	0.952	0.009	697	707	0.008	1.116	0.010	0.934	0.970
Received medical assistance during delivery (Births in the past 3 years)									
Urban	0.994	0.006	170	134	0.006	1.055	0.006	0.982	1.000
Rural	0.928	0.015	527	572	0.012	1.281	0.016	0.898	0.958
Total	0.940	0.013	697	707	0.009	1.369	0.014	0.915	0.966
Received postpartum check-up (Noninstitutional births in the past 3 years)									
Total	0.274	0.073	40	43	0.071	1.021	0.266	0.128	0.420
Had diarrhoea in the past 2 weeks (Children under 3 years)									
Urban	0.127	0.029	165	130	0.026	1.113	0.228	0.069	0.185
Rural	0.113	0.014	512	556	0.014	0.983	0.121	0.086	0.141
Total	0.116	0.012	677	687	0.012	1.005	0.107	0.091	0.141
Treated with ORS packets (Children under 3 with diarrhoea in past 2 weeks)									
Total	0.479	0.060	79	80	0.057	1.060	0.126	0.359	0.600
Taken to a health facility/provider for diarrhoea (Children under 3 with diarrhoea in past 2 weeks)									
Urban	0.712	0.122	21	17	0.099	1.236	0.171	0.468	0.957
Rural	0.795	0.059	58	63	0.053	1.103	0.074	0.678	0.913
Total	0.778	0.052	79	80	0.047	1.120	0.067	0.673	0.883

Table A.2 Sampling errors, Kerala, 1999 (contd.)

Variable/ residence	Value (R)	Standard error (SE)	Number of cases		Standard error assuming SRS (SER)	Design effect (DEFT)	Relative standard error (SE/R)	Confidence limits	
			Unweighted (N)	Weighted (WN)				R-2SE	R+2SE
Showing a vaccination card (Children age 12–23 months)									
Urban	0.692	0.059	59	47	0.060	0.990	0.086	0.573	0.811
Rural	0.618	0.045	181	197	0.036	1.246	0.073	0.528	0.709
Total	0.632	0.038	240	244	0.031	1.239	0.061	0.556	0.709
Received BCG vaccination (Children age 12–23 months)									
Urban	0.966	0.023	59	47	0.023	0.964	0.023	0.921	1.000
Rural	0.961	0.017	181	197	0.014	1.162	0.017	0.928	0.995
Total	0.962	0.014	240	244	0.012	1.160	0.015	0.934	0.991
Received DPT vaccination (3 doses) (Children age 12–23 months)									
Urban	0.931	0.032	59	47	0.033	0.966	0.034	0.867	0.995
Rural	0.868	0.033	181	197	0.026	1.267	0.038	0.802	0.934
Total	0.880	0.028	240	244	0.021	1.290	0.031	0.825	0.936
Received polio vaccination (3 doses) (Children age 12–23 months)									
Urban	0.931	0.032	59	47	0.033	0.966	0.034	0.867	0.995
Rural	0.873	0.031	181	197	0.026	1.202	0.035	0.811	0.934
Total	0.884	0.026	240	244	0.021	1.229	0.030	0.832	0.936
Received measles vaccination (Children age 12–23 months)									
Urban	0.917	0.036	59	47	0.036	0.989	0.039	0.846	0.988
Rural	0.829	0.047	181	197	0.029	1.629	0.056	0.735	0.923
Total	0.846	0.039	240	244	0.024	1.643	0.046	0.768	0.924
Fully vaccinated (Children age 12–23 months)									
Urban	0.849	0.044	59	47	0.047	0.949	0.052	0.760	0.937
Rural	0.784	0.046	181	197	0.031	1.474	0.059	0.692	0.876
Total	0.797	0.038	240	244	0.026	1.462	0.048	0.720	0.874
Received vitamin A (Children age 12–35 months)									
Urban	0.350	0.053	109	86	0.045	1.170	0.152	0.244	0.456
Rural	0.455	0.042	353	383	0.027	1.551	0.093	0.370	0.540
Total	0.436	0.036	462	469	0.023	1.528	0.082	0.364	0.507
Had reproductive health problem (Currently married women age 15–49)									
Urban	0.374	0.022	796	628	0.017	1.273	0.058	0.331	0.418
Rural	0.440	0.013	1882	2048	0.011	1.142	0.030	0.414	0.466
Total	0.424	0.011	2678	2675	0.010	1.203	0.027	0.401	0.447
Not involved in any decisionmaking (Ever-married women age 15–49)									
Urban	0.068	0.012	846	667	0.009	1.414	0.179	0.044	0.093
Rural	0.073	0.010	2038	2217	0.006	1.658	0.131	0.054	0.092
Total	0.072	0.008	2884	2884	0.005	1.632	0.109	0.056	0.088
Ever beaten or physically mistreated since age 15 (Ever-married women age 15–49)									
Urban	0.087	0.014	846	667	0.010	1.443	0.161	0.059	0.114
Rural	0.107	0.009	2038	2217	0.007	1.315	0.084	0.089	0.125
Total	0.102	0.008	2884	2884	0.006	1.353	0.075	0.087	0.118
Not worked in past 12 months (Ever-married women age 15–49)									
Urban	0.764	0.023	846	667	0.015	1.607	0.031	0.717	0.811
Rural	0.746	0.023	2038	2217	0.010	2.347	0.030	0.700	0.791
Total	0.750	0.018	2884	2884	0.008	2.251	0.024	0.714	0.786

