Using the Health and Retirement Survey to Investigate Health Disparities

Mark D. Hayward[1]
Professor of Sociology and Demography
The Pennsylvania State University

[1] The views expressed in this document are solely those of the author and do not necessarily reflect the opinions of the staff of the Health and Retirement Survey or the National Institute on Aging. The author can be contacted at: 601 Oswald Tower, Population Research Institute, The Pennsylvania State University, University Park, PA 16802. E-mail: Hayward@pop.psu.edu.

Executive Summary

The Health and Retirement Survey (HRS) is currently the foremost data vehicle to assess health disparities in the United States. A number of factors contribute to the HRS' advantages in tackling health disparities.

1. The HRS includes measures of the major domains of physical health (major fatal and nonfatal chronic diseases, functional problems, and disability) mortality, and affective and cognitive functioning. The HRS' coverage of both morbidity and mortality provides the means to assess disparities in the burden of major chronic diseases – differences in the basic functional problems and disability associated with chronic disease, differences in the length of life spent with chronic health conditions, and differences in the mortality consequences of chronic health conditions.
2. Unlike many community-based or clinical studies, the HRS is a nationally representative sample and a long-running panel – now over 10 years. Longitudinal data are essential for studies whose intent is to examine race/ethnic and socioeconomic differences in the development and consequences of chronic health problems. Such information cannot be supplied by either a localized or selective sample. Moreover, the length of the panel permits analysts to document disparities in the major health changes and mortality associated with aging.
3. The HRS begins at age 51 and goes through the end of the life span. Having a sample that begins in midlife is essential to investigate disparities in changes in health associated with aging and even a younger starting age would be better. The decade of the life beginning at age 50 is the one when the rates of onset of chronic health problems begin to rise. It is also the decade when socioeconomic and race/ethnic groups differentially lose people to disease and death. To assess how disparities in chronic health conditions unfold over the life cycle, it is thus important to begin observation at a young age to capture as many health events as possible.
4. The HRS has more information than any other survey on lifetime socioeconomic experiences and more complete data on social and economic circumstances for the survey period than any other nationally representative data set. A survey such as the HRS that collects information on all of the major health outcomes of old age, lifetime SES characteristics, and information about the mechanisms by which SES might work is essential to address questions about the nature and magnitude of health disparities in the population and questions about causality and mechanisms.
5. No other national survey provides the *combination* of attributes that are so important for assessing health disparities – the breadth of coverage of core health concepts, a long-running panel, the inclusion of persons from roughly 50 years of age until advanced old age, and key measures defining the nature of disparities in population health.

Although the HRS has numerous advantages, analysts using these data to examine health disparities should be cognizant of a number of issues.

1. Although the design is clearly is a longitudinal panel, changes in the health measures prior to 1998 and the 2-year observation window present methodological challenges in assessing health disparities.
2. The epidemiological and medical communities have criticized HRS' ability to document health disparities because of the absence of performance measures of functional ability or laboratory tests of chronic disease. This review examines the available evidence and offers some suggestions for interpreting the self-report measures.
3. Coverage of key concepts – particularly measures pertaining to childhood family structure during childhood (e.g., whether the respondent was reared primarily in a two (biological) parent family) and adult family relationships (e.g., marital quality and conflict) in adulthood – constrains investigators' ability to fully delineate both how these factors come into play in affecting health disparities and the net effects of other social conditions such as household wealth.
4. Frequent updates of the data due to errors frustrate analysts and are costly to the research enterprise. Similarly, variable naming conventions that result in changes across waves hinder efficient use by the research community. The RAND longitudinal file is an important contribution to the research enterprise, and updated documentation should be encouraged as future HRS waves become available.

Introductory Remarks

In this review, I focus exclusively on the strengths and weaknesses of the HRS to address the general research problem of health disparities. In doing so, I emphasize the coverage and quality of a set of core health measures that are measured longitudinally, design issues, and the inclusion of measures necessary to document the extent of major chronic health disparities in the American population.

Health Measures: Concept Coverage and Data Quality

Coverage

In assessing the health of the older population, the HRS relies on self-report measures of general health status, physical health, affective and cognitive functioning, and health risk behaviors. The breadth of chronic health conditions encompassed in these measures is a key element defining the HRS' value in documenting disparities in the major domains of population health. The HRS also tracks respondents' mortality, and HRS staff have begun to institutionalize the submission of respondent records to the National Death Index (NDI) for matching in order to more carefully identify the timing of death. The NDI also provides information to the HRS about the primary and tertiary causes of death, although this information is available to analysts only as a restricted data file. The addition of NDI information is a significant step in enhancing the value of the HRS as a vehicle to investigate health disparities defined by mortality experience. HRS' plans for the timely addition of NDI data to the dataset should be given a high priority.

