

**Health and Retirement Study
External Review
Lewis H. Kuller, MD, DrPH**

I have based this review on materials provided by the University of Michigan Health and Retirement Study (HRS). I have also reviewed the pink sheets from the Asset and Health Dynamics Among the Oldest Old (AHEAD) Study, documents provided by Dr. Robert Wallace regarding the focus of the AHEAD Study and some of the special studies, a statement to the HRS Data Safety Monitoring Board and, most important, published papers and unpublished manuscripts from the study including descriptions of the goals of the study and subsequent analyses. I have limited my review to the health components, as I am not an expert in other areas although I realize they are equal or of greater importance.

The HRS and the AHEAD Study and the recently added studies are very well designed and have continued to demonstrate good science. The follow up of participants has been superb. The primary goals of the study to link health, disease histories, symptomatology with functional status and disability, economic factors, retirement and health services utilization and impact on societal questions were well-planned. Clearly the study can answer very important questions that may have major implications for policy regarding social services and health care, especially for older citizens.

The study, however, has some very major limitations for measuring health effects: 1) primarily the absence of physiological measures of such things as blood pressure, lipids, bone density, body composition, blood-sugar, etc., 2) the history of disease being based on interview only, 3) lack of information on specific drug therapies and adherence and measurable changes in risk factors and 4) lack of measurement of preventive approaches.

The study, therefore, has very limited role in the study of the etiology or natural history of disease or in pathophysiology of risk factors as they relate to outcome. Since the beginning of the study in the early 1990s, there has been a dramatic increase in longitudinal studies of aging, many of which have been supported by the National Institutes on Aging. These studies provide much greater detail regarding etiology, natural history and studies of epidemiological risk factors. They, however, lack the important economic and social-behavioral variables of the HRS. Thus, it would be a big mistake for the HRS or the AHEAD Study to delve into the area of etiology, natural history and pathogenesis, given the weak data set.

The HRS and AHEAD Study have a unique role in evaluating the effects of changes in health as measured by function, such as Instrumental Activities of Daily Living (IADLs), Activities of Daily Living (ADLs), lower body measurements, cognition and sensory deficits as reported by the participants and the effects on

health services, retirement, economic factors, family services, community services, etc. The risk of disease is a function of both health risk factors and adherence to therapies that modify risk factors and host-susceptibility. The resulting impairments that lead to measured functional changes, as noted in the publications, are related to such things as family income, use of devices and other comorbidities. Thus, not all disease or any disability associated with disease clearly leads to the same functional changes or the functional changes lead to the same outcomes with regard to the use of health services and the impact on the family and the community. The absence of good measures of risk factors and adherence to therapy, especially risk factor changes, substantially limits the determinants of the functional changes. Thus, I think the study needs to focus on primarily longitudinal changes in functional changes including cognition, depression, vision, hearing, IADLs, ADLs, lower body function and incontinence and the subsequent effect on health services, retirement, the family, economic factors and survival.

It is a mistake for investigators to believe that the measurements of physiological functioning such as blood pressure, diabetes, arthritis and most of the diseases reported are objectively measured. The Commission on Chronic Illness in the 1950s clearly demonstrated the tremendous variation between reported history of disease or physiological functions and the actual measurements based on physical examination and many studies have since verified such findings. Furthermore, there are substantial biases since the reporting of disease is clearly related to the availability of health services, technology, the ability to make the disease diagnosis by a physician and the interpretation of the diagnosis by the patient or participant. Even reported diseases by a physician are inadequately reported and certainly there is bias in such reports.

The absence of good medical data, however, can in part be modulated by linkage with the Medicare data base and therefore it is very important to develop this linkage as best as possible as it will substantially enhance the value of the study. There are many investigators around the United States who have used the Medicare data base. For example, the studies from Northwestern University have been able to link the Medicare data base with the risk factor studies in Chicago and have provided important information about the relationship between risk factors and health and expenditures in the Medicare data base. Similarly, this study has an excellent opportunity of using the Medicare data base for both diagnoses and linking some of the important baseline and follow up information.