Table A.2 Sampling errors, Kerala, 1999 (contd.)

Variable/ residence	Value (R)	Standard error (SE)	Number of cases		Standard error assuming SRS (SER)	Design effect (DEFT)	Relative standard error (SE/R)	Confidence limits	
			Unweighted (N)	Weighted (WN)				R-2SE	R+2SE
Anaemic women (Ever-married women age 15–49)									
Urban	0.204	0.020	796	629	0.014	1.384	0.097	0.164	0.243
Rural	0.234	0.013	1925	2092	0.010	1.385	0.057	0.208	0.261
Total	0.227	0.011	2721	2721	0.008	1.401	0.050	0.205	0.250
Anaemic children (Children age 6–35 months)									
Urban	0.468	0.047	131	104	0.044	1.076	0.101	0.374	0.562
Rural	0.432	0.033	384	413	0.025	1.290	0.076	0.366	0.497
Total	0.439	0.028	515	517	0.022	1.264	0.063	0.384	0.494

Table A.2 Sampling errors, Kerala, 1999 (contd.)					
Variable/ residence	Value (R)	Standard error (SE)	Relative standard error (SE/R)	Confidence limits	
				R-2SE	R+2SE
Total fertility rate (Women age 15–49)					
Urban	1.510	0.088	0.058	1.333	1.686
Rural	2.071	0.090	0.043	1.891	2.251
Total	1.961	0.073	0.037	1.816	2.106
Age-specific fertility rate (Women age 15–19)					
Urban	0.013	0.009	0.640	0.000	0.031
Rural	0.041	0.007	0.176	0.026	0.055
Total	0.039	0.006	0.149	0.028	0.051
Age-specific fertility rate (Women age 20–24)					
Urban	0.128	0.012	0.093	0.105	0.152
Rural	0.179	0.009	0.052	0.160	0.197
Total	0.166	0.008	0.046	0.151	0.181
Age-specific fertility rate (Women age 25–29)					
Urban	0.097	0.012	0.120	0.074	0.120
Rural	0.137	0.010	0.074	0.117	0.157
Total	0.128	0.008	0.064	0.112	0.144
Age-specific fertility rate (Women age 30–34)					
Urban	0.042	0.009	0.206	0.024	0.059
Rural	0.039	0.006	0.143	0.028	0.050
Total	0.040	0.005	0.118	0.030	0.049
Age-specific fertility rate (Women age 35–39)					
Urban	0.022	0.010	0.453	0.002	0.041
Rural	0.014	0.004	0.316	0.005	0.023
Total	0.016	0.004	0.258	0.008	0.024
Age-specific fertility rate (Women age 40–44)					
Urban	0.000	0.000	NC	0.000	0.000
Rural	0.004	0.002	0.485	0.000	0.009
Total	0.003	0.002	0.487	0.000	0.006