The HRS's ability to document health disparities is also enhanced by the fact that it includes longitudinal information on *both* morbidity and mortality for a long-running panel – 10 years thus far. Although chronic-disease based morbidity and mortality are related, they are not isomorphic concepts. Morbidity and mortality intersect to define life-cycle health experiences such as:

- The length of life with and without major chronic health conditions
- The years of life lost due to particular health conditions (e.g., heart disease)
- The structure (i.e., age-specific prevalence) of health problems across major population subgroups
- The timing and nature of major health events leading to subsequent health problems and death

This type of information is important to researchers interested in documenting disparities in life-cycle health, the different ways that mortality and morbidity may intersect in leading to these disparities, and disparities in the burden of chronic health problems. This information is also useful to policy makers evaluating the potential implications at the population level and for particular population groups of changes in chronic disease experience, functional problems, and mortality.

The bulk of the HRS health measures refers to the major domains of physical health (i.e., major fatal and nonfatal chronic diseases, functional problems, and disability)

and mortality rather than affective and cognitive functioning. The physical measures correspond roughly to health concepts from the World Health Organization's classification scheme (the 1980 *International Classification of Impairments, Disabilities, and Handicaps* (ICIDH) (World Health Organization 1980) and Nagi's scheme (Nagi 1989, 1991; Verbrugge and Jette 1994) adopted by the Institute of Medicine.[2] Conceptually, the WHO and Nagi schemes overlap significantly (Verbrugge and Jette 1994). The ICIDH and Nagi's schemes are the basis of Verbrugge and Jette's (Verbrugge and Jette 1994) disablement framework. Because of the combination of long-running panel and the HRS' coverage of the major domains of health (and mortality), the HRS allows investigators to identify the onset of (diagnosed) major chronic diseases, functional problems associated with diseases, the possibility of the onset of co-morbid chronic disease conditions, changes in functioning, and whether an individual's health problems curtail major domestic and work activities (i.e., disability). All of these are important facets in evaluating health disparities in the population.

The HRS does not measure mental health as comprehensively as it does physical health, and the HRS presents justification in a documentation report of the affective health measures (e.g., affective functioning is not a well defined or easily quantifiable concept; many psychological concepts are hard to measure in a survey). The documentation report, ***Documentation of Affective Functioning Measures in the Health and Retirement Study***, is a valuable resource for researchers interested in examining health disparities in terms of affective function. A similar report is available for cognitive functioning.

Affective functioning is measured by a global measure of emotional health, whether a doctor has given the respondent a diagnosis of an emotional or nervous problem, and a modified CESD scale. More detailed information on affective functioning, provided by WHO's Composite International Diagnostic Interview - Short Form, is obtained once in 1996 for the HRS cohort and 1995 for the original AHEAD cohort. New cohorts added to the HRS also will receive the CIDI-SF during the baseline interview but not thereafter. Cognitive functioning is measured along a number of dimensions – memory (self-rated, recall, and working memory) and mental status (knowledge, language and orientation). Although the incorporation of cognitive functioning measures in a population survey is relatively novel, and some researchers might debate the relative merits of particular measures, the inclusion of a core set of

[2] The World Health Organization introduced the *International Classification of Functioning, Disability, and Health* (ICF) in 2001 World Health Organization. 2001. *International Classification of Functioning, Disability, and Health*. Geneva.. The ICF's changes in terminology make it difficult to explicitly compare to Nagi's scheme and Verbrugge and Jette's disablement process, although the ICF's conceptual framework embraces Verbrugge and Jette's ideas of how extra- and intra-individual factors influence environmental demands and individuals' capabilities. The ICF uses two umbrella terms, functioning and disability. Functioning encompasses body functions, activities executed by the individual, and participation in a life situation. Disability refers to the impairment of physiological functions, organ systems, activity limitation and participation restriction.

measures represents an important step in deriving a more global assessment of health disparities in the population.[3]

A strength of the HRS is its ability to track at least some aspects of affective and cognitive functioning, however limited, over time and evaluate how these types of functioning are associated with physical health. For example, declines in cognitive functioning can induce problems in IADL disability. Depressive symptoms are both a risk factor for subsequent declines in health and outcomes of a disease process curtailing normal activities. How these types of associations play out over the long-term is not well understood nor is it clear whether the associations hold across the major race/ethnic groups or across groups anchoring the ends of the socioeconomic ladder. Similarly, it is an open question whether socioeconomic and race groups disadvantaged in one aspect of health are disadvantaged in other aspects. Having measures of affective and cognitive functioning thus allow for a more nuanced analysis of health disparities.

