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:

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

in which

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

$$z_h = y_h - rx_h$$

where

h represents the stratum that varies from 1 to H,
m_h is the total number of PSUs selected in the hth stratum,
y_{hi} is the sum of the values of variable y in PSU i in the hth stratum,
x_{hi} is the sum of number of cases in PSU i in the hth stratum,
f is 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±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, Bihar, 1998–99

Variable	Estimate	Base population
Sex ratio	Ratio	*De facto* household population
Illiterate	Proportion	*De facto* household population age 6 and above
Have tuberculosis	Rate	1,000 *de jure* 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 past 3 years
Received iron and folic acid tablets or syrup	Proportion	Births in past 3 years
Received medical assistance during delivery	Proportion	Births in past 3 years
Received postpartum check-up	Proportion	Noninstitutional births in past 3 years
Had diarrhoea in past 2 weeks	Proportion	Children under age 3 years
Treated with ORS packets	Proportion	Children under age 3 with diarrhoea in past 2 weeks
Taken to a health facility/provider for diarrhoea	Proportion	Children under age 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, Bihar, 1998–99

Variable/ residence	Value (R)	Standard error (SE)	Number of cases: Unweighted (N)	Number of cases: Weighted (WN)	Standard error assuming SRS (SER)	Design effect (DEFT)	Relative standard error (SE/R)	Confidence limits: R-2SE	Confidence limits: R+2SE
Sex ratio (*De facto* household population)									
Urban	932	28.009	2,146	2,201	24.394	1.148	0.030	876	988
Rural	955	8.305	17,181	17,112	8.392	0.990	0.009	939	972
Total	953	8.060	19,327	19,314	7.937	1.016	0.008	937	969
Illiterate (*De facto* household population age 6 and above)									
Urban	0.261	0.025	3,615	3,707	0.012	2.040	0.097	0.210	0.311
Rural	0.540	0.010	27,941	27,827	0.005	2.064	0.019	0.520	0.560
Total	0.507	0.011	31,556	31,534	0.005	2.326	0.022	0.485	0.529
Have tuberculosis (1,000 *de jure* household population)									
Urban	6.290	1.504	4,223	4,330	1.443	1.043	0.239	3.282	9.299
Rural	10.347	0.763	34,478	34,339	0.627	1.216	0.074	8.821	11.874
Total	9.893	0.703	38,701	38,669	0.581	1.211	0.071	8.487	11.299
Salt iodized at 15 ppm or more (Households)									
Large city	0.773	0.121	75	77	0.049	2.483	0.156	0.532	1.000
Small city	0.837	0.054	159	163	0.029	1.838	0.065	0.729	0.945
Town	0.756	0.034	467	480	0.020	1.729	0.045	0.687	0.825
Rural	0.430	0.015	5,644	5,625	0.007	2.317	0.036	0.399	0.460
Total	0.469	0.016	6,345	6,345	0.006	2.478	0.033	0.438	0.500
Illiterate (Ever-married women age 15–49)									
Urban	0.462	0.037	687	718	0.019	1.959	0.081	0.388	0.537
Rural	0.801	0.011	6,337	6,306	0.005	2.153	0.013	0.779	0.823
Total	0.766	0.012	7,024	7,024	0.005	2.372	0.016	0.742	0.790
High school complete and above (Ever-married women age 15–49)									
Urban	0.278	0.035	687	718	0.017	2.023	0.124	0.209	0.347
Rural	0.063	0.007	6,337	6,306	0.003	2.179	0.106	0.049	0.076
Total	0.085	0.008	7,024	7,024	0.003	2.376	0.093	0.069	0.100
Currently married (Ever-married women age 15–49)									
Urban	0.942	0.007	687	718	0.009	0.828	0.008	0.928	0.957
Rural	0.949	0.004	6,337	6,306	0.003	1.361	0.004	0.941	0.956
Total	0.948	0.003	7,024	7,024	0.003	1.318	0.004	0.941	0.955
Number of children ever born (Currently married women age 15–49)									
Urban	3.303	0.090	647	677	0.089	1.009	0.027	3.122	3.483
Rural	3.284	0.036	6,015	5,984	0.031	1.180	0.011	3.212	3.357
Total	3.286	0.034	6,662	6,661	0.029	1.163	0.010	3.218	3.353
Number of living children (Currently married women age 15–49)									
Urban	3.046	0.080	647	677	0.081	0.991	0.026	2.885	3.207
Rural	2.850	0.029	6,015	5,984	0.026	1.121	0.010	2.791	2.909
Total	2.870	0.028	6,662	6,661	0.025	1.116	0.010	2.814	2.925
Have ever used any method (Currently married women age 15–49)									
Urban	0.448	0.030	647	677	0.020	1.551	0.068	0.387	0.509
Rural	0.257	0.010	6,015	5,984	0.006	1.813	0.040	0.236	0.277
Total	0.276	0.010	6,662	6,661	0.005	1.878	0.037	0.256	0.297

