Brief description of the MHSS: The MHSS was carried out as part of a program project, Healthy Aging in Rural Populations, supported by an award from NIA and NICHD to RAND (Omar Rahman, Harvard University, P.I.; Andrew Foster (Brown University), Paul Gertler (University of California, Berkeley) and Jane Menken (University of Colorado at Boulder), co-P.I.s) with subcontracts to the various institutions involved and to the International Centre for Diarrhoeal Disease Research, Bangladesh. It was also supported in part by a Mellon Foundation grant to the University of Pennsylvania. Data come from both areas (Comparison and Maternal and Child Health and Family Planning Area) in ICDDR,B's Demographic Surveillance System. Over 4500 households are included in the two subsamples - a primary sample of 2695 households and a secondary subsample of 1842 households chosen because of their relationship to a primary household.

The MHSS interviewed relatively large samples of men and women aged 15 years and collected data on a relative large sample of children. It provides extensive information on composition and economic status of the household, social networks, and characteristics of kin, on health (measures include days lost to illness, presence and duration of various symptoms, reported and observed activities of daily living), on education of children (including type of schools attended, dates, financial support for schooling) as well as migration, marriage, and fertility histories. The MHSS will be available as a public use data set early in 1999.

MHSS Sampling Design

The MHSS was carried out in the surveillance area over a period of eight months in 1996 as part of a Program Project funded primarily by the National Institute on Aging with additional support from the National Institute of Child Health and Human Development and the Mellon Foundation. The Project is a collaboration among investigators at RAND (including Omar Rahman, now at Harvard), the University of Pennsylvania (including Jane Menken, now at the University of Colorado at Boulder and Andrew Foster, now at Brown University), and the International Centre for Diarrhoeal Disease Research, Bangladesh (including Nizam Khan (now a graduate student at the University of Colorado)). The data set is expected to be available for public use in the near future. The Matlab Demographic Surveillance System conducted by ICDDR,B since the early 1960s provided the sampling frame for the MHSS and information going back to the 1974 census for all individuals in the sample households.

The MHSS collected extensive information on multiple domains, including health status, marriage, fertility, and migration, socioeconomic status, and social networks from approximately 38,000 individuals in a sample of over 7000 households.

A distinctive feature of the MHSS is its multistage sampling procedure which takes into account social structure in rural Bangladesh and was intended to permit study of relationships among family members who did not coreside. The first stage was a random sample of baris followed by a sample of households within the *bari*, then by a sample of individuals within the household. A *bari* is a locally well-recognized cluster of households around a central courtyard whose heads are usually assumed to be related. These households are believed to cooperate with one another to varying degrees. The latter is particularly true in the case of women who tend to leave the *bari* only rarely. *Bari* members have a strong sense of identification with the *bari* and most *bari*s are known by a name which refers to the

occupation of its most successful member (e.g. Sarker [clerk-manager], Patoyari [revenue collector]).

Although the above description characterizes the typical *bari*, there is a fair amount of heterogeneity in both physical characteristics and social organization. On average, a *bari* consists of 37 individuals living in 5-6 households. However *baris* range from single households (16%) to twenty-six, with population size varying correspondingly.

A major purpose of the MHSS was to provide information on the network of relationships between households related either by kinship or proximity. The *bari* sample is a multistage stratified sample that led to two sub-samples chosen in the following way. In the first stage, 2,597 *baris* were chosen at random from the total of about 8000 *baris*. As described above, the number of households in a *bari* may vary from one to more than twenty, with the mean being 5.5. For all *baris* with more than one household, the second stage consisted of picking a *first* and a *second* household. The *first* household was chosen at random and is part of the *primary* sub-sample. In *baris* with two households, the second automatically became part of the *secondary* sub-sample. For *baris* with three or more households, the second household was chosen on the basis of relationship to the head of the first household (with the presence of parents followed by sons and then brothers being the criteria for selection). In the absence of any related household within the *bari*, the second household was chosen at random. Data from the second household permits all participants in the inter-household relationship to provide information. Thus, the *bari* sample has two types of households (*first* and *second*) that constitute respectively the *primary* and *secondary* sub-samples of households.

Of the 2597 distinct *baris* selected, 1941 (75%) had two or more households while 656 (25%) had only one household. There are thus 2597 households in the *primary* sub-sample and 1941 in the *secondary*, totaling 4538. Within each household, all individuals aged 50 and above were selected for interview. For those between the ages of 15-49, if an individual was the head of the household, that person and his/her spouse was always chosen. Similarly if an individual in the 15-49 year age group was the spouse of an individual aged 50+, that person was also always chosen. From the remaining 15-49 year old members of the household (not falling into one of the above two relationship categories), one person at random was chosen. If that individual was currently married and the spouse was coresident, the spouse was also chosen. This procedure ensured that unmarried adults aged 15-49 (a group that is often under-represented in surveys) would have a finite probability of being selected. In point of fact, given household sizes of about five, there was a complete census in most cases of individuals aged 15 years and above. Preliminary counts show that, in the *bari primary* and *secondary* sub-samples, 11,190 men and women aged 15 years and over (7,761 between ages 15 and 49, and 3,429 aged 50 and over) were interviewed between April and November 1996 with respect to their social networks, health status, and socioeconomic information. The table below gives the breakdown by age and sex for these sub-samples and their total.

A second sample is drawn from the households of all women who participated in the Determinants of Natural Fertility Study (DNFS) in the mid to late1970s who could be located in 1996. About 2400 women (all those of reproductive age in four villages) were followed prospectively for nearly three years on a monthly basis. Detailed information is available on their reproductive status (breast feeding, menstruation, pregnancies and their terminations, births, etc.), their health as measured by anthropometry and some blood testing, and the

health of their children. Extensive analysis of this data set has improved our knowledge of the relationship between current health and fertility. Because of our interest in the effects of a woman's reproductive career on her later health and well-being, and because this group provided prospective information on women's health over 20 years prior to the MHSS, we decided to have a separate sample of their households, wherever they could be located in Bangladesh (1776 DNFS households participated in the MHSS; for the remainder, dates of death or out-migration are available). Again, respondents within the household were chosen for the *DNFS* sample in the way described above.

Preliminary counts for the *bari primary* and *secondary* sub-samples and the *DNFS sample* are:

Table 1.

	Males aged				Females aged			
	15-29	30-49	50+	Total	15-29	30-49	50+	Total
***Bari* Prim+Sec HH**	1565	1801	1736	5102	1902	2493	1693	6088
Primary HH	982	1103	1061	3146	1165	1527	1030	3722
Secondary HH	583	698	675	1956	737	965	663	2366
DNFS	888	571	962	2421	760	1044	988	2792

Because of the complex sampling designs, appropriate weights for each individual will be included in the final data set. Final weights were not yet available when these preliminary studies were carried out. However, calculations carried out with and without an early version of the weights yielded similar results. We have chosen to carry out these studies without weights. All analyses will, of course, be repeated when the final weights are available.