In summary, the HRS does a solid job of covering the major domains of health – this is important for developing a refined and multidimensional picture of health disparities in the population. Although researchers might debate the relative merits of measuring particular concepts or they might desire more detailed information on some aspects of health, general concept coverage is relatively complete – certainly more so than any other major longitudinal survey.

Future extensions in concept coverage? One issue that is not well addressed by the HRS – or any other longitudinal survey -- is that some of the traditional health measures – particularly measures of disability (e.g., ADLs, IADLs, and work and domestic disability) do not necessarily translate well in a conceptual sense to a longitudinal design or longitudinal analysis (Crimmins and Hayward 1997). This is a problem induced not by the exclusion of a major health problem in the survey but by the survey's longitudinal nature. At any particular survey wave, for example, disability is not exclusively a biomedical process but an outcome of dysfunction (organ system or bodily function) and the environmental demands on functioning. Disparities and changes in disability potentially reflect differences in the level of social and environmental support as well as physical or mental functioning. For example, marital status changes or the addition of technology can result in changes in individuals reporting that they need or get less help with tasks without improvements in the underlying biomedical process. Differentiating functional changes from environmental changes is thus important to identify disparities in the onset and "recovery" from disability.

I am *not* proposing changes in the current set of disability measures (these are a large fraction of the total number of physical health measures). To do so would seriously undermine the longitudinal integrity of the HRS. I am recommending, however, that the survey staff seriously consider a couple of options. First, survey staff, in consultation

[3] I am not sufficiently knowledgeable to comment on the issue of executive function. In the ideal sense, it is desirable to incorporate measures that are sufficiently sensitive to pick up differences in highly functional persons – particularly if these predict later declines. I suspect, however, that the reliability of these measures will be low – particularly if measures used in an experimental setting are adapted for survey use.

with a *team* of scientific experts, should develop a module aimed at assessing how dysfunctional changes combine with environmental changes to affect reports of disability. Second, presuming that these changes can be measured in a straightforward fashion using 2 or 3 items, survey staff should consider incorporating these measures into future waves of the HRS. Recently, a team of population health experts met at the National Center for Health Statistics and noted that interpreting changes in disability is a major problem in evaluating trends and disparities in population health (Progress in Estimating Active Life Expectancy, National Center for Health Statistics, October 9, 2002).

Data Quality

Presently, the HRS staff is releasing (or is in the process of releasing) documentation that evaluates the quality of the health measures in the HRS. Documentation reports are currently available for affective and cognitive functioning and a report is planned for physical health. These reports are valuable to the scientific community, and HRS staff should make these reports a high priority.

Reviewers on the health measures have been asked to comment on the tendency for the biomedical community to be critical of the self-report measures of health in the HRS. My personal experience validates this tendency. Although I cannot speak directly to the prevalence of this stance in the biomedical community, I have observed these criticisms on NIH study section (SNEM-3), among my biomedical colleagues, and even among NCHS staff. The common thread is that performance measures of functioning or laboratory tests of chronic disease (e.g., hypertension, diabetes II) are inherently preferable to self-report measures. Although it is difficult to interpret the snippets of comments, I suspect that much of the criticism is rooted in disciplinary culture.

Although there are a number of reasons to identify clinical chronic disease (or performance measures of functioning) in a population (e.g., unmet need for treatment, more effective targeting of preventing care, and screening), self-report information is also valuable. Individuals make decisions in their lives about retirement, health care, and so on based on recognized symptoms or diagnoses. I do not think that the social science community or the HRS has done an effective job in identifying the advantages of self-report measures of population health. Communication efforts in this vein, I think, would certainly help to balance and perhaps better contextualize critical comments by some in the biomedical community. Moreover, in developing a response to this criticism, I think it is essential to engage social scientists and biomedical researchers in a dialogue to establish a bridge between self-report and biomedical measures of health. Getting buy-in on the self-report measures by a core group of biomedical researchers would help to counteract criticism.

In determining the context of the HRS self-report measures of health, I think the HRS staff adopted a defensible position. The HRS chronic disease items, for example, were selected on the basis of several important properties: reliability based on past studies, comparability to other major health surveys such as the National Health Interview Survey, diseases with the highest prevalence and those associated with high medical costs

and caregiver burden (Myers, Juster and Suzman 1997). The National Health Interview Survey, as well as the NHANES, uses these items to monitor population changes in major disease morbidity prevalence and to gauge socioeconomic and race/ethnic disparities in health prevalence. These estimates have fueled the development of *Healthy People 2000* and *Healthy People 2010*.