Table A.2 Sampling errors, Bihar, 1998–99 (Contd.)

Variable/ residence	Value (R)	Standard error (SE)	Number of cases: Unweighted (N)	Number of cases: Weighted (WN)	Standard error assuming SRS (SER)	Design effect (DEFT)	Relative standard error (SE/R)	Confidence limits: R-2SE	Confidence limits: R+2SE
Currently using any method (Currently married women age 15–49)									
Urban	0.389	0.028	647	677	0.019	1.455	0.072	0.333	0.445
Rural	0.229	0.010	6,015	5,984	0.005	1.763	0.042	0.210	0.248
Total	0.245	0.010	6,662	6,661	0.005	1.804	0.039	0.226	0.264
Currently using any modern method (Currently married women age 15–49)									
Urban	0.354	0.025	647	677	0.019	1.356	0.072	0.303	0.405
Rural	0.209	0.009	6,015	5,984	0.005	1.732	0.043	0.191	0.227
Total	0.224	0.009	6,662	6,661	0.005	1.756	0.040	0.206	0.242
Currently using pills (Currently married women age 15–49)									
Urban	0.029	0.006	647	677	0.007	0.940	0.215	0.016	0.041
Rural	0.008	0.001	6,015	5,984	0.001	1.026	0.149	0.005	0.010
Total	0.010	0.001	6,662	6,661	0.001	1.052	0.129	0.007	0.012
Currently using IUD (Currently married women age 15–49)									
Urban	0.012	0.004	647	677	0.004	0.961	0.341	0.004	0.020
Rural	0.004	0.001	6,015	5,984	0.001	1.292	0.248	0.002	0.007
Total	0.005	0.001	6,662	6,661	0.001	1.233	0.207	0.003	0.007
Currently using condoms (Currently married women age 15–49)									
Urban	0.030	0.009	647	677	0.007	1.284	0.288	0.013	0.047
Rural	0.004	0.001	6,015	5,984	0.001	1.132	0.219	0.002	0.006
Total	0.007	0.001	6,662	6,661	0.001	1.265	0.185	0.004	0.010
Currently using female sterilization (Currently married women age 15–49)									
Urban	0.266	0.020	647	677	0.017	1.175	0.077	0.225	0.307
Rural	0.183	0.008	6,015	5,984	0.005	1.679	0.046	0.167	0.200
Total	0.192	0.008	6,662	6,661	0.005	1.650	0.041	0.176	0.208
Currently using male sterilization (Currently married women age 15–49)									
Urban	0.017	0.007	647	677	0.005	1.325	0.400	0.003	0.030
Rural	0.009	0.001	6,015	5,984	0.001	1.099	0.149	0.006	0.012
Total	0.010	0.001	6,662	6,661	0.001	1.147	0.142	0.007	0.012
Currently using rhythm/safe period (Currently married women age 15–49)									
Urban	0.015	0.004	647	677	0.005	0.913	0.290	0.006	0.024
Rural	0.008	0.001	6,015	5,984	0.001	1.165	0.167	0.005	0.011
Total	0.009	0.001	6,662	6,661	0.001	1.131	0.148	0.006	0.011
Using public source for modern method (Current users of modern methods)									
Urban	0.560	0.030	229	239	0.033	0.898	0.053	0.501	0.619
Rural	0.809	0.013	1,260	1,250	0.011	1.172	0.016	0.783	0.835
Total	0.769	0.013	1,489	1,490	0.011	1.229	0.017	0.742	0.796
Do not want any more children (Currently married women age 15–49)									
Urban	0.311	0.026	647	677	0.018	1.428	0.084	0.259	0.363
Rural	0.301	0.007	6,015	5,984	0.006	1.242	0.024	0.287	0.316
Total	0.302	0.007	6,662	6,661	0.006	1.260	0.023	0.288	0.316
Want to delay birth at least 2 years (Currently married women age 15–49)									
Urban	0.159	0.015	647	677	0.014	1.048	0.095	0.129	0.189
Rural	0.158	0.006	6,015	5,984	0.005	1.173	0.035	0.147	0.169
Total	0.158	0.005	6,662	6,661	0.004	1.158	0.033	0.148	0.169