The measures of cognition and of depression can be greatly strengthened by next-of-kin interviews. This is especially true of the measurement of substantial cognitive change which will be referred to shortly. The absence of measurements of drug therapy is a mistake and every attempt should be made to see if this can be rectified in the future. A major issue, for example, is whether lower education or absences of financial resources results in underutilization of effective medical therapies, i.e. pharmacological therapies or lack of adherence

to such therapies. There is growing evidence that limitation of adherence to therapies contribute substantially to the educational and socioeconomic differences in health outcomes and, in fact, investigators have suggested that measurements of risk factors and adherence to risk factor modifications and therapy explains most, if not all, of the educational variations. I would, therefore, recommend attempts to collect information about pharmacological therapies and other therapies. This is especially important in AHEAD Study. I believe it would be important to make a major effort to collect drug therapy, even if it means eliminating some other less important information.

A very important area that is not covered very well is the use of preventive services. Medicare now covers some of these preventive services. It is possible that absence of insurance (medigap) and/or education has an important impact on use of preventive services. The lack of use of preventive services increases risk of disease, disability and functional changes.

The Telephone Interview for Cognitive Status (TICS) measurement is probably the best measure of cognition available for this type of a study, i.e. telephone and home interviews. The important issue is being able to measure changes in the TICS score and not the cross-sectional data. It is, therefore, critical that interviews be completed for as many of the participants as possible and especially those who are showing declines in their TICS and for nonrespondents, an all-out effort should be made to obtain data from the proxies. This is also true for deaths between each survey examination and this is especially important in the AHEAD Study.

It is unclear about what instruments are being used for the proxy interviews. The Dementia Questionnaire (DQ) and the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) are both very good instruments and would then be comparable with other studies. The measurement of cognitive decline at high level of functioning is difficult, especially when it is limited to the TICS. However, we and others have found that even small changes in such measures as the 100-point 3-Mini-Mental State Examination (3MMSE) at the high end, i.e. a 5 point decline from 95-90 or from 100 to 95 over a one or two year period is a strong predictor of subsequent dementia. I doubt whether a telephone-based type of executive functioning test will add dramatically to the study. Rather, it would seem important to make certain that the TICS data is being collected as adequately and completely as possible.

The key to this study is longitudinal data. Therefore, it would be important to monitor the trajectory of the decline in TICS over time and its potential relationship to other variables. For example, are small changes in the TICS even in the high level group predictors of apparent clinical dementia as measured through hospitalizations for nursing homes, from family interviews or from other secondary sources of data? Are they also predictors of mortality? We and others have shown that they predict changes in driving habits and also perhaps

increase in frequency of accidents. Furthermore, these early changes in the TICS may have important implications in the changes in the social environment of the individual such as social contacts, functioning in the community and even their relationship to family members. It would be important to determine whether these very early changes in the TICS and modest changes in cognition are a driving force that is determining the changes in the interpersonal relationships of older individuals in relationship to their peers and family as has been suggested in several studies.

There is obviously a substantial problem in measuring cognition in relationship to education. There is a striking association between education and almost every other cognitive tests that are done. However, the decline in cognitive scores is still an important determinant and is the key, rather than the cross-sectional measurement. There is a statement in a paper by Freedman, et al., "Aggregate changes in severe cognitive impairment," *Journal of Gerontology*, 2001 that approximately 60% of the individuals classified with severe cognitive dysfunction on the TICS in one year recovered between exams. This is inconceivable and needs to be reviewed. If there is such a substantial change, then clearly there is a serious problem with the way the TICS is being used. Individuals with severe cognitive loss, i.e. change over time, do not recover or have only a small recovery in spite of various drug therapies. Thus, if this is the case, it may be that the classification of severe cognitive dysfunction is questionable, that the TICS was not measured properly, i.e. was probably monitoring depression or problems with special senses such as hearing or lack of concentration or participation in the interview. The measurement of cognition is clearly very important and there should be a careful evaluation of the longitudinal changes in the TICS and what they mean. Furthermore, it may be well worthwhile to consider expanding the TICS to the full battery so that it could be useful compared with other studies. It is certainly not a perfect instrument but may be the best that can be used and does provide useful information and correlate fairly well with the 3MMSE. Other tests of cognition could be added to the study but, again, they would have to be done by telephone or a limited number done by home interview and given the tremendous amount of evolving studies on dementia and cognition and the addition which includes such measures of magnetic resonance imaging (MRI) and genetic markers and now even addition of biochemical markers, it is highly unlikely that either the AHEAD Study or the HRS is going to make any major contribution to understanding the etiology and natural history of dementia.

