Ms Nandita Saikia* and P.N.Mari Bhat**

Quality of Data on Adult Mortality: An Assessment of NFHS I & II for India and Its different States

Introduction and context

There are three fundamental problems in the measurement adult mortality and its determinants that are not encountered in case of child mortality. Firstly, adult deaths are relatively rare events, and a large sample sizes are therefore needed to provide a precise mortality estimates. Secondly, unlike childhood deaths for which mother or guardian is the natural informant, for adult deaths there is no single relative or associate that can always be used as an appropriate informant. Thirdly, it may be difficult to obtain a reliable diagnosis through retrospective questioning for adult deaths than for the childhood deaths (Hayes et. al. 1989). Thus scientific research on various determinants of adult mortality in developing countries is hindered by the lack of good quality of data and also by the insufficient use of the information that have been collected already without considering the quality. This is factual in India also as the coverage of Vital Registration system of this country is very poor. Sample Registration System although appears to be consistent, it hardly provides any data regarding the background characteristics of the dead person. Unlike SRS, the National Family Health Survey provides data by the socio-economic variants, which gives a scope to analyze mortality condition more exclusively. The available information in the two large-scale surveys NFHS I & NFHS II conducted during 1992-1993 and 1998-1999 have not analyzed so far to explore this new domain of mortality. While the data provided by the two large scale surveys II & I on infant and child mortality have been much scrutinized, no research has been based on death above age group fifteen. Hence before drawing any inference on the basis of above-mentioned data, it will be appropriate to evaluate the quality of data in order to ensure the accuracy of the conclusions drawn. The National Family health Survey can be regarded as a benchmark in the history of collection of demographic data through surveys (Visaria et. al. 1999). There are some previous attempts to assess the quality of NFHS data. Bhat et. al. (1999) did one of the most extensive analysis about the findings of NFHS I data at regional level. They compared the estimates of female literacy and percentage of Muslims from NFHS 1991 census and I data and ensured that the high level of agreement of between these two sources with a very few exceptions. Singh et. al. (1999) matched the findings of Post Survey Check[1] with that of NFHS I and confirmed the high quality of NFHS I data on different demographic variables viz. age, sex ratio, literacy rate, fertility rate, knowledge of family planning method, maternal care, immunization etc. There is an apprehension among some demographers that there has been deterioration in the quality of data from NFHS I to NFHS II mainly attributable to engaging private consultancy firms for the data collection of the second phase of the survey compared with the first phase. James et. al. (2004) attempted to examine this opinion in the context of respondent's educational background and concluded that information gathered from uneducated respondents is more erroneous than from the educated groups. They further concluded that digit preference in reporting age of the household population and age at deaths of household members

*PhD Scholar, International Institute For Population Sciences, Mumbai, India.

*Former Director and Senior Professor, International Institute For Population Sciences, Mumbai, India.

[1] A post survey check of the households already surveyed in the NFHS I was carried out by the Institute for Research in Medical Statistics, New Delhi

is very common phenomenon in India and data collected by NFHS is not free from it. They have not found any difference on the quality of age information between NFHS I and NFHS II. Nevertheless there is no attempt still regarding the coverage of adult deaths on the basis of which we can infer conclusion about socio-economic determinants of adult mortality. Question may arise how assessment of quality of data will help if it is not possible to correct the already conducted survey. According to Preston, error assessment is useful as it indicates a clue to the degree of confidence that can be placed in demographic estimates.

Methodology

While talking about quality of data, two types of error come into picture; one is coverage error and another is content error. While the omission of some specific questions or omission of people or events is known to be coverage errors, the inaccuracy occurs from wrongly reported or recorded events are the content errors. All these type of errors distort the age sex data. As we will calculate the adult death rate or probability of death using the age sex distribution, it is necessary to make it free from all possible types of errors. Here we will measure the extent errors present in the age distribution of population understudy by applying Whipple's index to find the extent of preference for terminal digits 0 and 5 for the age group 12 to 62 so that depending upon the extent of errors a decision regarding adjustment can be made. Since Whipple's Index shows very rough data with reference to digit preference, single year age data has been smoothened for further analysis. After smoothing of the data, adult age specific death rate has been calculated by applying the technique of Lexi's diagram to back project usual resident age structure two years ago to the date of survey. This calculated death rate in turn has been compared with SRS report to test out discrepancies of these two sources of mortality. We followed the single year's cohort to the too years back and added the number of deaths in last two years by giving appropriate weight with the help of Lexi's diagrams. Thus the number of persons two years ago to the survey has been computed in the following way:

No of Persons Two Years Ago at Age x = $P_{x+3} + 0.25*D_x + 0.5*D_{x+1} + 0.25*D_{x+2.}$

Where P_{x+3} =Number of Usual Residents at age x+3 at the Survey date.