Table A.2 Sampling errors, Bihar, 1998–99 (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	2.802	0.074	662	693	0.041	1.799	0.026	2.655	2.949
Rural	3.312	0.028	5,883	5,856	0.015	1.892	0.008	3.256	3.367
Total	3.258	0.028	6,545	6,549	0.014	1.983	0.008	3.202	3.313
Ideal number of sons (Ever-married women age 15–49)									
Urban	1.541	0.057	662	693	0.033	1.730	0.037	1.428	1.655
Rural	1.925	0.021	5,878	5,851	0.011	1.935	0.011	1.883	1.966
Total	1.884	0.021	6,540	6,544	0.010	2.014	0.011	1.842	1.926
Ideal number of daughters (Ever-married women age 15–49)									
Urban	1.107	0.034	662	693	0.022	1.516	0.031	1.040	1.175
Rural	1.294	0.012	5,878	5,851	0.008	1.480	0.009	1.270	1.319
Total	1.274	0.012	6,540	6,544	0.008	1.541	0.009	1.250	1.299
Visited by health/family planning worker (Ever-married women age 15–49)									
Urban	0.018	0.005	687	718	0.005	1.072	0.304	0.007	0.029
Rural	0.025	0.003	6,337	6,306	0.002	1.553	0.123	0.019	0.031
Total	0.024	0.003	7,024	7,024	0.002	1.520	0.116	0.018	0.029
Received no antenatal check-up (Births in past 3 years)									
Urban	0.306	0.048	245	258	0.032	1.493	0.157	0.209	0.402
Rural	0.656	0.018	2,703	2,689	0.010	1.829	0.027	0.620	0.692
Total	0.625	0.018	2,948	2,947	0.010	1.847	0.028	0.590	0.661
Received iron and folic acid tablets or syrup (Births in past 3 years)									
Urban	0.461	0.049	245	258	0.032	1.527	0.106	0.364	0.558
Rural	0.220	0.012	2,703	2,689	0.008	1.461	0.053	0.197	0.243
Total	0.241	0.012	2,948	2,947	0.008	1.540	0.050	0.217	0.265
Received medical assistance during delivery (Births in past 3 years)									
Urban	0.519	0.053	245	258	0.035	1.535	0.103	0.412	0.625
Rural	0.206	0.015	2,703	2,689	0.008	1.791	0.072	0.177	0.236
Total	0.234	0.015	2,948	2,947	0.008	1.788	0.064	0.204	0.264
Received postpartum check-up (Noninstitutional births in past 3 years)									
Urban	0.100	0.022	143	150	0.025	0.893	0.225	0.055	0.145
Rural	0.101	0.007	2,347	2,330	0.006	1.191	0.074	0.086	0.115
Total	0.100	0.007	2,490	2,480	0.006	1.173	0.070	0.086	0.115
Had diarrhoea in past 2 weeks (Children under age 3 years)									
Urban	0.156	0.024	232	244	0.024	1.014	0.155	0.108	0.205
Rural	0.179	0.009	2,463	2,448	0.008	1.162	0.050	0.161	0.197
Total	0.177	0.008	2,695	2,692	0.007	1.147	0.048	0.160	0.194
Treated with ORS packets (Children under age 3 with diarrhoea in past 2 weeks)									
Urban	0.281	0.084	36	38	0.075	1.126	0.299	0.113	0.449
Rural	0.143	0.017	440	439	0.017	1.014	0.119	0.109	0.177
Total	0.154	0.017	476	477	0.017	1.022	0.111	0.120	0.188
Taken to a health facility/provider for diarrhoea (Children under age 3 with diarrhoea in past 2 weeks)									
Urban	0.389	0.060	36	38	0.081	0.744	0.155	0.269	0.510
Rural	0.513	0.025	440	439	0.024	1.038	0.049	0.462	0.563
Total	0.503	0.024	476	477	0.023	1.018	0.047	0.455	0.550

