Aging and health care are the emerging policy issues in the Third World. However, we currently do not have the data to address these issues because economic status and health have not been integrated into a single survey design. This article discusses the rationale for the principal features of an emerging new international survey design which includes integration of younger and older families; reliance on retrospective data; intensive measurement of economic status, health outcomes and utilization and intergenerational transfers; and the combination of a household and community survey.
Key Words: International data, Health and economics, Intergenerational transfers, Retrospective data

Measuring Health and Economic Status of Older Adults in Developing Countries

James P. Smith, PhD

The central research questions on aging cross national boundaries. Even more so than in the United States, however, inadequate data have been the major limitation on aging research in developing countries. In most Third World countries, research and data collection has focused on the other end of the life-cycle: fertility and family planning or infant nutrition and mortality. The major multinational surveys — the World Fertility Surveys (WFS) and the Demographic and Health Surveys (DHS) — provide little information about what life is like for older people in these countries. Unfortunately, the policy issues of the next decades will reflect a very different time bomb than the one for which policy makers have been preparing. The loud ticking that they should now be hearing is the count of the aging of their populations. However, they have little data to prepare themselves for that demographic reality.

In light of this reality, a growing number of researchers have been involved in significant new data collection efforts in many developing countries. While these efforts have been unrelated administratively, they have been linked by a common set of themes and concerns. First, they all strive to have adequate samples of the elderly population. More importantly, they have learned the painful lessons of the earlier U.S. surveys of the elderly by integrating health, social support, and economic modules into a single survey design. For this integrative objective, these surveys have been explicitly designed to parallel the two new major social science aging surveys in the United States — the Health and Retirement Survey (HRS) and the Assets and Health Dynamics of the Oldest Old (AHEAD). Finally, they do not repeat another earlier mistake by recognizing that the elderly do not live in isolation and that the generations must be linked. While there are a number of such efforts, the character of these surveys will be illustrated herein with the RAND Indonesian Family Life Survey (IFLS).

Policy Framework

These new surveys were guided by a specific policy framework. Most governments are trying to effect some outcome — improved health status or reduced mortality are only two of many I could mention. For example, they can open new health centers, change user fees, promote an immunization program, and so on. But between those policy instruments and goals lies the behavioral reaction of households and service providers. Well-intentioned policy often fails because we do not anticipate or understand these important behavioral reactions. To put it most simply, if individuals don't respond to policy changes, no social goal will be achieved.

The implication is that our surveys must strive for state-of-the-art measurement of the final outcomes of interest (health, income, schooling, etc.), the proximate determinants of those outcomes (health care utilization), the attributes of the individuals and families who make these decisions, and the programmatic determinants of these behaviors, including the price variables emphasized by economists. As I will discuss later, these factors should be measured at both the individual and community levels.

One of the departures of these surveys from the current norm is that equal weight is given to the health and economic modules. Disciplinary specialization has meant that health and economic surveys have been fielded separately in developing countries with little substantive overlap. The consequence of this separation has been that neither the health nor economic surveys have been able to fully achieve their own stated objectives. The key risks to successful aging rest in a complex two-way interaction between wealth and good health and the social and

[1]This research was supported by grant P-01-AG08291 from the National Institute on Aging.
[2]Address correspondence to Dr. James P. Smith, RAND Corporation, 1700 Main Street, Santa Monica, CA 90407.

financial support networks set up within families. If anything, the potential impact of health on income is even more direct and immediate in developing countries. When physical labor is most important, poor health means that energy levels are difficult to sustain, bouts of sickness are more frequent, and crop output falls. But causation may well run in the other direction. Better-off farmers have better diets, and their nutrition levels are higher. Both directions of influence are plausible and both may well be quantitatively large. However, unless surveys of the elderly give sufficient weight to both the health and economic domains, neither the determinants of health outcomes nor those of economic performance will be well understood.

I will highlight four innovative features of these new emerging international aging surveys using the IFLS as a prototype. These four features are its sample design, use of retrospective data, substantive instruments, and its dual reliance on community and household data.

The Indonesian Context

One point that should not be controversial is that aging and health is the emerging issue in the Third World. The demands placed on health care systems in developing countries will grow dramatically in the coming decades. Not only will the capacity of their systems have to expand dramatically, but the type and mix of services will have to change to serve an older population with a markedly different pattern of diseases.