Table A.2 Sampling errors, Kerala, 1999 (contd.)					
Variable/ residence	Value (R)	Standard error (SE)	Relative standard error (SE/R)	Confidence limits	
				R-2SE	R+2SE
Neonatal mortality (5-year period preceding survey)					
Urban	12.478	6.054	0.485	0.371	24.585
Rural	14.156	5.243	0.370	3.670	24.642
Total	13.806	4.313	0.312	5.181	22.431
Infant mortality ${}_1q_0$ (5-year period preceding survey)					
Urban	15.544	6.660	0.428	2.224	28.863
Rural	16.514	5.422	0.328	5.669	27.358
Total	16.316	4.483	0.275	7.350	25.283
Child mortality ${}_4q_1$ (5-year period preceding survey)					
Urban	3.065	3.107	1.014	0.000	9.280
Rural	2.403	1.465	0.610	0.000	5.333
Total	2.563	1.326	0.517	0.000	5.215
Under-five mortality ${}_5q_0$ (5-year period preceding survey)					
Urban	18.561	7.055	0.380	4.451	32.671
Rural	18.876	5.760	0.305	7.356	30.397
Total	18.838	4.759	0.253	9.321	28.355
Crude death rate (Based on Household Questionnaire)					
Urban	5.681	0.758	0.133	4.166	7.196
Rural	6.012	0.606	0.101	4.801	7.223
Total	5.934	0.501	0.084	4.932	6.936
Crude birth rate (Based on women's birth history)					
Urban	14.814	0.968	0.065	12.879	16.750
Rural	19.744	0.964	0.049	17.816	21.671
Total	18.834	0.766	0.041	17.302	20.367
NC: Not calculated because denominator is 0.000 SRS: Simple random sample					

APPENDIX B

DATA QUALITY TABLES

The purpose of this appendix is to provide the data user with an overview of the general quality of the NFHS-2 data. Whereas Appendix A is concerned with sampling errors and their effects on the survey results, the tables in this appendix refer to possible *nonsampling* errors: for example, 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 workload; noncooperation of the respondent in providing information; or refusal to have children measured for height and weight or tested for anaemia. A description of the likely magnitude of such nonsampling errors is provided in this appendix.

The distribution of the *de facto* household population by single years of age and sex is presented in Table B.1. In many (but not all) cases, the respondent for the household questionnaire was the head of the household. It is well documented that ages are poorly reported in most parts of India. Ages are of little relevance for much of the rural population in particular, and no amount of probing will ensure that ages are properly recorded. In interviewer training for NFHS-2, a great deal of emphasis was placed on obtaining as accurate information as possible on ages and dates of events. Nevertheless, it is clear that age reporting in NFHS-2 shares the same problems inherent in all Indian censuses and surveys. Heaping on ages ending in 0, 2, 5, and 8 is considerable and is particularly severe in the older age groups. Another measure of the quality of the NFHS-2 age data is the percentage of persons whose ages were recorded as not known or missing. In Kerala, information on age was missing for only 3 persons out of 14,506 persons who stayed in the sample households the night before the interview.

Table B.2 examines the possibility that some eligible women (that is, ever-married women age 15–49) were not properly identified in NFHS-2. In some surveys, interviewers may try to reduce their workload by pushing women out of the eligible age range or recording ever-married women as never married so that they will not have to be interviewed. If such practices were being followed to a noticeable extent, Table B.2 would normally show (1) a shortage of ever-married women in the 45–49 age group and an excess in the 50–54 age group or (2) an unusually low proportion of ever-married women by age. Neither of these patterns is evident in the NFHS-2 data. It can, therefore, be concluded that there was no concerted effort to misidentify eligible women in NFHS-2 in Kerala.

One traditional measure of the quality of data is the extent to which information is missing on key variables. Although completeness of responses does not necessarily indicate that the results are accurate, the existence of missing information for a large number of cases would suggest that data collection was not carried out with sufficient care. In NFHS-2 in Kerala, woman's education and the age at death for children born in the past 15 years who died are recorded in every case. Almost complete information is available for age at first marriage and prevalence of diarrhoea in the two weeks preceding the survey (Table B.3). Missing information is somewhat higher for the month of birth, but the year of birth is recorded in every case in which the month is missing. Data on height or weight for children is missing in 12 percent of the cases. Many children could not be measured because they were not at home or they were ill at the time of the survey. In some cases when the child was at home, either the child refused to be measured

Table B.1 Household age distribution

Single-year age distribution of *de facto* household population by sex (weighted), Kerala, 1999