Comparability in these items across the national surveys provides the HRS with an important benchmark to assess the reliability of HRS estimates, and it is important that the forthcoming documentation report make such a comparison. In doing so, I also recommend a careful evaluation of the sample composition of the various surveys. For example, how do the socioeconomic characteristics of Hispanics or African Americans in the HRS compare to minority group SES in the NHIS or NHANES? The same sort of evaluation should be done for disparate SES groups across the surveys. Sample differences could potentially distort the benchmarking of the HRS estimates to surveys such as the NHIS and NHANES. A careful evaluation of reliability will contribute significantly to making a case for the HRS self-report measures of health.

The design of the HRS does not allow the evaluation of systematic mis-reporting (over or under) of health conditions measured by the HRS. In general, my reading of the scientific literature suggests that self-reports of chronic disease (has a doctor ever told you …?) in population surveys of health are reasonably accurate, i.e., the correspondence between self-reports that a doctor has told the respondent that they have a disease and physician records are high. Accuracy varies somewhat with the type of disease, e.g., the accuracy of self-reports of cancer and diabetes typically exceed 90-95% with slightly less accuracy for cardiovascular disease. In using any of these conditions, it is important to be mindful that whether a respondent reports a doctor's diagnosis is related to health care and the severity of the disease symptoms. These factors suggest that self reports are likely to lead to the under-reporting of *disease* conditions in the population -- persons who are infrequent users of health care and/or who display mild symptoms are least likely to report a doctor's diagnosis. Self-reports of chronic disease in population health surveys are thus likely to represent a lower bound of (clinical) disease prevalence in the population. This point is frequently lost in discussions of the relative merits of self-report measures.

An important issue with regard to measuring disparities in diseases via doctor diagnosis is whether groups with a lower level of health care use (e.g., groups lower in SES) are less likely to report a chronic condition. Scientific evidence on this point is limited. What little evidence exists, however, suggests that the race/ethnic patterns (i.e., relative differences) observed in the self-report data are consistent with clinical measurement. For example, Mitchell et al. (Mitchell et al. 1991), using clinical measures, observed that Mexican American men's prevalence of myocardial infarction was significantly lower than that for white non-Hispanic men. Mexican American women, however, did not differ significantly from white non-Hispanic women in their prevalence of myocardial infarction. This pattern, although not the levels, is similar to that reported for ethnic differences in heart disease by Kington and Smith (1997) using the HRS/AHEAD. There also is some evidence that validity of self-reported

hypertension is relatively high for blacks and whites (Giles, Croft and Keenan 1995; Vargas et al. 1997), although the validity of hypertension reporting among Hispanics has been questioned (Vargas et al. 1997).[4] Studies based on NHANES II indicate that blacks and whites in the HRS age range with clinical hypertension are equally likely to report having hypertension (Andersen, Mullner and Cornelius 1987; Drizd, Dannenberg and Engel 1986; Hadden and Harris 1987). Among persons with clinical diabetes, blacks are slightly less likely to report diabetes than whites. Bound et al. (1995) performed similar analyses for educational groups and observed a similar pattern. Overall, the empirical evidence consistently points to a pattern of under-reporting that either accurately reflects race/ethnic disparities in clinical disease or slightly understates race/ethnic disparities.

Clearly, there is insufficient evidence at this juncture to make conclusive statements about the reliability and validity of the self-report health data. However, given the potential importance of the HRS in documenting health disparities, I recommend that the HRS staff pursue two options:

1. Conduct a validation study using Medicare administrative records: Medicare data are able to provide detailed information about dates of service, diagnosis (ICD-9) and procedures. Because Medicare information is only available for a subset of HRS respondents (i.e., persons 65 years of age and older who do not participate in HMOs), an analysis will necessarily be restricted to persons aged 65 years and older. At a minimum, the Medicare data offer an opportunity to carry out a validation study of the self-reports of chronic disease with the Medicare diagnostic information and the degree to which validity varies across the major race/ethnic and socioeconomic status groups. Note, however, that this type of validation study does not capture the degree to which persons with clinical disease report having a disease. Rather, it refers to the accuracy of reports of whether "a doctor has told the R that they have a disease." A validation study of the overall degree of disease underreporting requires self-reports of disease and clinical measures for a population. The NHANES IV is likely to be an important dataset to support this sort of validation study.
2. Conduct a reliability study to assess comparing HRS prevalence rates of the major chronic conditions with other nationally representative health surveys. This type of benchmarking activity should be sensitive to possible survey differences in composition within the major race/ethnic groups and socioeconomic categories and design differences across surveys.
3. Involve biomedical researchers in a dialog, perhaps has co-investigators in these studies, to develop a bridge to the biomedical community.