This point can be illustrated with one developing country, Indonesia, but it applies to many others. One reason for their increased concern about the elderly is purely demographic. As in many other developing countries, Indonesian birth rates have been declining rapidly and will continue to do so for the foreseeable future. At the same time, mortality risks at all ages have been falling rapidly. The demographic consequence is certain — a dramatic tilt in the age structure toward the old. In Indonesia, the size of the population of those age 60 and over will quadruple over the next 30 years, while that of younger people will stabilize. These shifts in Indonesia's age structure combined with other likely demographic trends — increased urbanization and rising incomes — foretells a need for a very different public health system than the one that exists there today. As demonstrated in Figure 1, health risks will shift dramatically from the acute infectious diseases of the young — diarrhea, respiratory infections, and measles — toward the more chronic noninfectious diseases of the old — heart disease and cancer.

Figure 2 highlights the pressures that are being placed on the Indonesian medical system by the aging of their population. Inpatient admissions will almost quadruple among those age 60 and over. With this dramatic shift in patient mix, the Indonesian health system may well have to be revamped from top to bottom. Both the existing personnel and the physical facilities are geared for a much different population of patients than that which will exist tomorrow. Moreover, this tomorrow is not in some distant future — a problem safely put aside for the next generation. Mortality rates for acute infectious diseases fell by a third between 1985 and 1990, and chronic disease mortality may well outrank deaths due to infectious disease by 1995.

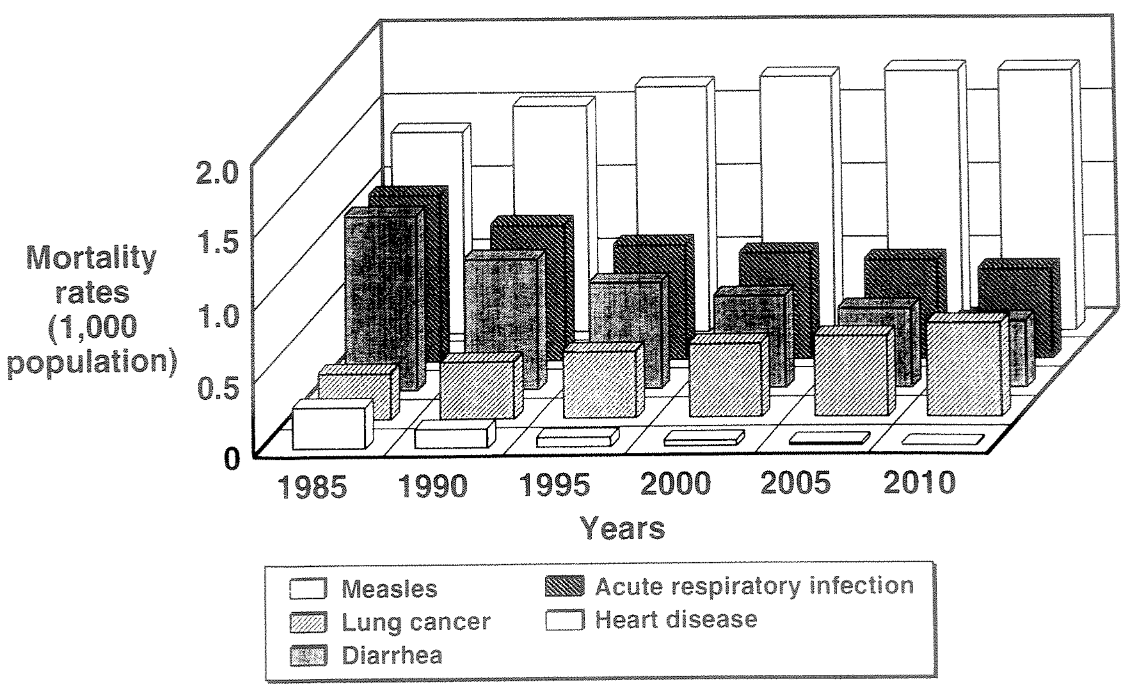

Figure 1. Specific causes of mortality, present and projected.