D_x = Number of Deaths at age x in last two years prior to the survey.

Finally, ASDRs = (Number of Deaths in the age range x to x+n) / (Number of Person Years lived in the age range x to x+n)

Where person years lived has been calculated using age structure of two points of time i.e. at the survey date and two years back to the survey date.

To check the consistency of NFHS data with SRS, a summery index $_{45}q_{15}$ for each State has been calculated by the formula below.

$$_nq_x = (n * {}_nm_x) / (1 + (n- {}_na_x)).$$

Where $_na_x = n + 1/ {}_nm_x - n/(1- \exp(-n^{*} {}^nm^x))$ and $^nm^x$ is age specific death rate.

Results and discussion

Table A presents Whipple's index for male and female separately for India and other fifteen States for the age group 12 to 62. Under the assumption that the number of persons decreases linearly as age advances, the expected value of Whipple's index ranges from 100 to 510 for the age group 12 to 62. When the age reporting is accurate (no heaping or avoidance at any ages), then we expect the sum of the persons with ages ending 0 and 5 to be exactly 10/51 th of the total number of persons in the age group 23 to 62. On the contrary, if we imagine the other extreme case when all people report their ages ending with 0 and 5, then in such a situation the

index will be 510. However, if there is dislike or avoidance for the ages ending in zero and five, then the index will vary between 0 and 100. Quality of data assumed to be vary bad or highly inaccurate with respect to age heaping if this index is higher than 175. Calculated Whipple's index values fall within highly inaccurate category for males and females from NFHS data for national and sub national except Bihar in NFHS II showing significant digit preference of the ages ending with digits 0 and five exists. It varies between 192 to 359 for NFHS I and 171 to 359 for NFHS II. Here we have divided States into two categories based on Whipple's index. What is noticeable is that some States like Gujarat and Kerala are constantly showing lower index for both sex in comparison to other States. Similarly, States like Bihar and Punjab, although not constantly, fall under good States with respect to this index. Significant differential exists between males and female in digit preference as we scrutinize the data within States for male and female separately.

Table A

Male

NFHS I		NFHS II	
Good States (WI<250)	Bad States (WI>250)	Good States (WI<250)	Bad States (WI>250)
Gujarat(248)	A.P.(333)	A.P.(229)	Assam(335)
Kerala(194)	Assam(359)	Bihar(171)	Haryana(282)
Punjab(250)	Bihar(278)	Gujarat (228)	Karnataka(304)
Tamil Nadu(244)	Haryana(297)	Kerala (136)	MP(299)
	Karnataka(331)	Maharashtra(237)	Orissa(258)
	M.P.(330)	Punjab(208)	Rajasthan(347)
	Maharashtra(322)		T.N(278)
	Orissa(261)		UP (359)
	Rajasthan(338)]		WB(355)
	UP(288)		India(271)
	WB(342)		
	India(292)		

Female

NFHS I		NFHS II	
Good States (WI<250)	Bad States (WI>250)	Good States (WI<250)	Bad States (WI>250)
Bihar(192)	A.P.(322)	Bihar(173)	A.P.(255)
Gujarat(217)	Assam(301)	Gujarat (188)	Assam(285)
Kerala(193)	Punjab(272)	Kerala (227)	Haryana(282)
Tamil Nadu(238)	Haryana(276)	Maharashtra(237)	Karnataka(271)
Rajasthan(246)]	Karnataka(252)	Punjab(237)	MP(270)
UP(221)	M.P.(330)	Haryana(236)	Rajasthan(343)]
	Maharashtra(345)	Orissa(231)	T.N(285)
	Orissa(268)		UP (321)
	WB(262)		WB(295)
	India (261)		Maharashtra(346)