Age	Male		Female		Age	Male		Female	
	Number	Percent	Number	Percent		Number	Percent	Number	Percent
< 1	122	1.8	138	1.8	38	130	1.9	135	1.8
1	132	1.9	120	1.6	39	68	1.0	70	0.9
2	124	1.8	115	1.5	40	137	2.0	155	2.1
3	123	1.8	121	1.6	41	40	0.6	58	0.8
4	148	2.1	145	1.9	42	98	1.4	105	1.4
5	136	2.0	108	1.4	43	61	0.9	77	1.0
6	124	1.8	128	1.7	44	61	0.9	65	0.9
7	133	1.9	122	1.6	45	142	2.0	127	1.7
8	157	2.3	103	1.4	46	60	0.9	72	1.0
9	119	1.7	106	1.4	47	53	0.8	68	0.9
10	139	2.0	147	2.0	48	98	1.4	103	1.4
11	129	1.9	114	1.5	49	57	0.8	49	0.7
12	161	2.3	154	2.0	50	116	1.7	91	1.2
13	156	2.2	161	2.1	51	34	0.5	69	0.9
14	148	2.1	169	2.2	52	86	1.2	90	1.2
15	134	1.9	164	2.2	53	42	0.6	44	0.6
16	140	2.0	144	1.9	54	33	0.5	31	0.4
17	167	2.4	153	2.0	55	101	1.4	102	1.4
18	198	2.8	179	2.4	56	39	0.6	30	0.4
19	144	2.1	138	1.8	57	35	0.5	24	0.3
20	140	2.0	172	2.3	58	51	0.7	59	0.8
21	106	1.5	114	1.5	59	17	0.2	21	0.3
22	138	2.0	139	1.8	60	98	1.4	129	1.7
23	121	1.7	143	1.9	61	21	0.3	18	0.2
24	142	2.0	145	1.9	62	33	0.5	50	0.7
25	112	1.6	166	2.2	63	27	0.4	27	0.4
26	106	1.5	146	1.9	64	20	0.3	24	0.3
27	111	1.6	129	1.7	65	100	1.4	108	1.4
28	138	2.0	144	1.9	66	14	0.2	22	0.3
29	68	1.0	127	1.7	67	27	0.4	20	0.3
30	132	1.9	173	2.3	68	28	0.4	43	0.6
31	57	0.8	65	0.9	69	9	0.1	13	0.2
32	97	1.4	136	1.8	70+	288	4.1	340	4.5
33	75	1.1	104	1.4	Don't				
34	78	1.1	90	1.2	know/				
35	151	2.2	166	2.2	missing	1	0.0	2	0.0
36	71	1.0	117	1.6					
37	71	1.0	86	1.1	Total	6,973	100.0	7,533	100.0

Note: The *de facto* population includes both usual residents and visitors who stayed in the household the night before the interview.

or the mother refused to allow the child to be measured. Data on women's haemoglobin levels are available for 94 percent of respondents and data on children's haemoglobin level are available for 88 percent of children. Before undertaking haemoglobin measurements, a separate 'informed consent' statement was read to the respondent explaining that participation in the haemoglobin testing was completely voluntary. At this point, some women declined to take part in the anaemia testing and/or to have their children participate. Given the voluntary nature of the measurements, the high response rates for height and weight and anaemia testing are impressive.

Another measure of data quality is the completeness and accuracy of information on births. Table B.4 examines the distribution of births by calendar year to identify any unusual patterns that may indicate that births have been omitted or that the ages of children have been displaced. Overall, 86 percent of living children listed in the birth history had complete birth dates recorded, as did 60 percent of children who had died. The completeness of data on birth

Table B.2 Age distribution of eligible and interviewed women					
Age distribution of the <i>de facto</i> household population of women age 10–54 and of interviewed women age 15–49, and percentage of eligible women who were interviewed (weighted), Kerala, 1999					
Age	All women	Ever-married women	Interviewed women		Percent interviewed
			Number	Percent	
10–14	745	0	NA	NA	NA
15–19	778	110	99	3.4	90.4
20–24	713	432	386	13.4	89.4
25–29	712	633	582	20.2	92.1
30–34	567	531	504	17.5	94.9
35–39	575	552	524	18.2	94.9
40–44	460	446	419	14.5	93.7
45–49	419	408	374	13.0	91.6
50–54	325	314	NA	NA	NA
15–49	4,223	3,113	2,889	100.0	92.8