[4] The reliability of self-reported chronic disease using the HRS appears to be quite high. Race/ethnic reporting of heart disease in the HRS, for example, is highly consistent with self-reports of heart disease based on the National Health Interview Survey National Center for Health Statistics. 1999. *Health, United States, 1999*. Hyattsville, MD..

HRS Design Implications for Assessing Health Disparities

The HRS design offers a number of unique benefits for capturing health disparities. I will enumerate these briefly below.

- The HRS is a very large nationally representative survey, combined with over-samples of African Americans and Hispanics. These features allow analysts to assess group differences in health as well as conduct within-group analyses (e.g., within race/ethnic group) of health disparities. An example of the latter type of analysis is whether the SES gradient in health is comparable across major race/ethnic groups. This type of study cannot be conducted using community-based or clinical studies.
- The longitudinal design of the HRS is essential to examine population disparities in the development and consequences of chronic health problems. This type of information cannot be obtained in localized or selective samples. The fact that the panel is long-running (and will continue to be long-running) also offers opportunities to conduct research on disparities in the onset of major diseases, subsequent co-morbidity, functional changes and disability associated with disease, and mortal outcomes.
- The HRS begins at age 51 and includes respondents through the end of the life span. Many datasets on older populations have a lower age-eligibility criterion of age 65 or 70. Health disparities, however, refer to *differences* in the age-related health experiences, i.e., some groups may experience health problems at significantly younger ages than other groups. Those groups more prone to experience health problems at younger ages are also likely to experience higher rates of mortality, leaving behind an increasingly select population at the older ages. Having a sample in midlife, such as the HRS, thus provides a means to assess health disparities which potentially occur relatively early in the life cycle. Indeed, taking a casual look at the U.S. Vital Statistics for persons aged 45-54 years of age shows rather remarkable levels of chronic disease deaths among African Americans (National Center for Health Statistics 1999) – particularly compared to whites – suggesting that including younger ages would be even better to capture disparities in chronic disease problems.

<u>Criticisms of the Longitudinal Design</u>

1. The 2-year observation interval: One criticism of the longitudinal design has to do with the 2-year intervals between interview waves. The 2-year observation interval, combined with the static nature of most of the health measures (e.g., functional status at the time of the interview), points to potential problems in identifying health changes that are unobserved in the interval. This has implications for examining health disparities to the extent that unobserved health experiences differentially occur for major population groups. For example, Laditka and Wolf argue that a significant number of disability transitions are missed in a 2-year observation interval (Laditka and Wolf 1998). Crimmins and Hayward note similarly that some fatal disease episodes ending in death progress rapidly (e.g., certain cancers, heart attack deaths) such these disease experiences are missed by a two-year observation window (Crimmins et al. 1999; Hayward,

Crimmins and Zhang 2000), e.g., a person may report having no fatal disease condition at baseline and then be dead before observation at the subsequent interview. Moreover, Hayward and Crimmins note that short fatal disease episodes appear to occur disproportionately among African Americans and Hispanics in the HRS sample, suggesting a potential distortion of health disparities if this is not taken into account (Hayward et al. 2000).

Laditka and Wolf (1998), as well as other demographers (e.g., Brouard), have developed new statistical methods based on embedded Markov chains as a methodological means to grapple with the problem of capturing unobserved disability events in the observation interval. Note, however, that the statistical approach presumes that the transitions are "real" – e.g., methodologists have not seriously questioned the interpretation of recovery from disability. As noted early in the report, changes in reports of disability status may reflect underlying changes in physical functioning, modifications in the environment, or both. I suggested strategies to improve the interpretation of the disability measures.

In grappling with the problem of unobserved fatal disease episodes, Hayward and Crimmins made use of the NDI cause-of-death data to identify unobserved disease experience among decedents. Causes of death were used to impute morbidity transitions during the observation interval. Crimmins' and Hayward's work demonstrates the utility of the cause-of-death data in refining the interpretation of the morbidity data. In addition, the Medicare administrative records could also be used in a similar fashion to identify disease experience within observation intervals for a significant subset of the HRS respondents.

2. Changes in measuring the core health concepts. The HRS is notorious in some scientific circles for changes in item wording, question format, and response categories for a number of health measures – particularly the ADLs, IADLs and Nagi functioning items. This problem principally refers to waves 1, 2 and 3. After wave 3, the measurement of the health items has been relatively stable.