Table A.2 Sampling errors, Bihar, 1998–99 (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	0.203	0.052	76	80	0.046	1.123	0.255	0.100	0.306
Rural	0.171	0.014	806	802	0.013	1.043	0.081	0.143	0.199
Total	0.174	0.013	882	882	0.013	1.047	0.077	0.147	0.201
Received BCG vaccination (Children age 12–23 months)									
Urban	0.635	0.074	76	80	0.055	1.347	0.117	0.487	0.783
Rural	0.351	0.020	806	802	0.017	1.178	0.057	0.312	0.391
Total	0.377	0.020	882	882	0.016	1.220	0.053	0.337	0.417
Received DPT vaccination (3 doses) (Children age 12–23 months)									
Urban	0.370	0.063	76	80	0.055	1.134	0.169	0.245	0.495
Rural	0.230	0.017	806	802	0.015	1.122	0.073	0.196	0.263
Total	0.242	0.016	882	882	0.014	1.136	0.068	0.209	0.275
Received polio vaccination (3 doses) (Children age 12–23 months)									
Urban	0.540	0.052	76	80	0.057	0.914	0.096	0.436	0.645
Rural	0.397	0.021	806	802	0.017	1.193	0.052	0.356	0.438
Total	0.410	0.020	882	882	0.017	1.183	0.048	0.371	0.449
Received measles vaccination (Children age 12–23 months)									
Urban	0.398	0.063	76	80	0.056	1.127	0.159	0.272	0.524
Rural	0.143	0.014	806	802	0.012	1.148	0.100	0.115	0.172
Total	0.166	0.015	882	882	0.013	1.182	0.090	0.136	0.196
Fully vaccinated (Children age 12–23 months)									
Urban	0.224	0.038	76	80	0.048	0.796	0.169	0.148	0.300
Rural	0.099	0.012	806	802	0.011	1.094	0.116	0.076	0.122
Total	0.110	0.011	882	882	0.011	1.078	0.103	0.088	0.133
Received Vitamin A (Children age 12–35 months)									
Urban	0.194	0.040	142	150	0.035	1.130	0.204	0.115	0.273
Rural	0.093	0.009	1,622	1,613	0.007	1.241	0.097	0.075	0.111
Total	0.102	0.009	1,764	1,762	0.007	1.238	0.089	0.083	0.120
Had reproductive health problem (Currently married women age 15–49)									
Urban	0.371	0.025	647	677	0.019	1.319	0.068	0.321	0.421
Rural	0.451	0.010	6,015	5,984	0.006	1.634	0.023	0.430	0.472
Total	0.442	0.010	6,662	6,661	0.006	1.624	0.022	0.423	0.462
Not involved in any decisionmaking (Ever-married women age 15–49)									
Urban	0.104	0.018	687	718	0.012	1.577	0.177	0.067	0.141
Rural	0.139	0.009	6,337	6,306	0.004	2.042	0.064	0.121	0.156
Total	0.135	0.008	7,024	7,024	0.004	2.008	0.061	0.119	0.151
Ever beaten or physically mistreated since age 15 (Ever-married women age 15–49)									
Urban	0.189	0.024	687	718	0.015	1.593	0.126	0.142	0.237
Rural	0.275	0.008	6,337	6,306	0.006	1.506	0.031	0.258	0.292
Total	0.266	0.008	7,024	7,024	0.005	1.534	0.030	0.250	0.282
Not worked in past 12 months (Ever-married women age 15–49)									
Urban	0.869	0.015	687	718	0.013	1.130	0.017	0.840	0.899
Rural	0.720	0.011	6,337	6,306	0.006	2.034	0.016	0.698	0.743
Total	0.736	0.011	7,024	7,024	0.005	2.044	0.015	0.714	0.757