Note: The *de facto* population includes both usual residents and visitors who stayed in the household the night before the interview. For all columns, the age distribution is taken from ages reported in the Household Questionnaire. The total number of interviewed women in this table differs from the total number in earlier tables because this table uses household weights rather than women's weights for the calculations.
NA: Not applicable

Table B.3 Completeness of reporting			
Percentage of observations with missing information for selected demographic and health indicators (weighted), Kerala, 1999			
Indicator	Reference group	Percentage missing information	Number of cases
Birth date	Births in past 15 years		
Month only		9.41	3,774
Month and year		0.00	3,774
Age at death	Deaths to births in past 15 years	0.00	127
Age at first marriage	Ever-married women age 15–49	0.04	2,884
Woman's education	Ever-married women age 15–49	0.00	2,884
Anthropometry	Living children age 0–35 months		
Height		12.16	699
Weight		7.76	699
Height or weight		12.16	699
Woman's haemoglobin level	Ever-married women age 15–49	5.67	2,884
Child's haemoglobin level	Living children age 6–35 months	11.72	589
Diarrhoea in past 2 weeks	Living children age 1–35 months	0.16	687

Table B.4 Births by calendar year

Number of births, percent with complete birth date, sex ratio at birth, and calendar year ratio for children still alive at the time of the survey (L), children who died by the time of the survey (D), and total children (T), by calendar year (weighted), Kerala, 1999

Calendar year	Number of births			Percent with complete birth date ¹			Sex ratio at birth ²			Calendar year ratio ³		
	L	D	T	L	D	T	L	D	T	L	D	T
1999	85	0	85	100.0	NC	100.0	1,135	NC	1,135	NA	NA	NA
1998	238	4	242	100.0	74.6	99.5	1,074	329	1,053	NC	NC	NC
1997	259	5	264	98.8	100.0	98.8	805	626	801	116.5	130.0	116.7
1996	208	3	211	97.9	100.0	97.9	950	572	943	83.7	61.0	83.3
1995	237	5	242	98.3	78.1	97.9	1,016	281	992	105.5	145.0	106.1
1994	241	4	245	96.4	100.0	96.5	775	233	762	100.7	88.3	100.5
1993	242	4	246	94.3	73.1	93.9	1,007	837	1,004	101.0	49.3	99.2
1992	239	13	251	92.2	91.2	92.1	994	526	964	99.9	296.3	103.3
1991	235	4	240	94.7	49.1	93.9	856	982	858	99.3	37.4	96.4
1990	236	11	246	91.6	59.7	90.2	773	394	752	99.2	178.6	101.1
1989	240	8	248	89.2	58.3	88.2	1,018	617	1,003	98.3	58.1	96.2
1988	253	16	268	86.1	57.9	84.4	977	1,539	1,003	103.0	175.8	105.5
1993-97	1,187	21	1,208	97.1	89.4	97.0	902	465	892	NA	NA	NA
1988-92	1,202	51	1,253	90.7	65.7	89.7	920	757	913	NA	NA	NA
1983-87	1,288	73	1,361	82.5	59.5	81.3	1,082	974	1,076	NA	NA	NA
1978-82	1,193	65	1,258	80.7	60.6	79.7	890	1,041	898	NA	NA	NA
1977 or earlier	1,015	96	1,111	76.4	48.8	74.0	957	635	924	NA	NA	NA
All	6,208	310	6,518	86.4	59.7	85.2	956	781	947	NA	NA	NA

NA: Not applicable

NC: Not calculated because full-year data were not collected for 1999 or because the percentage with complete birth dates and the sex ratio at birth for children who have died are undefined.

¹Both year and month of birth given

² $(B_f/B_m) \times 1000$, where B_f and B_m are the numbers of female and male 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 B.5 Reporting of age at death in days				
Distribution of reported deaths under 1 month of age by age at death in days and percentage of neonatal deaths reported to occur at age 0–6 days, for births occurring during five-year periods preceding the survey (weighted), Kerala, 1999				
Age at death (days)	Years preceding survey			
	0–4	5–9	10–14	0–14
< 1	4	7	5	17
1	3	10	7	20
2	0	1	3	4
3	3	0	7	10
4	2	1	2	5
5	0	0	1	1
6	1	0	1	2
7	1	0	3	4
8	0	0	1	1
9	0	0	1	1
10	0	1	1	2
11	0	0	0	0
12	0	0	0	0
13	0	0	0	0
14	1	0	1	2
15	0	0	1	1
16	0	0	0	0
17	0	0	0	0
18	0	0	0	0
19	0	0	0	0
20	0	0	0	0
21	0	0	0	0
22	0	0	1	1
23	0	0	1	1
24	0	0	1	1
25	0	0	0	0
26	0	0	0	0
27	0	0	0	0
28	0	0	0	0
29	0	0	0	0
30	0	0	0	0
0–30	15	21	38	74
Percent early neonatal ¹	87.5	95.1	70.0	80.7