It is imperative that the HRS refrain from future changes in the measurement of the health status items to preserve the longitudinal integrity of the measures. Longitudinal integrity is fundamental to minimizing measurement error in identifying health changes, and identifying health changes in the population are a major reason justifying the value of the HRS. Previous design changes have hampered researchers' use of the HRS to examine disparities in age-related health experiences, and they have spawned a cottage industry among HRS methodologists to create "cross-walks" between measures and complicated statistical algorithms to resolve problems introduced by changes in the health measures. The consequences of the changes in measurement have been expensive to say the least.

Although some item design changes were obviously necessitated by interview mode (e.g., in person vs. telephone), other reasons for the design changes are not transparent. In light of the expertise on the HRS staff, as well as on the HRS Monitoring

Committee, these design changes are surprising to say the least. Although it is tempting to assume that future design changes will be minimal based on recent history, the loss of Dr. Regula Herzog to the HRS health team cannot be underestimated. Replacing Dr. Herzog will undoubtedly be a difficult task. Without outstanding leadership in this area, the measurement of the core health concepts could drift again in the HRS. Members of the HRS Monitoring Committee should be more heavily involved in reviewing plans for measuring health in future waves of the HRS.

3. Population coverage: The sample size of the HRS and over-samples of African Americans and Hispanics are essential to evaluate health disparities defined by race/ethnicity and socioeconomic status. Nonetheless, it is not clear that the sample size of the Hispanic subsample is sufficient to identify potential within group disparities in health. For example, other national surveys point to significant health and mortality differences between Mexican Americans, Cubans, and Puerto Ricans. Complicating the within-Hispanic group comparison is that a significant portion of the Hispanic subsample are immigrants. Given the policy importance of the growing Mexican American population, the HRS may want to consider expanding the Hispanic subsample to increase the number of Mexican Americans (currently at about 65% of the Hispanic subsample).

Other important population subgroups are also not well represented in the HRS, particularly Asian Americans and Native Americans. The small number of respondents who report being Asian American or Native American necessarily reflects their small numbers in the total population. Nonetheless, these groups represent important policy targets – hence, the lack of group-level health information and within-group variation in health. Moreover, according to the U.S. Census Bureau, the number of elderly Asian Americans is projected to grow significantly over the next several decades and increase as a proportion of the total population. Without specific attention to Asian Americans, the HRS could potentially offer only a circumscribed view of future health disparities. The HRS staff should consider drawing over-samples for groups such as Asian Americans in order to monitor their health. Should the HRS move in this direction, an important issue is the recognition of Asian subgroups (e.g., Chinese, Samoans, etc.) in terms of sample size. Like Hispanics, there is considerable heterogeneity in health among Asians according to country of origin.

HRS Coverage of Concepts Defining Social and Economic Disparities in Health

Compared to other nationally representative longitudinal surveys, the HRS contains the most complete data on individuals' lifetime socioeconomic experiences and the most complete data on adult social and economic circumstances for a long-running panel. The HRS is especially well suited to investigate health disparities along the SES continuum and the bi-directional relationship between socioeconomic achievements and health over the life cycle that is associated with late-life health. The HRS was strengthened in these areas with the addition in 1998 of information on childhood socioeconomic conditions and health experiences.

The HRS is not strong, however, in investigating health disparities linked, for example, to subjective evaluations of economic conditions, and non-economic family and household conditions over the life cycle. I am *not* suggesting that the HRS staff introduce such measures on an *ad hoc* basis or based on my own professional opinion about the desirability of adding these concepts. This would lead to a survey spinning out of control. I am suggesting, however, that should a body of scientific evidence point in the direction of adding new items to capture concepts important to understanding health disparities, that a systematic and scientific evaluation be conducted prior to redesigning the survey. Moreover, such an evaluation should consistently be done by a team of researchers – not a single investigator -- to ensure consensus on the incorporation of the theoretically relevant measures. I will use two cases to illustrate this point.

1. Case 1: the incorporation of childhood conditions. As noted previously, the HRS was strengthened by the addition of childhood measures of SES and health experiences. Unclear, however, is whether other theoretically relevant childhood measures should or could have been included. Given that this is a burgeoning area of research, it is paramount that the right set of measures be included to maximize the usefulness of the HRS redesign.

As far as I can tell from the HRS website, the decision regarding what measures to include was based on the analysis of the 1996 HRS topical module. The results of this analysis are described in a single working paper by Irma Elo. Although I am NOT questioning the quality of Elo's work, her paper is necessarily circumscribed by its attention to only a subset of health outcomes, the measures chosen to define childhood conditions, and model specification. No single paper can address all the relevant issues – nor should it. A particularly glaring omission in the 1998 HRS childhood measures, for example, is the absence of information about childhood family structure. A growing body of research using European and American data points to strong associations between childhood family structure and late-life health, net of lifecycle SES. Other research in family sociology and human development documents the long-term consequences for health and economic achievement of marital quality and conflict in childhood (highly stressful conditions). My basic point is that involving a broader range of scholars in making redesign changes could (hopefully) minimize the omission of key measures.