Table A.2 Sampling errors, Bihar, 1998–99 (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	0.596	0.029	571	637	0.021	1.418	0.049	0.538	0.654
Rural	0.639	0.012	5,652	5,586	0.006	1.809	0.018	0.615	0.662
Total	0.634	0.011	6,223	6,223	0.006	1.771	0.017	0.613	0.656
Anaemic children (Children age 6–35 months)									
Urban	0.807	0.037	149	167	0.032	1.153	0.046	0.732	0.882
Rural	0.813	0.013	1,511	1,490	0.010	1.333	0.016	0.786	0.840
Total	0.813	0.013	1,660	1,657	0.010	1.309	0.015	0.787	0.838

Table A.2 Sampling errors, Bihar, 1998–99 (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	2.748	0.167	0.061	2.414	3.081
Rural	3.588	0.080	0.022	3.427	3.749
Total	3.489	0.074	0.021	3.341	3.636
Age-specific fertility rate (Women age 15–19)					
Urban	0.072	0.009	0.125	0.054	0.090
Rural	0.119	0.005	0.044	0.108	0.129
Total	0.113	0.005	0.042	0.103	0.123
Age-specific fertility rate (Women age 20–24)					
Urban	0.200	0.019	0.094	0.162	0.238
Rural	0.226	0.006	0.029	0.213	0.239
Total	0.223	0.006	0.027	0.211	0.235
Age-specific fertility rate (Women age 25–29)					
Urban	0.160	0.020	0.123	0.120	0.199
Rural	0.182	0.006	0.034	0.170	0.195
Total	0.180	0.006	0.033	0.168	0.192
Age-specific fertility rate (Women age 30–34)					
Urban	0.088	0.020	0.222	0.049	0.127
Rural	0.115	0.007	0.058	0.102	0.129
Total	0.112	0.006	0.057	0.099	0.124
Age-specific fertility rate (Women age 35–39)					
Urban	0.030	0.010	0.322	0.011	0.050
Rural	0.053	0.005	0.095	0.043	0.063
Total	0.050	0.005	0.091	0.041	0.059
Age-specific fertility rate (Women age 40–44)					
Urban	0.000	0.000	NC	0.000	0.000
Rural	0.020	0.004	0.173	0.013	0.027
Total	0.018	0.003	0.175	0.011	0.024
Age-specific fertility rate (Women age 45–49)					
Urban	0.000	0.000	NC	0.000	0.000
Rural	0.003	0.002	0.577	0.000	0.006
Total	0.002	0.001	0.576	0.000	0.005

Table A.2 Sampling errors, Bihar, 1998–99 (Contd.)

Variable/ residence	Value (R)	Standard error (SE)	Relative standard error (SE/R)	Confidence limits	
				R-2SE	R+2SE
Neonatal mortality rate (5-year period preceding survey)					
Urban	32.751	10.480	0.320	11.791	53.711
Rural	47.810	3.703	0.077	40.404	55.216
Total	46.476	3.495	0.075	39.487	53.465
Infant mortality rate $_1q_0$ (5-year period preceding survey)					
Urban	53.300	10.956	0.206	31.388	75.212
Rural	74.810	4.579	0.061	65.652	83.969
Total	72.924	4.293	0.059	64.338	81.511
Child mortality rate $_4q_1$ (5-year period preceding survey)					
Urban	15.803	8.037	0.509	0.000	31.877
Rural	36.493	2.759	0.076	30.974	42.012
Total	34.686	2.628	0.076	29.429	39.942
Under-five mortality rate $_5q_0$ (5-year period preceding survey)					
Urban	68.261	13.476	0.197	41.308	95.213
Rural	108.573	5.515	0.051	97.543	119.603
Total	105.081	5.207	0.050	94.666	115.495
Crude death rate (Based on Household Questionnaire)					
Urban	8.317	1.056	0.127	6.206	10.429
Rural	11.641	0.560	0.048	10.521	12.761
Total	11.269	0.527	0.047	10.215	12.323
Crude birth rate (Based on women's birth history)					
Urban	22.329	1.314	0.059	19.701	24.957
Rural	28.813	0.585	0.020	27.644	29.982
Total	28.106	0.542	0.019	27.023	29.190