The measurement of depression based on the Centers for Epidemiological Studies-Depression (CES-D) scale, as given, does not appear to be very good. Depression is a major problem in the elderly, especially with the highest rates of suicide among older men. There is growing interest in the whole issue of the ability to identify depression in older people and, I believe, it would be very worthwhile to review the measurement of depression to look at the consistency of the measurement over time and whether the use of the depression scale as

provide in the study in any way has any external validity. Thus, it would be possible to evaluate the CES-D measurements as given in comparison to hospitalizations for depressions using the Medicare data base or to suicide and also the consistency of the measurements over time. It is truly unfortunate that there are no data on the use of various types of drug therapies in relationship to depression and also about physician-diagnosed depression and how this related to both treatment and follow up. This is an extremely important issue and may well affect many of the other variables that are being included in the study.

Heart failure is the leading cause of hospitalization among the elderly and is a major cause of disability and mortality and has a huge effect on hospitalization rates, physicians' visits and on health care costs. The relatively simple questionnaires available to measure disability and functional status associated with congestive heart failure (CHF) and some of these standardized questionnaires might be considered for the study. The Medicare data base again will be of great value in terms of looking at the issue of CHF and hospitalizations for CHF and, in fact, hospitalizations and treatment for cardiovascular disease is the major driving force for cost among the elderly in the United States and for out-of-pocket expenses. Thus, somewhat greater emphasis should be made to try to evaluate the functional status and changes related to CHF.

Stroke again is a major cause of disability and has a unique implication in terms of functional loss and disability since it results in an acute and dramatic reduction in functional status with often minimal recovery afterwards as opposed to coronary artery disease, in which the changes in function status are much less dramatic and recovery is much more substantial. There are three aspects of stroke that would be worth exploring: 1) attempt to try and measure incident stroke, i.e. new strokes and perhaps linking these with the Medicare records, 2) further evaluate the disability and functional changes after a stroke and how this relates to both health services and use of nursing home facilities, home health services, family caregivers and 3) relationship of stroke to cognitive decline after a stroke. About 25-30% of incident stroke cases develop dementia usually within 12-24 months after the stroke and this is rapidly becoming the most serious component of the stroke diagnosis. Stroke is much more common in the less-educated and in the minority populations and it may be very worthwhile to try and gain an understanding of the impact of stroke in less educated individuals and also individuals with less resources and how they deal with this specific problem. This may be an important issue.

There are papers talking about differences in arthritis in relationship to race and sex and educational factors and the uses of various types of therapies. Unfortunately, the measurement of arthritis in these studies is extremely limited. There is a great deal of arthritis, obviously, which is not diagnoses and many cases of arthritis which is a "waste paper diagnosis" by a physician or by a respondent who attributes any type of "pain in the legs or arms" to arthritis. Thus, arthritis is often considered a cause for a variety of functional changes and

disability which relate to problems to upper and lower body movement and functional changes. Some of this may be arthritis, probably a great deal of it, but some of it is due to other disease problems both muscle disease, neurological disease as well as disease in the joints. Thus, the focus on lower body functioning may in some ways may be a better approach than trying to understand arthritis.

I would focus on the study of weight loss in this cohort, especially the older individuals. Cachexia could be defined as a five or ten percent weight loss. The prevalence and outcome of cachexia could be an important addition to the study, especially in relationship to depression, cognition, dental health, IADLs, ADLs, education, social factors, family, housing, etc.