¹Deaths during the first 6 days divided by deaths during the first 30 days

dates for both surviving and nonsurviving children is satisfactory overall and good in recent years.

Many surveys that include both demographic information and health information for children below a specified age have been subject to a substantial amount of age displacement. In particular, there is often a tendency for interviewers to ‘age’ children out of the eligible period for the health questions. This problem was well known before NFHS-2 began; therefore, interviewer training stressed this issue to try to reduce the extent of biases due to age displacement. The annual data on the number of births in Table B.4 can be examined to see if there is an abnormally large decline in the number of births after January 1996, the cutoff point for the health questions and measurements made on young children in the survey. It is typical for the annual number of births to fluctuate somewhat, so small annual fluctuations are to be expected. In Table B.4, a sharp drop in the annual number of births in the calendar year 1996 is observed, but there is no evidence of substantial avoidance of births after 1996.

Table B.5 presents information on the reporting of age at death in days. Results from the table suggest that early infant deaths have not been seriously underreported in Kerala, because the ratios of deaths under seven days to all neonatal deaths are consistently high (a ratio of less than 25 percent is often used as a guideline to indicate underreporting of early neonatal deaths). The ratios are 88 for the period 0–4 years, 95 for the period 5–9 years, and 70 for the period 10–14 years preceding the survey. Table B.6 shows the percentage of infant deaths that occurred during the neonatal period. These percentages are also quite high, especially in the most recent period, suggesting that there is no major omission of early deaths.

One problem that is inherent in most retrospective surveys is heaping of the age at death on certain digits, e.g., 6, 2, and 8 months. Misreporting of age at death will bias estimates of the age pattern of mortality if the net result of misreporting is the transference of deaths between age segments for which the rates are calculated. For example, an overestimate of child mortality relative to infant mortality may result if children dying during the first year of life are reported as having died at age one year or older. Thus, heaping at 12 months can bias the mortality estimates because a certain fraction of these deaths, which are reported to have occurred after infancy may have actually occurred during infancy (that is, at ages 0–11 months). In this case, heaping would bias the infant mortality rate downward and the child mortality rate upward.

Table B.6 Reporting of age at death in months				
Distribution of reported deaths under two years of age by age at death in months and percentage of infant deaths reported to occur at age under one month, for births occurring during five-year periods preceding the survey (weighted), Kerala, 1999				
Age at death (months)	Years preceding survey			
	0–4	5–9	10–14	0–14
< 1	15	21	38	74
1	0	0	3	3
2	1	0	3	4
3	1	1	2	4
4	0	1	1	2
5	0	1	2	3
6	0	1	2	3
7	0	1	2	3
8	0	3	1	4
9	0	2	0	2
10	0	0	0	0
11	0	0	0	0
12	0	1	0	1
13	1	0	1	2
14	0	0	2	2
15	0	1	0	1
16	0	0	0	0
17	0	0	0	0
18	0	1	1	2
19	0	0	0	0
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	0	0	0	0
0–11	17	32	55	103
Percent neonatal ¹	88.8	66.4	69.2	71.5

¹Deaths during the first month divided by deaths during the first year

An examination of the distribution of deaths under one year, shows some heaping on 1, 3, 4, and 7 days at death (Table B.5), but an examination of the distribution of deaths under age two years during the 15 years before the survey by month of death (Table B.6) indicates absolutely no heaping of deaths at any month of age. The strong emphasis on this problem during the training of interviewers for the NFHS-2 fieldwork appears to have paid off to some extent. This brief check on internal consistency of the Kerala NFHS-2 data suggests that there is little underreporting of deaths during the time periods for which mortality rates are estimated and that other problems associated with retrospective data collection have also been largely minimized.