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 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 to 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 Bihar, information on age was missing for only 6 persons out of 37,713 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 Bihar.

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 Bihar, the extent of missing information is very low for age at death, age at first marriage, woman's education, and prevalence of diarrhoea in the two weeks preceding the survey (Table B.3). Missing information is higher for the month of birth of children born in the past 15 years. It is important to note, however, that the year of birth is reported almost in every case in which the month is missing. Data on height and weight of children and woman’s haemoglobin level are available for more than 88 percent of the members of the respective reference groups. Missing information is highest (24 percent) for children’s haemoglobin level. The response rates are

Table B.1 Household age distribution

Single-year age distribution of *de facto* household population by sex (weighted), Bihar, 1998–99

	Male		Female			Male		Female	
Age	Number	Percent	Number	Percent	Age	Number	Percent	Number	Percent
< 1	514	2.7	465	2.5	38	256	1.3	234	1.3
1	477	2.5	459	2.5	39	128	0.7	159	0.9
2	465	2.4	478	2.6	40	289	1.5	191	1.0
3	516	2.7	467	2.5	41	128	0.7	121	0.7
4	567	2.9	538	2.9	42	254	1.3	232	1.3
5	651	3.4	582	3.2	43	68	0.4	84	0.5
6	711	3.7	640	3.5	44	100	0.5	124	0.7
7	477	2.5	496	2.7	45	260	1.3	175	1.0
8	723	3.7	662	3.6	46	153	0.8	149	0.8
9	469	2.4	426	2.3	47	66	0.3	109	0.6
10	659	3.4	605	3.3	48	223	1.2	150	0.8
11	379	2.0	364	2.0	49	99	0.5	48	0.3
12	684	3.5	635	3.4	50	221	1.1	64	0.3
13	351	1.8	377	2.0	51	120	0.6	67	0.4
14	465	2.4	447	2.4	52	162	0.8	150	0.8
15	383	2.0	328	1.8	53	62	0.3	87	0.5
16	447	2.3	437	2.4	54	64	0.3	99	0.5
17	239	1.2	281	1.5	55	149	0.8	148	0.8
18	508	2.6	471	2.6	56	123	0.6	156	0.8
19	249	1.3	320	1.7	57	40	0.2	63	0.3
20	353	1.8	370	2.0	58	129	0.7	135	0.7
21	225	1.2	297	1.6	59	60	0.3	62	0.3
22	410	2.1	426	2.3	60	248	1.3	136	0.7
23	219	1.1	234	1.3	61	63	0.3	77	0.4
24	291	1.5	339	1.8	62	139	0.7	129	0.7
25	353	1.8	352	1.9	63	40	0.2	41	0.2
26	350	1.8	388	2.1	64	49	0.3	46	0.2
27	158	0.8	172	0.9	65	164	0.9	103	0.6
28	339	1.8	404	2.2	66	58	0.3	54	0.3
29	163	0.8	257	1.4	67	22	0.1	36	0.2
30	387	2.0	314	1.7	68	80	0.4	75	0.4
31	145	0.8	201	1.1	69	25	0.1	31	0.2
32	377	2.0	340	1.8	70+	493	2.6	295	1.6
33	114	0.6	134	0.7	Don't				
34	164	0.8	184	1.0	know/				
35	399	2.1	295	1.6	missing	2	0.0	4	0.0
36	313	1.6	278	1.5					
37	83	0.4	107	0.6	Total	19,314	100.0	18,399	100.0

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

acceptable for the height and weight and child haemoglobin level since in any survey many children cannot be measured because they are not at home or they are ill at the time of the survey. In some cases when the child was at home, either the child refused to be measured or the mother refused to allow the child to be measured because of cultural beliefs. 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.