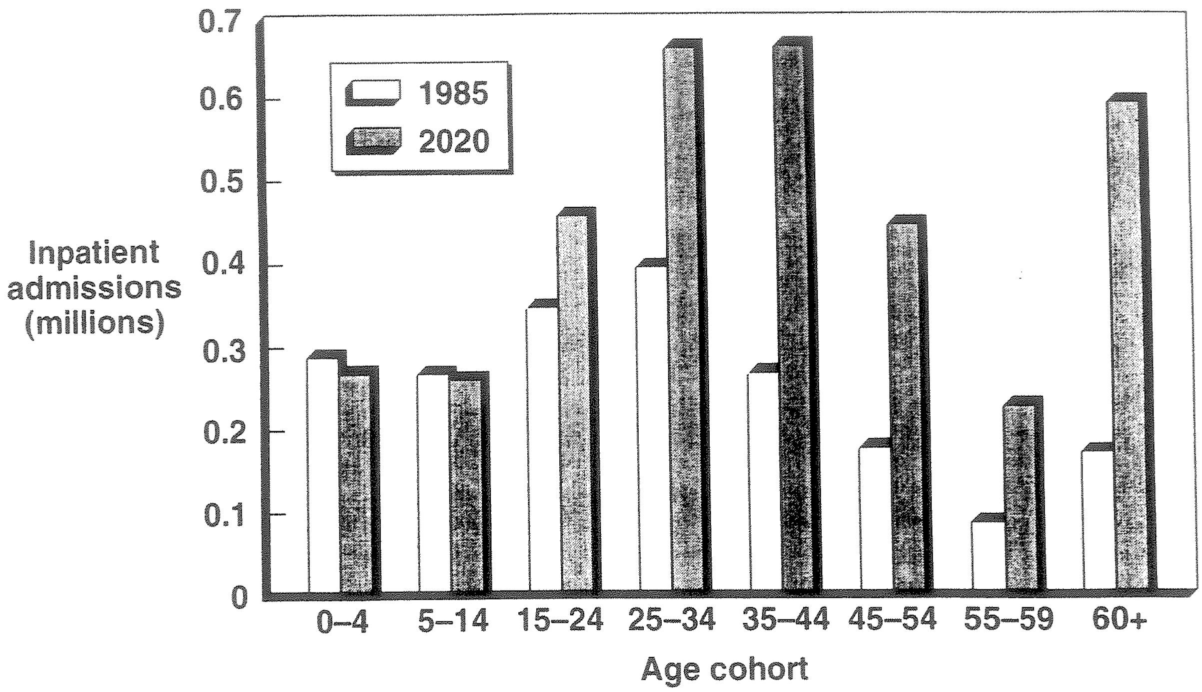

Figure 2. Inpatient admissions by age cohort; Indonesia, 1985 vs 2020.

Sample Design

One unique feature of these new international aging surveys rests on their sample design. Too often, surveys of the elderly are treated as isolated samples in which only people above a given age are eligible. For both substantive and cost reasons, this is a mistake. Surveys of the elderly population are much more cost-effective when they are combined with the more traditional demographic surveys that continue to dominate scientific work in these countries. The substantive argument is even stronger because it is impossible to understand the behavior of the elderly population unless their mutual dependency with the younger adult generation is monitored.

My point is illustrated in Table 1, which lists household sample sizes in the RAND IFLS. The first row mimics the conventional demographic survey of the 1970s and 1980s, the World Fertility Surveys or Demographic and Health Surveys being the most important prototypes. Their primary sampling units are currently married or ever married women (EMW). In the case of IFLS, our sample consists of 5,173 such ever married women less than age 50. This conventional demographic sample was to be fielded anyway, but we decided to augment it with a sample geared to the older population. Many of them were not difficult to find.

For example, 1,835 of these households contained a senior member (an individual over age 50). While these older persons were there for the asking, they are always ignored in fertility-oriented demographic surveys. In fact, a randomly selected elderly person and spouse were interviewed so that there were more than 2,500 elders present in the households with an EMW less than 50. Since a large fraction of survey costs involve preparing the instruments and contacting households, the marginal costs of including these older people were low.

Table 1. Sample Design, Indonesian Family Life Survey

Households	With Person Aged > 50	Without Person Aged > 50	Total
With an EMW[a]	1835	3338	5173
Without an EMW	1287	540	1827
Total	3122	3878	7000

[a]EMW = ever married woman.