Since our aim is to observe the quality of data on adult mortality, ASDRs rate has been calculated for early adult, adult and late adult for age groups 15-39, 40-59 and 60+ respectively. Thus table B presents estimated age specific death rate of above-mentioned age groups calculated from NFHS and SRS which shows the level of agreement is quite good for above age 15 in all India. Further It has been observed that the level of agreement between these two sources is better in NFHS I than that of NFHS II. Estimated ASDRs for age above 15 from NFHS I appear to be more consistent with SRS for both sex than that from NFHS II. It is quite interesting to notice that in the early adulthood viz. 15-39, females experience higher mortality than their male counterparts whereas during adult or late age, mortality among male is quite higher than that of females. Higher maternal mortality may be one possible reason for higher mortality among females during early adult group. It is observed that for the early adult and adult age group, estimates are quite good as level of consistency is better in these age groups than that of the late adult age group. However, during late adulthood, coverage of deaths in NFHS appears to be higher than that of SRS especially for the estimates from NFHS II data.

Table B. Age Specific Death Rate classified by early adult, adult and late adult age group calculated from NFHS and comparison with SRS

Age Group	NFHS II			SRS 1998			NFHS I			SRS 1992		
	All	M	F	All	M	F	All	M	F	All	M	F
15-39	3.6	3.5	3.6	2.60	2.7	2.6	3.3	3.1	3.6	2.8	2.6	3.1
40-59	10.2	11.8	8.5	8.50	9.9	7.0	9.1	10.1	7.6	9.1	11.1	8.1
60+	62.3	65.6	60.8	52.8	58.8	47.3	59.5	60.6	59.0	54.3	57.5	51.5
15+	11.5	12.6	10.7	9.30	10.2	8.5	8.74	9.3	8.3	9.6	9.8	9.3

Table C and table D disclose correlation between NFHS and SRS estimates of $_{45}q_{15}$ before and after smoothing the single year age data. It seems correlation between NFHS I and SRS 1991-1995 for males is very poor but the same for NFHS II and SRS 1996-2000 is moderately high (r=0.0586 table C). As it is evident from table, correlation between SRS and NFHS is always stronger for females then males, we can remark that mortality data for females from NFHS is better than that of males. Another important point that draws our attention is that correlation between SRS 1991-1995 estimates of $_{45}q_{15}$ and NFHS II estimates is very strong for both sex,

Table C Correlation between NFHS and SRS estimates of $_{45}q_{15}$ before smoothing

	Correlations of $_{45}q_{15}$ among males			
source	NFHS I	NFHS II	SRS1991-1995	SRS 1996-2000
NFHS I	1			
NFHS II	-0.079	1		
SRS1991-1995	0.083	0.631	1	
SRS 1996-2000	0.062	0.586	0.928	1
	Correlations $_{45}q_{15}$ among females			
Source	NFHS I	NFHS II	SRS1991-1995	SRS 1996-2000
NFHS I	1			
NFHS II	0.599	1		
SRS1991-1995	0.766	0.881	1	
SRS 1996-2000	0.712	0.860	0.967	1

Table D Correlation between NFHS and SRS estimates of 45q15 after Smoothing

	Correlations of 45q15 among males			
source	NFHS I	NFHS II	SRS1991-1995	SRS 1996-2000
NFHS I	1			
NFHS II	-0.079	1		
SRS1991-1995	0.083	0.631	1	
SRS 1996-2000	0.062	0.586	0.928	1
	Correlations 45q15 among females			
Source	NFHS I	NFHS II	SRS1991-1995	SRS 1996-2000
NFHS I	1			
NFHS II	0.599	1		
SRS1991-1995	0.766	0.881	1	
SRS 1996-2000	0.712	0.860	0.967	1

