HRS EXTERNAL REVIEW
LONGITUDINAL ASPECTS OF HEALTH

Submitted by

Karen Bandeen-Roche, Ph.D.
Associate Professor of Biostatistics, Johns Hopkins University
November 1, 2002

I.   **Overview**

         This report reviews the quality of longitudinal data collected in the Health and Retirement Study (HRS) and its associated studies. It considers broad aspects of HRS data utility for longitudinal analyses, and of their strength for investigating longitudinal aspects of health specifically. Throughout the report, ***longitudinal data are defined to be repeated measurements on sampling units*** (typically, individuals) ***over time***. Observe that this definition includes time-to-event data, where sampling units are known not to have had a defining event at some baseline time, and to either have the event or be censored at some later time.

         The rest of the report begins with general discussion on the import and usefulness of longitudinal data, and on criteria that longitudinal measurements must satisfy to yield high quality inferences on longitudinal hypotheses (Section II). It then briefly reviews the published research that has utilized longitudinal data from the HRS family of studies, evaluates the HRS data sets with respect to each of the criteria laid out in Section II and for their convenience of usage, and makes recommendations for strengthening the data or its usefulness. It concludes with synthesizing remarks.

II.   **Introduction: The Import and Usefulness of Longitudinal Data**

         Longitudinal data uniquely inform the study of individual-level change over time.[1] Thus, they may be used to ***delineate natural histories*** and evaluate their variability. More, it is well recognized that ***longitudinal data have greater strength to inform questions of causality and mechanisms*** than cross-sectional data. This claim can be understood by considering what causality entails. One useful definition states that characteristic A causes characteristic B if, in isolation from external factors, a change in A is accompanied by a change in B[2]. Thus, causality incorporates three components: association between A and B, precedence of A then B, and persistence of the association holding factors other than A and B constant. A variety of study designs may establish association. In the absence of external knowledge that clearly establishes precedence of cause and effect, precedence may only be established by longitudinal observation. Longitudinal data are not sufficient to guarantee isolation, but they do so more strongly than cross-sectional data by providing sampling units' prior characteristics as controls for their later characteristics.

         Arguably, ***ideal longitudinal measurement would record health and other characteristics continuously, accurately, precisely, and completely*** over a time interval of interest. As this ideal is infeasible in large-scale observational studies, one must typically grapple with shortcomings of available longitudinal data relative to the ideal. As a first shortcoming, data are typically ***measured at relatively widely spaced intervals***. This shortcoming threatens ability to establish temporal precedence due to the potential for feedback effects, such that change in A → change in B → change in A; for example, decline in socioeconomic status may lead to poor health, which in turn may compromise earning ability, hence further worsen socioeconomic status.   It threatens ability to detect association and achieve isolation due to the potential for unmeasured intervening influences and events in the interim.

As a second shortcoming, ***measurement may be inaccurate, imprecise, or subject to changes*** over time, which may lead to biased assessment of association. Non-differential misclassification of categorical outcomes and measurement errors in explanatory variables mute the strength of associations. ***Differential measurement is arguably the greatest measurement-related threat*** to inferences that can be drawn using longitudinal data, because the resulting biases in assessing associations may only be corrected with instrumental or validating information.[3][4] Differential measurement occurs when explanatory variables are correlated (associated) with errors with which outcome characteristics are measured, or when errors in measuring explanatory variables are associated with outcomes, or both. In this case observed associations between explanatory and outcome variables may reflect anomalies of measurement rather than relationships among the characteristics that are targeted by measurement, and resulting biases for estimating the latter may be in any direction. Finally, high quality measurement also has to do with assessing targeted aspects of socio-economic status or health with a ***sufficiently broad range of measures***, and collecting ***good measures of change in explanatory variables*** as well as outcomes. Such measures are generally needed to convincingly establish the precedence requirement for arguing causality.