To obtain a random sample of persons over age 50, we only had to contact 1,287 additional households not part of our normal demographic sample. These were households with seniors who did not live with an ever married woman less than 50. Many of these seniors were married and both spouses were interviewed. Thus, at a cost of less than 1,300 new households, a sample of more than 4,300 seniors was obtained. More than 2,000 of these people were over 60 and more than 800 will be more than 70 years old. One reason this sample design is appealing is its cost-effectiveness. The cost of conducting these two samples separately was 75 percent higher than combining them. The marginal cost of the elderly subsample was one-fourth of the cost of a stand-alone elderly sample.

This design was not driven by convenience alone, but it is important for substantive reasons as well. In most Third World countries, the lives of the generations are so intertwined that one cannot speak coherently about the welfare of the older generation without reference to their substantial interaction with the younger generations. The scope of these interactions spans living arrangements, shared economic resources including extensive remittances and transfers, and time and care during episodes of poor health or bad economic times. We have already seen that a large fraction of the senior sample lives in a household with a young married woman. However, even when the generations live apart, they often behave as a geographically extended household, assisting each other with life's vagaries. Surveys of the elderly that ignore these generational links may be flawed.

Retrospective Data

Perhaps the most controversial feature of our approach is its extensive reliance on long-term retrospective data. Until recently, such data had a poor reputation with doubts cast on the ability to recall events two months ago, much less a decade or more in the past. While memory decay is a legitimate concern, the potential gains from retrospective data are also large. Long-term prospective panel data are simply not currently available in these countries, and even in the best of circumstances, will not be for a decade or two in the future. In addition, the success of the original RAND Malaysian survey and the current waves of DHS have reopened the issue of the quality of retrospective data. Most demographers now accept that complete fertility and marital histories are both feasible and desirable. But what about less salient aspects of people's lives?

Questions about events decades ago are unlikely to yield as reliable a response as asking what happened last week or last year. Unfortunately, while opinions are often strongly held, there exists little scientific evaluation of the quality of retrospective questions. The two RAND Malaysian Family Life Surveys give us the potential of quantitatively evaluating the quality of long-term recall data. The first wave, MFLS-1, collected information on about 1,200 women and 900 of their spouses during 1976–77. As part of the second wave, which took place in 1988–89, updated information was gathered on the households originally surveyed 12 years earlier. The degree of correspondence between these two histories for the pre-1977 time span allows us to evaluate the quality of the retrospective data. This evaluation can be conducted on any of the life histories contained in the data, but I will summarize here the results of a recent study of the migration histories (Smith, Thomas, & Karoly, 1993).

In the Malaysian surveys, respondents were asked to recall the date and location of all migrations that lasted 3 months or more. Our evaluation simultaneously illustrates the problems and potential of long-term recall questions. The type of memory decay that raises concern about recall data with time is certainly evident in the data. For example, the longer ago the migration occurred, the less likely a respondent reported both a month and year. When the move took place more than 40 years ago, only one in seven respondents told us the month and year. In contrast, for migrations within the last five years, over 70% gave a specific month and year. Similarly, 61% (68%) of all male (female) migrations reported in the original survey were also remembered in the follow-up. At first blush, the statistics in the last paragraph give credence to a pessimistic verdict on the usefulness of long-term recall data. This verdict is premature because we were able to isolate which migration events people remember and which they forget. Respondents remembered salient moves, those linked with other important life events such as the start or breakup of a marriage, the birth of a child, or moves that lasted for a long time. In contrast, migrations that dim in memory as time passes were short duration or short distance moves often made by the young before they were married. This kind of mobility is not successfully retrieved by retrospective data. But for many purposes these may be less important moves for analysis, especially for the older population.

How useful then are retrospective migration data? A summary index involves the fraction of time respondents report themselves in the same location in both surveys. This fraction is 86% for all respondents and more than 90% among those over age 50. These positive summary evaluations are not meant to imply that long-term recall data are an easy panacea, for we have already seen that they certainly are not perfect. In addition, less positive bottom lines will surely be made for other outcomes of interest. However, until recall data for other outcomes are scientifically tested, the current presumption against long-term retrospective questions is premature.