2. Case 2: marriage and health. The scientific literature has long shown a strong association between marriage and mortality (Kitagawa and Hauser 1973; Lillard and Panis 1996; Lillard and Waite 1995), and a recent study by Pienta using HRS data reports similar associations between marriage and a range of chronic health conditions (Pienta, Hayward and Jenkins 2000). Generally, the empirical associations indicate that married persons are healthier along a number of dimensions compared to divorced and widowed persons. Never married persons, a growing segment of the older population, appear to be more like married persons in terms of health, but the literature provides only a preliminary cut on this issue. Marital status effects typically remain even after controlling for factors such as socioeconomic status and lifestyle behaviors, pointing to its importance in contributing to health disparities in the population. A range of

competing interpretations has been advanced to explain these associations – assortative mating, marital investment, stresses associated with divorce and widowhood, and selection of less healthy persons into divorce and widowhood. The scientific literature, as yet, offers little conclusive evidence on the role of these mechanisms in accounting for the association between marital status and health.

Because the HRS is a long-running panel, the HRS offers a unique opportunity to examine the associations between marital status, changes in marital status, chronic health conditions, and mortality. Less clear, however, is the capability of the HRS to support research on the mechanisms underlying health disparities defined by marriage, divorce, and widowhood. I am *NOT* recommending that the HRS should move in this direction. No study can be all encompassing. However, as researchers increasingly recognize the importance of marriage to understanding health disparities, the HRS staff may ultimately find themselves considering redesign issues that respond to this general research question. Moreover, it is likely that potential shortcomings of the SES paradigm to account for race/ethnic disparities, for example, may motivate researchers to pay closer attention to the role of marriage in defining health disparities.

Should a redesign of this sort be considered, I think it is essential to form a team of experts to assist in the redesign. The issues are complex and interdisciplinary – economists, family sociologists, and developmental psychologists all contribute to a growing literature in the area. Any redesign should incorporate a wide range of theoretically important measures within the constraints of time and space of the overall survey. I have used a specific example to make this general point, but the point needs to be emphasized to help ensure the highest quality of data available to the general scientific community.

<u>Adding geo-coded information to assess health disparities</u>: I think there is potentially great scientific value in geo-coding the residence of respondents and potentially the residences of kin and health care providers. Geo-coding this information, particularly as it might change across survey waves, would provide researchers with new tools to identify how older persons navigate their social environment to obtain health care, how the spatial environment influences older persons' social support networks, and the role that the special environment plays in determining the nature and volume of family support.

I am very much in favor of geo-coding as much information about respondents' social work and activities as much as possible (e.g., respondents' residences and places of work, locations of health care providers, residences of kin and close friends). I think this should be a high priority, since this sort of geo-coding would allow for the address-matching of a wide range of extant data. The value of the HRS staff creating a file is less clear to me (e.g., a file based on census data indicators for some particular geographic unit). In part, I am less enthusiastic about this activity because the address-matching capabilities of a range of software packages make this a relatively easy task for individual investigators. In part, my lack of enthusiasm stems from the fact that it is not clear which census measures should be included on a geo-coded file and the appropriate

geographic unit. Generally the larger the geographic unit, the less proximate the place characteristics are to the individual respondent, and the less value this is to the research activity.

In summary, I see significant value in geo-coding information about respondents' residences, health care provides, and friends and family. This would provide the research community with the flexibility it needs to address a wide range of scientific issues revolving around the spatial environment and health. I would place a lower priority on the HRS staff creating a data file containing geographic measures whose links to health disparities are not well articulated either theoretically or empirically.

Summary

In my opinion, the HRS is the best available dataset to assess health disparities at the population level. The advantages of the HRS stem from a combination of key factors – concept coverage, sample size and representativeness, a long-running panel, the inclusion of a relatively young population group, and a rich array of measures of lifetime socioeconomic conditions. No other dataset offers all of these advantages.

I also have noted, however, that the HRS can be improved in a number of ways that would make it even more valuable as an investigative tool to examine health disparities. I have offered specific suggestions about clarifying the interpretation of some of the health measures (particularly disability and functioning), assessing the reliability and validity of the self-report health measures, identifying changes in health during the 2-year observation windows, expanding the race/ethnic subsamples, and evaluating strategies for including additional concepts defining social and economic disparities in health. Most, if not all of my suggestions, avoid tampering with the current measures in the HRS. Preserving the longitudinal integrity of the HRS should be a top priority.