Thirdly, ***data rarely are completely ascertained*** in prospective observational studies, and depending on the mechanism that underlies data being incomplete, ***naively treating their incompleteness may severely distort findings on strength of associations and sabotage the ability to isolate associations*** to hypothesized causal factors. Rubin[5] has usefully classified mechanisms underlying "missingness" as being of three types: missing completely at random (MCAR), missing at random (MAR), or non-ignorable (NI). In the context of longitudinal data, MCAR means that missingness is not associated with any factor under study; MAR means that missingness may be associated with measured explanatory variables or past responses; and NI means that missingness may depend on the *unobserved*, missing data. The ***NI mechanism poses the severest threat to the validity of findings, because missingness can be neither verified nor excluded as NI unless external information is collected***. While models for NI missing data can be formulated,[6][7][8][9] inferences from these models are sensitive to assumptions that cannot be verified without external information.[10] In contrast, a variety of approaches preserve validity of inferences if incomplete data are MCAR or MAR, including maximum-likelihood fitting.[11][12][13] Researchers frequently invest substantial effort to produce analyses of whether and how missingness is associated with observed variables. Such analyses usefully identify factors that must be considered in analyses designed to account for MAR incomplete data but do not address NI missingness.

III.   Longitudinal Usage of HRS Data to Date

As one method for evaluating the usefulness of the HRS and associated data sets for longitudinal research on health, I surveyed the peer-review literature for the extent and quality of articles reporting such research. Four hundred and eighty-seven publications are currently listed in the bibliography posted on the HRS website, many of which are not peer-reviewed. Table 1 summarizes the dates of early and final public release for the Core HRS and AHEAD data sets. A key observation is that 3rd- and later wave data have only been publicly available in finalized

form since mid-late 2001, but they have been available in early form since 1997 for HRS and 1999 for HRS 4/AHEAD 3. HRS articles employing longitudinal data were identified both by searching the bibliography, and by a current PubMed search, on approximately a dozen keywords that identify longitudinal data or analyses (longitudinal, panel, change, decline, recovery, incidence, transition, mortality, dynamic, etc.). This method may have missed some articles targeted by the search, but I believe it has captured most. I read approximately 2/3 (19) of the articles I identified, in order of their convenience for me to obtain at my institution. During this stage I excluded a few papers identified through a longitudinal-oriented keyword, but that clearly did not employ longitudinal data, from my survey.

Table 1
Listed[1] Public Availability Dates, Follow-up HRS[2] and Associated Core Data

Study	Wave 2		Wave 3		Wave 4	
	Early	Final	Early	Final	Early	Final
HRS	6/95	6/98	5/97	5/01	6/99	12/01
AHEAD	5/97	5/02	6/99	12/01	5/01	9/02

1 HRS Website, 10/02
2 HRS Wave 5 not listed: availability as for AHEAD Wave 4

The article survey I conducted identified ***30 peer-reviewed "longitudinal" publications***, the most recent of which appeared in August, 2002 (Appendix). Even accounting for data availability dates, ***these comprise a remarkably low percentage of the 490+ existing papers***, counting articles not yet appearing in the HRS bibliography due to their recency. Moreover, ***only 2 of the 19 articles I read utilized outcome data from more than two waves***. With one exception, the remaining articles did not employ techniques specific to longitudinal data, but rather regressed a single incidence, change category, change score, or transition-in-change-score outcome on explanatory variables. The next sentences refer to all such outcomes as "change." The several papers that I regarded most highly regressed change over a later period (e.g. 1994-6) on change in explanatory variables over a previous period (e.g. 1992-4); those that I regarded least highly regressed outcome change on values of explanatory variables measured after the beginning of the time interval under investigation; most regressed two-year outcome change on baseline explanatory variables, or on their baseline values and concurrent change. ***Probability weighting was applied (or not) very inconsistently***, and many of the papers that may have used the study weights described the specific usage unclearly. Finally, very ***few of the papers I surveyed seemed to have accounted for potential within-couple outcome correlations***.

To summarize, my estimation is that the longitudinal potential of the HRS data has not been well exploited to date. This observation is in line with comments on pink sheets from the 1999 NIH review of the HRS renewal, that "awareness and use of [the HRS] data among non-economists is less than it should be...." The next sections explore shortcomings and inconveniences that may put off potential users.

IV.   Longitudinal Data Shortcomings in the HRS Data

IV.1.   Wave spacing

HRS and AHEAD waves have been scheduled at 2-3 year intervals. While a two year interval is approximately in line with a number of other prospective population-based studies of older adults, the length of the interim periods would seem to allow for substantial fluctuations in health and life circumstances in older adults, as well as their external circumstances. Unmeasured sequences of events would threaten the strength of inferences that derived from HRS data by muddying both strength and isolation of associations and by complicating arguments of temporal precedence. HRS and its related studies certainly collect retrospective information about interim events, but there are *two ways in which HRS investigators might strengthen the studies' ability to establish precedence and disentangle feedback effects*:

•   They might design a substudy in which an *appropriate subset of study participants are evaluated substantially more frequently than in the primary study*. The goal would be to delineate the sorts of event sequences that are missed in two-year intervals and by retrospective report. As one example of this idea, the Women's Health and Aging Study included such an investigation.[14]

•   They might *make information on the timing of societal shifts (e.g. Medicare policy changes) more accessible to potential users of HRS data*. I hesitated to raise this point, because ascertaining such information is surely the responsibility of individual researchers. I have raised it because so many of my colleagues with whom I have discussed the HRS data over the course of my review raised it with me.