Table B.2 Age distribution of eligible and interviewed women

Age distribution of the *de facto* household population of women age 10–54 and of interviewed women age 15–49, and percentage of eligible women who were interviewed (weighted), Bihar, 1998–99

Age	All women	Ever-married women	Interviewed women		Percent interviewed
			Number	Percent	
10-14	2,427	14	NA	NA	NA
15-19	1,836	685	660	9.4	96.3
20-24	1,666	1,446	1,385	19.8	95.8
25-29	1,573	1,536	1,481	21.1	96.4
30-34	1,173	1,168	1,130	16.1	96.8
35-39	1,074	1,072	1,033	14.7	96.4
40-44	751	751	719	10.3	95.7
45-49	631	630	599	8.6	95.1
50-54	466	466	NA	NA	NA
15–49	8,704	7,288	7,008	100.0	96.1

Note: The *de facto* population includes all 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), Bihar, 1998–99

Indicator	Reference group	Percentage missing information	Number of cases
Birth date	Births in past 15 years		
Month only		1.20	16,022
Month and year		0.04	16,022
Age at death	Deaths to births in past 15 years	0.46	1,786
Age at first marriage	Ever-married women age 15–49	0.69	7,024
Woman's education	Ever-married women age 15–49	0.00	7,024
Anthropometry	Living children age 0–35 months		
Height		10.22	2,732
Weight		10.25	2,732
Height or weight		10.52	2,732
Woman's haemoglobin level	Ever-married women age 15–49	11.65	7,024
Child's haemoglobin level	Living children age 6–35 months	23.59	2,161
Diarrhoea in past 2 weeks	Living children age 1–35 months	0.35	2,692

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, 99 percent of living children listed in the birth history had complete birth dates recorded, as did 97 percent of children who had died. The completeness of data on birth dates for both surviving and nonsurviving children is excellent, particularly in recent years. The annual data on the number of births can be examined to see if there is an abnormally large decline in the number of births after January 1995, 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. However, the sharp drop in the annual number of births between 1992–94 and 1995–98 (particularly for nonsurviving children) suggests that there has been some omission of recent births or displacement of birth dates that could result in an underestimate of both fertility and infant mortality rates for 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 asking 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. Apparently, the training was not entirely successful in avoiding this type of problem, however.

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 Bihar, 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 77 for 0–4 years, 78 for 5–9 years, and 71 for 10–14 years preceding the survey. Although there was no severe underreporting of early neonatal deaths in NFHS-2, there was some misreporting of age at death due to a preference for reporting the age at death at 6, 8, 12, 15, and 20 days (Table B.5).

Table B.6 shows the percentage of infant deaths that occurred during the neonatal period. These percentages are also consistently high for the different time periods preceding the survey, 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, 12, and 18 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 or older. Thus, heaping at 12 months can bias the mortality estimates because a certain fraction of these deaths 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.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), Bihar, 1998–99