Measurement Issues

These new international aging surveys face many complex questions about how to appropriately measure the concepts. While these issues are present in every part of the survey instrument, I will briefly discuss here our approach to three of the more central and least settled — health, income, and intergenerational transfers.

Because health is multidimensional and inherently difficult to measure, we pursued a multidimensional measurement approach. One common set of questions that proved not to be useful were probes about the past or current prevalence of specific diseases, such as cancer, heart disease, or the like. Because of limited physician contact, such diagnoses are rare. Even when physicians were visited, there was a cultural bias not to inform seriously ill patients about specific diseases. Instead, we asked about symptoms of illness such as headache, coughs, nausea, fever, or diarrhea during the last month. In measuring current health conditions, we rely heavily on self-reported health measures — overall health status, activities of daily living (ADLs), and limitations on daily activities. ADLs have been extensively tested for validity and reliability in the United States (Ware, Davies-Avery, & Brook, 1980) and have been shown to be strongly correlated with self-assessed health and economic status. These functional measures were extensively tested. With appropriate sensitivity to the different cultural contexts, both the physical and emotional measures appear to be equally useful in Third World settings. Especially in developing countries, anthropometric measurement has been particularly useful in predicting economic performance. In the IFLS, height and weight are measured for all children younger than 5 and for the biological parents of these children.

Our health module concludes with extensive detail on the principal behavioral proximate determinant — outpatient and impatient utilization. Our instruments are comprehensive and span type of health care — physician, nurse, traditional practitioner, hospital, local health clinic, self; time and money costs of treatment; and type of treatment. While we experimented with other options, a 4-week recall period for outpatient and one year for inpatient care did as well as any other.

I argued earlier for the critical importance of income and wealth for behavioral analysis of the elderly. In spite of this, most demographic and health surveys have not collected adequate income data. While the reasons are varied, an often cited one is the difficulty of measuring income in settings where markets are informal and most income is not received in monetary forms. We are following three approaches: The first relies on a detailed questionnaire about labor supply, occupation, and wages on all jobs in the month prior to the survey. A persistent finding by economists is that current income is often a very inadequate measure of the long-term economic status of households. Based on extensive pretesting, retrospective questions were asked about occupation and work for the first full-time job after school completion, the job held 10 and 20 years ago, and for each of the 5 years preceding the survey. Because a substantial fraction of Indonesian households still work in agriculture, our second approach measured gross revenues and costs for the enterprise as a whole. A similar approach was followed for all households who owned nonagricultural enterprises.

Most economists now argue that the best income measure in developing countries is aggregate household consumption. Unfortunately, such data are difficult to collect. Many government expenditures surveys take 3 hours or more to administer, and the best current alternative, the World Bank Living Standard Measurement Surveys (LSMS), took 50 minutes. Neither is a practical alternative in surveys with other priorities. Fortunately, while we were in our design phase, the World Bank was simultaneously conducting an experiment on the periodic consumption module of SUSENAS (National Social and Economic Survey) a well-established government survey in Indonesia. This consumption module contained more than 300 individual items, and took one and one-half hours on average to administer. The Bank wanted consumption as part of the regular core. To evaluate the feasibility, the Bank conducted a pilot study of 8,000 households in which both the current long form and a shorter form (25 items and 15 minutes to administer) were tested. The results were remarkable. The mean of the two consumption distributions were within one percent of each other, and one could not reject the hypothesis that the two distributions were identical. The most important implication was that one could obtain adequate measures of household consumption at reasonable survey time cost. Our consumption module is a compromise between the short and long forms (50 items) and takes about 24 minutes to administer to the average household.