Table E Comparison of $_{45}q_{15}$ from NFHS I and SRS 1991-1995

	C.1992 Male				C.1992 Female			
State/Country (1)	NFHS I (2)	SRS 1991-1995 (3)	Rank Difference (4)	Ratio= NFHSI /SRS (5)	NFHS I (6)	SRS 1991-1995 (7)	Rank Difference (8)	Ratio= NFHS I /SRS (9)
A.P.	0.259 (9)	0.267(5)	-4	0.969	0.198(9)	0.208(6)	-3	0.953
Assam	*0.378(1)*	0.307(1)	0	1.230	0.368(1)	0.287(1)	0	1.282
Bihar	0.298(3)	0.262(8)	5	1.138	0.359(2)	0.263(2)	0	1.365
Gujarat	0.277(6)	0.266(6.5)	0.5	1.042	0.189(11.5)	0.187(9)	-2.5	1.009
Haryana	0.271(7)	0.24(11.5)	4.5	1.130	*0.258(3)*	0.165(13)	10	1.564
Karnataka	0.248(10)	0.265(7)	-3	0.937	0.217(6)	0.184(10.5)	4.5	1.182
Kerala	*0.284(5)*	0.209(14)	9	1.357	0.08(15)	0.099(15)	1	0.805
MP	0.214(14)	0.28(3)	-11	0.767	0.189(11.5)	0.236(4)	-7.5	0.799
Maharashtra	0.262(8)	0.24(11.5)	3.5	1.094	0.192(10)	0.184(10.5)	0.5	1.043
Orissa	*0.243(12)*	0.296(2)	-10	0.821	0.211(7)	0.232(5)	-2	0.909
Punjab	0.218(13)	0.251 (10)	-3	0.867	0.127(14)	0.154(14)	0	0.825
Rajasthan	0.185(15)	0.261(9)	-6	0.709	0.128(13)	0.181(12)	-1	0.709
T.N.	0.312(2)	0.266(6.5)	4.5	1.173	0.233(5)	0.196(9)	4	1.190
UP	0.244 (11)	0.272(4)	-7	0.896	0.207(8)	0.237(3)	-5	0.873
WB	0.294 (4)	0.238(13)	9	1.235	0.252(4)	0.201(8)	4	1.254
India	0.262	0.260		1.006	0.217	0.209		1.038

Table F Comparison of $_{45}q_{15}$ from NFHS II and SRS 1996-2000

	C.1998 Male				C.1998 Female			
State/Country	NFHSII	SRS 1996-2000	Rank Difference	Ratio=NFHS II /SRS	NFHS II	SRS 1996-2000	Rank Difference	Ratio=NFHS II /SRS
A.P.	0.315(3)	0.284(4)	1	1.108	0.251(5)	0.204(6)	1	1.229
Assam	0.281(7)	0.316(1)	-6	0.889	0.311(1)	0.294(1)	0	1.057
Bihar	0.309(4)	0.248(9)	5	1.247	*0.306(2)*	0.249(3)	1	1.230
Gujarat	0.272(9)	0.249(8)	-1	1.092	*0.229(7)*	0.170(10.5)	3.5	1.349
Haryana	0.191(14)	0.241(13)	-1	0.792	0.168(13)	0.157(13)	0	1.070
Karnataka	0.300(6)	0.274(5)	-1	1.094	0.200(10.5)	0.182(9)	-1.5	1.099
Kerala	*0.140(15)*	0.199(15)	0	0.705	0.091(15)	0.095(15)	0	0.958
MP	0.306(5)	0.291(2)	-3	1.050	0.293(3)	0.244(4)	1	1.202
Maharashtra	0.248(11)	0.244(11.5)	0.5	1.018	*0.221(8)*	0.164(12)	4	1.351
Orissa	0.264(10)	0.289(3)	-7	0.914	0.286(4)	0.251(2)	-2	1.140
Punjab	0.231(13)	0.244(11.5)	-2.5	0.947	0.166(14)	0.170(10.5)	-3.5	0.977
Rajasthan	0.275(8)	0.246(10)	2	1.116	0.200(10.5)	0.156(14)	3.5	1.280
T.N.	0.345(1)	0.263(6)	5	1.310	0.203(10)	0.187(8)	-2	1.084
UP	0.322(2)	0.257(7)	5	1.255	0.239(6)	0.224(5)	-1	1.065
WB	0.242(12)	0.232(14)	2	1.045	0.188(12)	0.194(7)	-5	0.968
India	0.290	0.258	-	1.122	0.236	0.198	-	1.192