IV.2.   Measurement errors

The HRS data possess, or have been strengthened with, *several very advantageous features related to the measurement of health and functioning*. First, they provide an *excellent breadth* of health and functioning measures, considering that they derive from a survey and a study that is heavily focused on economic issues. Second, HRS investigators have recently provided *excellent working reports* on the measurement of affective disorders[15] and cognition[16] by their data. These reports discuss assessment content, psychometric properties, completeness, and changes over time, and they make specific recommendations on handling instances of imprecise and differential measurement. Third, the HRS website provides a convenient engine for *cross-referencing measures across waves*.

However, the HRS data also have features that raise *serious concerns related to differential measurement*. These include changes in the primary mode of measurement (face-to-face versus telephone, between waves 1 and 2); proxy respondents; assessment trends, such as practice effects, and changes in question wording or assessment content. Based on the literature,[17] [18] my discussions of the HRS data with colleagues, and my own review, *sub-optimal health assessment in the early HRS waves and subsequent changes in assessment are most troubling*. These are sufficiently severe that the HRS documentation report on affective

disorders recommends commencing studies involving depression with the second wave of the study. Upon examining the physical functioning portion of the questionnaire, *I would doubt the validity of longitudinal analyses that did not exclude the first and possibly the second HRS wave*. Specifically, the 1992 stem was: "How difficult is it for you to...(say, walk several blocks)." In contrast, the stem beginning in 1996 was: "...do you have any difficulty with walking several blocks," and the 1994 stem blended the 1992 and 1996 versions. It seems plausible that persons with health impairments of mild-to-moderate severity might tend to respond "a little difficult" when presented with the 1992 stem, in contrast to "*not at all* difficult," but respond "no" to the less sharply delineated "do you have any difficulty" in the 1996 stem. In that case, analyses relating health status and change to functional status and change might well overstate functional consequences of worsening of impairments from moderate to severe, and understate the functional consequences of onset of early-to-moderate impairments. However, *commencing analyses at later study waves also presents threats to validity of findings*, because study samples of older adults likely become less representative of the populations they are designed to represent over time due to study attrition.

I have *five recommendations* related to measurement issues in HRS:

- *More area-specific documentation reports* of the sort that have been developed for measures of affective functioning and cognition are needed. The need may be most pressing with respect to measures of physical functioning and disability.

- Users of HRS data would be aided by a *documentation report on measurement issues* that cuts across substantive areas. It should cover both the less avoidable issues, such as practice effects and the need for proxy respondents, and changes in wording and assessment. Without a unified document of this type, it is hard for investigators to keep track of all the various measurement-related threats to the validity of their longitudinal findings and tackle the most severe of these in their analyses.

- HRS investigators should consider conducting *substudies to document the consequences of wording and assessment changes*, particularly with respect to the assessment of functioning. These might be analytic, like those that have been reported in the documentation reports on affective disorder and cognition measures, or they might take the form of a multi-trait multi-method assessment study.[19] Where the latter is feasible, findings could inform the creation of measurement error models[3] for earlier assessments with respect to their later versions, hence increase investigators' ability to assess sensitivity of their findings to differential measurement effects.

- *Future changes of assessment must be avoided* unless very nearly necessitated by scientific advancement. Where changes are necessitated, new measures should be preceded by their old counterparts in at least a subsample of participants in one wave.

- *Continuing efforts are needed to keep existing resources correct and up-to-date*. I located errors in the cross-referencing engine (e.g., some Wave 1 ADL-type questions mis-located to Health Care Utilization rather than Health Status), and the documentation

report on affective functioning only extends through the third HRS Wave.