REFERENCES

Andersen, R.M., R.M. Mullner, and L.J. Cornelius. 1987. "Black-White Differences in Health Status: Methods or Substance?" *Milbank Quarterly* 65(Suppl 1):72-99.

Bound, J., M. Schoenbaum, and T.A. Waidmann. 1995. "Race and Education Differences in Disability Status and Labor Force Attachment in the Health and Retirement Study." *The Journal of Human Resources* 30(Supplement):S227-S267.

Crimmins, E.M.and M.D. Hayward. 1997. "What Can We Learn about Competence at Older Ages from Active Life Expectancy?" Pp. 1-22 in *Social Structural Mechanisms for Maintaining Competence in Old Age*, edited by K.W. Schaie, S. Willis, and M.D. Hayward. New York: Springer.

Crimmins, E.M., M.D. Hayward, H. Ueda, and Y. Saito. 1999. "Life With and Without Heart Disease Among Men and Women Over 50." Paper presented at the Annual Meetings of the Gerontological Society of America, San Francisco, CA.

Drizd, T., A.L. Dannenberg, and A. Engel. 1986. "Blood Pressure Levels in Persons 18-74 years of Age in 1976-80, and Trends in Blood Pressure from 1960 to 1980 in the United States." *Vital and Health Statistics* 11:1-68.

Giles, W., J.B. Croft, and N.L. Keenan. 1995. "The Validity of Self-reported Hypertension and Correlates of Hypertension Awareness Among Blacks and Whites Within the Stroke Belt." *American Journal of Preventive Medicine* 11:163-169.

Hadden, W.C.and M.I. Harris. 1987. "Prevalence of diagnosed diabetes, undiagnosed diabetes, and impaired glucose tolerance in adults 20-74 years of age." *Vital and Health Statistics* 11(237):1-55.

Hayward, M.D., E.M. Crimmins, and Z. Zhang. 2000. "The Racial Burden of Chronic Disease: A Life Cycle Model of Heart Disease Experience." Paper presented at the Annual Meetings of the Population Association of America, Los Angeles, CA.

Kitagawa, E.M.and P.M. Hauser. 1973. *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology*. Cambridge, MA: Harvard University Press.

Laditka, S.B.and D.A. Wolf. 1998. "New Methods for Analyzing Active Life Expectancy." *Journal of Aging and Health* 10:214-241.

Lillard, L.A.and C.W.A. Panis. 1996. "Marital Status and Mortality: The Role of Health." *Demography* 33:313-327.

Lillard, L.A.and L.J. Waite. 1995. "'Til Death Do Us Part: Marital Disruption and Mortality." *American Journal of Sociology* 100:1131-1156.

Mitchell, B.D., H.P. Hazuda, S.M. Haffner, J.K. Patterson, and M.P. Stern. 1991. "Myocardial Infarction in Mexican-Americans and Non-Hispanic Whites. The San Antonio Heart Study." *Circulation* 83:45-51.

Myers, G.C., F.T. Juster, and R.M. Suzman. 1997. "Introduction." *The Journals of Gerontology* 52B(Special Issue):v-viii.

Nagi, S.Z. 1989. "The Concept and Measurement of Disability." Pp. 1-15 in *Disability Policies and Government Programs*, edited by E.D. Berkowitz. New York: Praeger.

—. 1991. "Disability Concepts Revisited: Implications for Prevention." Pp. 309-327 in *Disability in America: Toward a National Agenda for Prevention*, edited by A.M. Pope and A.R. Tarlov. Washington, DC: Institute of Medicine, National Academy Press.

National Center for Health Statistics. 1999. *Health, United States, 1999*. Hyattsville, MD.
Pienta, A., M.D. Hayward, and K.R. Jenkins. 2000. "Patterns of Health and Marriage in Later Life." *Journal of Family Issues* 21:559-586.
Vargas, C.M., V.L. Burt, R.F. Gillum, and E.R. Pamuk. 1997. "Validity of Self-reported Hypertension in the National Health and Nutrition Examination Survey III, 1988-1991." *Preventive Medicine* 26(5 Pt 1):678-685.
Verbrugge, L.M.and A.M. Jette. 1994. "The Disablement Process." *Social Science and Medicine* 38(1):1-14.
World Health Organization. 1980. *International Classification of Impairments, Disabilities, and Handicaps*. Geneva.
—. 2001. *International Classification of Functioning, Disability, and Health*. Geneva.