Table E and F give a picture of comparison of $_{45}q_{15}$ from NFHS I and II with SRS for the respective time period. In addition rank has assigned in descending order of adult mortality to each State for both NFHS and SRS estimates. At the national level, the absolute value of 45q15 from both the sources is very close to each other. Again if we assign rank to India, rank of 45q15 will coincide for NFHS and SRS in table E and F except slight difference in table E for male adult. At the State level, it is found that Assam accounts for highest adult mortality as it is placed in the first rank for both male and female during NFHS I which shows high level of agreement with SRS estimates. During this period Rajasthan and Kerala shows lowest mortality among male and female adults respectively. However, according to SRS 1991-1995, Kerala is the State with lowest mortality for both male and female. Table E also reveals that lowest mortality is found in Kerala for the both sex and also from the both sources. Hence 45q15 for Kerala from NFHS I data is dubious for us where we observe a high level of disagreement in rank as well as in actual figure from those two different sources. The absolute difference between NFHS I and SRS 1991-1995 estimate of 45q15 is always less than 0.05 except in a few cases viz. in Assam, Kerala and Orissa for male adults and in Haryana female adults. States like Punjab, Gujarat, Maharashtra, Andhra Pradesh and Karnataka display consistently better concordance in rank than the other States during this period. The differential between male and female is quite evident in the rank differences shown in the column (4) and (8) respectively. Further column (5) and (9) provide evidence that the NFHS estimates are smaller than that of SRS in States like Andhra Pradesh, Madhya Pradesh, Orissa, Punjab, Rajasthan and Uttar Pradesh.

As mentioned earlier, Kerala stands for the lowest mortality for both sex according NFHS II as well as SRS 1996-2000. Assam again shows highest mortality for male and female according to SRS 1996-2000 but difference in rank in male mortality between NFHS and SRS is quite high. However when we consider the absolute difference of 45q15 between these two sources, we found it as low as 0.035.During this period also, States like Kerala, Haryana, Gujarat, Karnataka, Andhra Pradesh, Maharashtra and Punjab are showing very slight difference in the rank assigned by the two sources. The important conclusion we can draw by comparing table E and F is that the concordance in rank during NFHS II is better than that of NFHS I for both male and female. Further unlike NFHS I, $_{45}q_{15}$ from NFHS II is always higher than the corresponding SRS figures in almost all cases which may be due to better coverage of adult deaths in NFHS II.

Summary: From the above discussion, it is clear that quality of data found to be very bad or highly inaccurate with respect to age heaping in the terminated by either zero or five. Therefore we have smoothed the data for further analysis. A fairly good level of concurrence has been observed for broad age group especially for early adult, adult and late adult age group for both NFHS I and NFHS II. Correlation between SRS and NFHS is always stronger for females then males; we can conclude that mortality data for females from NFHS is better than that of males. Comparison of $_{45}q_{15}$ estimates from NFHS and SRS shows that the data quality of NFHS II is superior to that of NFHS I as NFHS II estimates are more consistent with SRS estimate which give us the assurance to analyze these data with further objective.

References

Bhat, Mari P N 1995 "On the Quality of Birth History Data Collected in National Family Health Survey", *Demography India* vol. 24(2) Pp. 245-58.

Bhat, Mari P N and Zavier, (1993) "Findings of National Family Health Survey: Regional Analysis", *Economic and Political Weekly* vol.34 (42): Pp243

Hayes, Richard., Mertens, T., Lockett, G. and Rodrigues, L. (1989) "Causes of Adult Deaths in Developing Countries Areview of Data and Methods" working paper Population and Human Resource Department

International Institute for Population Sciences (IIPS). 1995. *National Family Health Survey, India 1992-93*. IIPS, Mumbai and ORC Macro.

--------. 2000. *National Family Health Survey, India 1998-99*. IIPS, Mumbai and ORC Macro

James, K.S. and Rajan, Irudaya (2004) "Respondents and Quality of survey Data", *Ecerronomic and Political Weekly* vol. 39 No 7 Pp 659-663.

Preston, S. H, Elo T., and Stewart, Q. 1999 "Effects of Age Misreporting on Mortality Estimates at Older Ages" Population Studies. Vol 53. : 165-177.

Preston, S.H, Heuveline, P. and Guillot, M. (2001) "*Demography : Measuring and modeling population processes*" Blackwell Publisher , Oxford

Sample Registration System Statistical Report 1992. Registrar General, India, New Delhi.

Sample Registration System Statistical Report 1998. Registrar General, India, New Delhi.

Singh, P. 1999 "National Family Health Survey , 1992-1993 Post Survey Check", *Economic and Political Weekly* vol.*34* No 42 &43 Pp 3084-3088.

SRS BASED ABRIDGED LIFE TABLES 1990-94 AND 1991-95 Registrar General, India, New Delhi

SRS BASED ABRIDGED LIFE TABLES 1996-2000 AND 1997-2001 Registrar General, India, New Delhi