IV. 3.   Missing Data

The sources of data incompleteness in the HRS and its related studies cut broadly across prospective observational studies: participants may drop out, miss waves of evaluation, or refuse to participate in selected assessments within waves of evaluation. Moreover, death or other events may censor measurements. In a study of older adults, one expects these mechanisms to be informative, hence to pose serious potential threats to validity of findings.

The HRS data possess, or have been strengthened with, *several very advantageous features related to data completeness and accounting for incompleteness*. First, *initial study response and subsequent retention rates are good* for a study of an older population, ranging from 70-80% and 80-90+% respectively. To the study's credit, most persons who declined to participate in a wave continued to be invited for future waves. Second, *sampling weights have been provided for all study waves*, which may mitigate concerns over initiating analyses at post-baseline waves. Third, study researchers appear to be working hard and thoughtfully to *maximize retention,[20] [21] item response[22] [23] and completeness[24]*, and to publish *recommendations on handling missing data*.[9] [20] They have documented that attrition is related to declining health and to structural aspects of interviewing.[20] Broadly, these features represent some of the best outcomes and efforts I have seen in comparable studies, and the investigators are to be commended.

Nonetheless, *three concerns persist* related to missing data in the HRS studies. First, to judge from my literature review (§ III), *substantive researchers have not generally been utilizing the recommendations and tools* that HRS investigators have been providing to treat data incompleteness, including accruing knowledge on the determinants of missing data, methodological recommendations, and sample weighting. Second, *the sample weights do not incorporate information on the factors that appear primarily associated with dropout*, so that they likely do not well protect against related biases. Finally, while the delineation of factors related to attrition may reasonably extrapolate to hypotheses about NI missingness, it *only directly informs MAR missingness*. HRS investigators might consider the following recommendations:

• *Synthesizing recommendations and description of tools for handling missing data in a format that is accessible to substantive researchers*, perhaps as a documentation report. Recommendations should include advice on handling factors known to associate with attrition in analyses, so that findings are valid under MAR with respect to these factors.

• *Developing a supplemental set of sampling weights* that incorporate factors known to be associated with attrition in HRS.[20] In my opinion, these would support valuable sensitivity analyses for the effects of attrition on longitudinal findings.

• *Conducting substudies or implementing methods to investigate factors underlying NI missingness*. As an example of a substudy, one might build Medicare expenditure

histories for a sample of those having dropped out of the HRS and its related studies. This is widely possible for deceased persons, and consent might be obtainable from persons still living. As an example of a method, researchers might design a "foot in the door" questionnaire, such that those declining to participate further might be persuaded to respond to a carefully selected *few* questions on overall health, functioning, and economic status. This strategy has been used to advantage elsewhere.[25]

- ***Efforts to maximize ascertainment of death***. I understand that work to assess the quality of ascertainment through the National Death Index is underway (personal communication, Dan Hill, 10/29/02). As death is a "competing" rather than "independently censoring" risk, the completeness of information on this event may materially affect the validity of researchers' longitudinal analyses.

IV. 4. Other Complexities

Moving beyond threats to validity of findings, inconveniences may raise barriers to working with a given longitudinal data set. The **HRS investigators seem to have worked hard to make the use of its data reasonably convenient**. I found it easy to download data and tracking files, codebooks, and documentation from the HRS website, and to format the data in the Stata statistical package. Downloadable information includes explicit advice on which weights to apply in various analyses. The investigators are to be commended for its efforts in this area. I would judge that mere convenience of obtaining data is not an off-putting point for researchers wishing to conduct analyses.

There are three areas in which potential users may be discouraged from using HRS data, and on which I have recommendations:

- The data set and its documentation comprise a ***vast amount of information, and sorting through it is cumbersome***. HRS investigators might consider the development of an overarching monograph to introduce users to the data. The Women's Health and Aging Study provides an example where such a document has proven invaluable.[26]

- My impression is that many ***researchers are unsure when sample weighting should be applied and intimidated by doing it.*** The "Sampling Weights" document that is downloadable from the HRS website advises users to conform to "best practice" in their disciplines in deciding when to weight. This may be inadequate: for example, if African-Americans were at risk for particularly severe health declines relative to the population average, not accounting for their over-sampling would introduce biases into analyses of health declines. This example highlights that sample weighting may be neither necessary nor sufficient to avoid biases in structural analyses; rather, factors that determine deviations from simple random sampling might be controlled, or the heterogeneity that they induce acknowledged, in analyses. Providing data users with ***more explicit advice on what were the factors that determined departures from simple random sampling in HRS, and what variables might be used to handle them in analyses***, would be useful.