	Number of births			Percent with complete birth date[1]			Sex ratio at birth[2]			Calendar year ratio[3]		
Calendar year	L	D	T	L	D	T	L	D	T	L	D	T
1999	118	5	123	100.0	100.0	100.0	826	0	767	NA	NA	NA
1998	982	65	1,047	99.9	100.0	99.9	930	974	933	NC	NC	NC
1997	859	60	918	99.6	100.0	99.6	934	809	926	NC	NC	NC
1996	884	100	984	99.4	100.0	99.5	1,007	692	970	103.1	157.0	106.8
1995	856	68	923	99.4	100.0	99.4	916	835	910	84.8	52.6	81.1
1994	1,135	157	1,292	99.1	96.6	98.8	930	1,202	960	115.4	157.4	119.2
1993	1,111	132	1,244	99.2	94.5	98.7	950	940	949	95.1	82.2	93.6
1992	1,202	165	1,367	98.8	96.1	98.5	890	1,012	904	120.1	122.1	120.3
1991	890	138	1,027	98.5	97.7	98.4	963	1,038	973	73.1	83.2	74.4
1990	1,231	166	1,397	98.4	96.9	98.2	921	1,091	940	143.0	123.3	140.4
1989	832	131	963	98.7	97.6	98.6	912	1,420	968	72.6	87.0	74.3
1988	1,062	136	1,198	98.8	100.0	98.9	897	1,109	919	137.2	117.7	134.7
1993–97	4,844	518	5,362	99.3	97.6	99.1	947	925	944	NA	NA	NA
1988–92	5,217	735	5,952	98.6	97.6	98.5	915	1,117	937	NA	NA	NA
1983–87	3,750	574	4,323	98.7	97.4	98.5	977	1,124	996	NA	NA	NA
1978–82	2,739	466	3,204	98.7	96.0	98.3	849	1,005	870	NA	NA	NA
1977 or earlier	2,418	612	3,029	98.5	97.0	98.2	814	887	828	NA	NA	NA
All	20,066	2,974	23,040	98.9	97.2	98.7	912	1,008	924	NA	NA	NA

NA: Not applicable
NC: Not calculated because full-year data were not collected for 1998 and 1999 (the survey began during 1998)
[1] Both year and month of birth given
[2] (B_f/B_m)x1000, where B_f and B_m are the numbers of female and male births, respectively
[3] $[2B_x/(B_{x-1}+B_{x+1})]$x100, 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 early neonatal deaths for births during five-year periods preceding the survey (weighted), Bihar, 1998–99

	Years preceding survey			
Age at death (days)	0–4	5–9	10–14	0–14
< 1	45	59	42	147
1	54	73	65	192
2	20	29	16	65
3	16	36	11	63
4	11	17	10	38
5	12	14	9	35
6	22	21	13	57
7	9	10	5	24
8	7	10	14	31
9	8	8	3	19
10	4	11	4	19
11	3	3	2	8
12	2	5	10	17
13	0	2	1	3
14	0	0	1	1
15	4	5	11	20
16	0	0	1	1
17	1	0	0	1
18	1	2	4	7
19	0	0	0	0
20	5	6	2	13
21	1	2	2	5
22	0	2	3	5
23	0	0	0	0
24	1	2	0	3
25	1	0	2	3
26	1	0	0	1
27	0	0	0	0
28	1	0	0	1
29	0	0	0	0
30	5	0	3	8
0–30	236	317	235	787
Percent early neonatal[1]	77.0	78.4	71.0	75.8

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

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 there is some heaping of deaths at 6, 12, and 18 months of age. Digit preference appears not to be serious enough to alter substantially the mortality rates calculated here. Because the extent of heaping on 12 months is minor, probably due to the strong emphasis on this potential problem during training of the interviewers, adjustment of the infant and child mortality rates is unnecessary.

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 neonatal deaths for births during five-year periods preceding the survey (weighted), Bihar, 1998–99

	Years preceding survey			
Age at death (months)	0–4	5–9	10–14	0–14
< 1	236	317	235	787
1	22	22	16	61
2	19	24	19	62
3	16	21	14	51
4	8	15	7	30
5	7	5	5	17
6	15	23	14	53
7	8	13	6	27
8	6	9	9	24
9	2	5	17	24
10	7	7	5	19
11	4	11	9	24
12	10	8	13	31
13	6	12	8	26
14	1	12	11	24
15	3	5	5	13
16	2	3	4	8
17	1	2	4	7
18	4	17	16	36
19	0	2	2	4
20	0	0	1	1
21	1	3	0	4
22	1	1	3	5
23	2	0	1	3
1 year	7	13	7	27
Missing	3	1	0	4
0–11	350	472	356	1,178
Percent neonatal[1]	67.3	67.1	66.0	66.8

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