There has been considerable interest in the effects of education, social class in the HRS and also in the AHEAD Study. However, the papers are superficial, contain relatively little or no new information and certainly cannot get out the reasons for the educational differences in health, disease, utilization of health services and also trying to deal with the racial and ethnic differences. It is well known and has been recognized for generations that the educational, socioeconomic and racial variables are very important and must be included in much of the analysis. The study, however, lacks the strength to determine the reasons for these differences, i.e. related to risk factors, adherence to risk factor therapies, to host-genetic factors, to other environmental factors which may be adversely affecting the health of the community. It is important to look at education, race as well as sex as variables as they relate to the impact of functional changes, disability, cognitive change, etc. on retirement, economic factors, utilization of health services, etc.

Several other key points:

First, the publication record for the study to date is not very good especially with regard to health related areas. Many health related researchers in aging, in epidemiology or in preventive medicine or even in clinical medicine have no familiarity with these studies. There have been only a few papers in mainstream, high quality journals. The papers tend to be relatively superficial and descriptive such as papers on obesity and mortality, even on the relationship between health insurance and utilization of health services.

There is, unfortunately, only very limited information on the longitudinal analysis and no major linkages with the Medicare data base. Thus, the study investigators need to make a major effort to increase the utilization of this very important data set. I do not believe they can sit by and wait for "users." It also is extremely important that users of this data set, as a requirement, should get their papers published in peer-reviewed journals and not as reports in various conferences, symposium, etc. I believe it would be worthwhile to provide limited funding for investigators in the health field especially to generate specific

hypotheses using the data research analysis plans that could then be peer-reviewed especially using the longitudinal data and which a small stipend or grant would be provided for writing up the data and publication.

Second, there is also a need for better definition of testable hypotheses that will link the health economics, social-behavioral, health service utilization, etc. For example, do you see these early functional changes prior to the onset of any cognitive changes? This is an important question.

The longitudinal component of this study, as noted, is critical. Therefore, maximum resources should be committed to the success of the follow up study. The cross sectional data is interesting but is much less valuable.

The health measurements should focus primarily on measures of functional status and changes in functional status and less so on specific disease which are hard to measure or on physiological measures such as blood pressures, which is not measured. There needs, however, to be a major effort to link the Medicare files with this important data set. Both Part A and Part B of Medicare should be utilized and further methods be used to deal with the large numbers of individuals who may be in Medicare HMOs – how to obtain their data base, if possible. As noted, it would be useful to review the Commission on Chronic Illness Report from the 1950-1960s which studied very large samples in Baltimore and Hunterton County, New Jersey and evaluated specific medical examinations, interview data and screening of risk factors. As noted, I believe it is very unfortunate that the study lacks data on drug therapies and adherence to these therapies. For example, loss of a job may result in depression, decreased adherence to therapies such as for blood pressure, diabetes, hyperlipidemia, etc. and increased risks of clinical disease. However, stress is related to loss is a job, so-called allostatic load could increase the risk of disease given a selected blood pressure level, diabetes, etc. or third, that drugs given for the various treatments of risk factors could modulate behavior resulting in depression or in slowing of activities and subsequent loss of job which might appear to the subject to be involuntary but are related to the lack of productivity of the worker.

In spite of these limitations, the study of functional changes as reported by the respondents over time, the impact on the use health services and occupational changes on the family and retirement and the use of economic resources are still a very important issue that can be studied very effectively by the HRS and AHEAD Study and their related projects. The limitations of the self-reported data should be recognized by the investigators who should shy away from the physiological measurements and focus much more on functional status and disability with regards to health issues. Disease information is of limited value especially now with the advent of very sophisticated medical technology and the variation of which may be a major determinant of the prevalence of specific diseases. This study will have a major advance when it can provide both high quality longitudinal information and linkages with the Medicare data base.

Furthermore, as recommended, there is a need for measurement of pharmacological and nonpharmacological therapies and at least some crude estimate of adherence to such therapies and reasons for lack of adherence.