- ***Conducting explicitly longitudinal analyses with a large and complex data set is daunting***, however conveniently the data may be obtained. A ***series of papers providing case-study applications of explicitly longitudinal methods to real scientific questions in HRS data*** is needed. Very few such papers seem to have been published.[9] Clearly the expertise to produce them exists among HRS investigators with quantitative expertise, in the Institute for Social Research, and in the University of Michigan departments of statistics and biostatistics. Valuable visibility might be gained by inviting quantitatively oriented gerontologists at various locations to write papers, prominently including members of the recently formed Gerontological Society of America interest group in Measurement and Statistics. To summarize, the sort of papers just proposed might go a long way toward heightening awareness and strengthening usage of the longitudinal HRS data.

V.   Summary

The data from HRS and its associated studies offer many strengths for conducting longitudinal analyses of health status in older adults. In addition to stellar financial, employment, and family measures, they include a solid ***breadth of health and functioning measures***. Many measures were obtained repeatedly at each wave of evaluation. Valuable thought and substantial effort have been invested in maintaining ***high study participation and item completeness rates,*** and these efforts are reflected in initial recruitment and subsequent follow-up rates and in item quality. ***Investigators have produced extremely valuable reports*** to inform users on the measurement quality of selected assessments and factors associated with study attrition. Up-to-date ***sampling weights are available*** for each study wave, and the ***data are conveniently accessible*** and documented for concordance across waves. To summarize, the data lay the foundation for an outstanding resource with which to investigate the role of economic factors in the natural history of health and functioning in older adults.

Usage of the longitudinal data collected by HRS has not yet approached its potential. In this report, I identified three broad strategies that might broaden and strengthen data usage:

- The production of ***additional documentation reports*** to advise users on treatment of incomplete or differential measures, as well as ***methodological case study papers*** to demonstrate powerful uses of HRS data.

- The ***design and conduct of substudies*** to elucidate short-term change, NI missingness mechanisms, and inform the correction of differential measurement biases.

- ***Avoiding future changes in mode of assessment*** wherever possible.

APPENDIX
Peer-Reviewed HRS Papers That Are Clearly Longitudinal

Baker, David W., Joseph J. Sudano, Jeffrey M. Albert, Elaine A. Borawski and Avi Dor.
2001. "Lack of Health Insurance and Decline in Overall Health in Late Middle Age."
New England Journal of Medicine, 345(15):1106-12.

Choi, Namkee G. and L. Schlichting-Ray. 2001. "Predictors of Transitions in Disease and
Disability in Pre- and Early-Retirement Populations." *Journal of Aging and Health*,
13:379-409.

Clark, Daniel O., Timothy E. Stump, and Frederic D. Wolinsky. 1998. "Predictors of Onset
of and Recovery From Mobility Difficulty Among Adults Aged 51-61 Years." *American
Journal of Epidemiology* Vol. 148(No. 1):63-71.

Damush, T.M., T.E. Stump, and D.O. Clark. 2002. Body-mass Index and 4-year Change in
Health-Related Quality of Life. *Journal of Aging and Health*, 14(2):195-210.

Ekerdt, David J., Jennifer Hackney, Karl Kosloski, and Stanley DeViney. 2001. "Eddies in
the Stream: The Prevalence of Uncertain Plans for Retirement." *Journals of
Gerontology:Social Sciences* 56B(3):S162-S170.

Flippen, Chenoa and Marta Tienda. 2000. "Pathways to Retirement: Patterns of Labor Force
Participation and Labor Market Exit Among the Pre-Retired Population by Race,
Hispanic Origin, and Sex." *The Journals of Gerontology, Series B: Psychological and
Social Sciences* 55B(1):S14-S27.

Fonda, Stephanie J., Elizabeth C. Clipp, and Geaorge L. Maddox. 2002. "Patterns in
Functioning Among Residents of an Affordable Assisted Living Housing Facility." *The
Gerontologist*, 42(2):178-87.

Fonda, Stephanie J. and A.Regula Herzog. 2001. "Patterns and Risk Factors of Change in
Somatic and Mood Symptoms Among Older Adults." *Annals of Epidemiology* 11:361-68.

Fonda, Stephanie J., Robert B. Wallace, and A. Regula Herzog. 2001. "Change in Driving
Patterns and Worsening Depressive Symptoms Among the Elderly." *Journals of
Gerontology, Social Sciences*, 56B(6):S343-351.