There are two problems that have bedeviled measurement and interpretation of transfer data — their episodic nature and the opportunity set over which the transfers are defined. Most surveys in the United States, as well as in developing countries, collect demographic data only on those family members who currently live in the household. When they deal with non-coresident kin, often only the attributes of those who actually gave or received transfers are recorded. This partial recording of the opportunity set of people who can give transfer can make the eventual transfer data uninterpretable. For example, a large transfer of help in the form of hours of time from a highly educated male sibling is more unusual if there are four low-schooled sisters close by. Since it was essential that at least the full household opportunity set was defined, our roster includes a basic set of demographic data for all children, parents, and siblings whether they reside in the household or not. A very detailed set of characteristics is available for all relatives living in the household. For those outside the household, the detail varies by the relationship. We ask for non-coresident parents, their age, schooling, work status, occupation, location, health status, and degree of contact; for children, we ask their age, sex, schooling, location, and marital status; and for siblings, we ask their age and sex. Hours of time and money transfers in and out of the household were then asked about as a subset of this complete list of potential relations. This subset included up to four non-coresident children and siblings, for whom a more complete list of attributes was obtained. The additional attributes for children were their work status, occupation, extent of contact, health status, and land ownership. The corresponding added sibling characteristics were schooling, marital status, work status, occupation, location, health status, and extent of contact. To select the four children or four siblings, interviewers asked about the oldest and youngest and the two closest in age to the respondent. While other reasonable selection algorithms exist, it is essential that they not be selected based on the outcome of interest. In particular, while it might seem tempting to do so, relatives should not be selected based on the existence or size of the transfers. In addition to this one-on-one recording of transfers, the total amount of transfers for all nonlisted children and siblings was recorded. Transfers are also difficult to measure given their episodic character. Many transfers are triggered by other catastrophic events, such as an illness in the extended family, a severe economic shock, marriage, or death. In the IFLS, we opted for a two-track approach to this problem. The first relied on capturing all time and money transfers that took place within a recent time period (the last year). This track has the advantage of completeness and places the least strain on recall. The second track attempted to capture "domain specific" transfers associated with other salient events: marriage (dowry and bride prices), divorces (assets at separation), migration (remittances and job help), illness, death of a parent (inheritances), and economic shocks. Not only are significant transfers associated with these events, but imbedding transfer questions inside these salient events is the best aid in prompting memory of these transfers.

Community Level Data

The last dimension of the survey involves its dual emphasis on household and community level data. Measurement of community characteristics is critical, because they contain the instruments that can be manipulated by public policy, e.g., the prices, availability, and quality of the services provided at health facilities and the kind of environment the people live in. Since health is a stock that reflects current as well as past investments, it is not only the kind of current community that respondents live in, but what their communities were like in the past. We started with a detailed review with our Indonesian colleagues and the relevant ministries of the major parameters of public policy programs and how they have changed over time. This review shaped all our attempts to measure characteristics of the community. To measure the key community level variables, we have followed three complementary approaches.

The first step involved a community questionnaire given to the village head or "informed" leader. This questionnaire describes a number of key dimensions of the community, including its history and economic base, characteristics of public programs (agricultural aid and loans), infrastructure (electricity, water supply, irrigation, transportation, and financial institutions), the number and type of education and health facilities available, average housing quality and rent, the prices of stable crops, and wages of daily workers. Second, we conducted a major survey of all family planning clinics, schools, and health facilities that were mentioned by respondents in the household survey. Among other things, this health facility survey measures such critical variables as distance and transportation cost from the village, type and date of disease epidemics, and the history and type of immunization and other public health programs.

In addition, a comprehensive description of the quality of the facility and its staff was obtained. This description included a list of all medical services provided at the facility; numbers, specialty, and training of staff; type, price, and quality and availability of specific services provided; medical equipment present; and specific drug prices and availability. Finally, the household data will be linked with other secondary Indonesian data that have been measuring community characteristics. Indonesia has conducted a number of studies over time at the community level, i.e., PODES (Village Infrastructure Survey) and SAKERNAS (National Workforce Survey).

Conclusion

Because of their rapidly changing age structures, aging is fast becoming the emerging policy concern in developing countries. However, these countries will have even less time than was available in the U.S. to adapt their institutional structures to the health and income problems of their elderly. Fortunately, there are a number of new aging surveys in several developing countries that have the potential of filling the gap. In many ways, these surveys match or are superior to the best aging data sets in this country. This article highlighted some of the more innovative features of these new international surveys in order to increase their visibility across the gerontological community.

References

Smith, J. P., Thomas, D., & Karoly, L. A. (1993). *Migration in retrospect: Differences between men and women*. Unpublished paper.

Ware, J., Davies-Avery, A., & Brook, R. (1980). *Conceptualization and measurement of health status for adults in the health insurance study: Volume VI. Analysis of relationships among health status measures.* R-1987/6-HEW. Santa Monica, CA: RAND.

Received January 5, 1994
Accepted May 8, 1994