Franks, MM., A.M. Pienta, and L.A. Wray. 2002. "It Takes Two: Marriage and Smoking Cessation in the Middle Years." *Journal of Aging and Health*, 14(3):336-54.

Freedman, Vicki A., Hakan Aykan, and Linda G. Martin. 2002. "Another Look at Aggregate Changes in Severe Cognitive Impairment: Further Investigation into the Cumulative Effects of Three Survey Design Issues." *Journals of Gerontology, Series B*, 57(2):S126-S131.

———. 2001. "Aggregate Changes in Severe Cognitive Impairment Among Older Americans: 1993 and 1998." *Journals of Gerontology, Social Sciences*, 56B(2):S100-S111.

Gallo, William T., Elizabeth H. Bradley, Michele Siegel, and Stanislav V. Kasl. 2001. "The Impact of Involuntary Job Loss on Subsequent Alcohol Consumption by Older Workers." *Journals of Gerontology:Social Sciences*, 56B(1)S3-9.

———. 2000. "Health Effects of Involuntary Job Loss Among Older Workers: Findings From the Health and Retirement Study." *Journals of Gerontology: Social Sciences,* 55B(3):S131-40.

Hill, Daniel H. and Robert J. Willis. 2001. "Reducing Panel Attrition: A Search for Effective Policy Instruments." *Journal of Human Resources*, 36(3):416-438.

Hughes, Mary E. and Linda J. Waite. 2002. "Health in Household Context: Living Arrangements and Health in Late Middle Age." *Journal of Health and Social Behavior*, 43(1):1-21.

Johnson, Richard W. and Stephen Crystal. 2000. "Uninsured Status and Out-of-Pocket Costs at Midlife." *Health Services Research*, 35(5):911-932.

———. 1997. "Health Insurance Coverage at Midlife: Characteristics, Costs, and Dynamics." *Health Care Financing Review* 18(3):123-48.

Lee, Myoung J. and Satoru Kobayashi. 2001. "Proportional Treatment Effects for Count Response Panel Data: Effects of Binary Exercise on Health Care Demand." *Health Economics*, 10(5):411-428.

Nothwehr, Faryle and Tim Stump. 2000. "Health-Promoting Behaviors Among Adults with Type 2 Diabetes: Findings from the Health and Retirement Study." *Preventive Medicine*, 30(5):407-414.

Olson, Janice A. 1999. "Who Is "62 Enough?" Identifying Respondents Eligible for Social Security Early Retirement Benefits in the Health and Retirement Study." *Social Security Bulletin* 62(3):51-56.

Ostbye, Truls, D.H. Taylor, K.M. Krause and L.V. Scoyoc. 2002. "The Role of Smoking and Other Modifiable Lifestyle Risk Factors in Maintaining and Restoring Lower Body Mobility in Middle-Aged and Older Americans: Results from the HRS and AHEAD." *Journal of the American Geriatrics Society*, 50(4):691-699.

Ostbye, Truls, D.H. Taylor and Jung Sang-Hyuk. 2002. "A Longitudinal Study of the Effects of Tobacco Smoking and Other Modifiable Risk Factors on Ill Health in Middle-Aged and Older Americans: Results from the Health and Retirement Study." *Preventive Medicine*, 34:334-335.

Ostermann, Jan and Frank A. Sloan. 2001. "Effects of Alcohol Consumption on Disability Among the Near Elderly: A Longitudinal Analysis." *Millbank Quarterly*, 79(4):487-515.

Perreira, K.M. and Frank A. Sloan. 2002. "Excess Alcohol Consumption and Health Outcomes: A 6-Year Follow-Up of Men Over Age 50 from the Health and Retirement Study." *Addiction*, 97(3):301-310.

———. 2001. "Life Events and Alcohol Consumption Among Mature Adults: A Longitudinal Analysis." *Journal of Studies on Alcohol*, 62(4):501-508.

Reville, R.T., J. Bhattacharya and L.R. Sager-Weinstein. 2001. "New Methods and Data Sources for Measuring Economic Consequences of Workplace Injuries." *American Journal of Industrial Medicine*, 40(4):452-463.

Sloan, Frank A. and C.J. Conover. 1998. "Life Transitions and Health Insurance Coverage of the Near Elderly." *Medical Care* Vol. 36(No. 2).

Sloan, Frank A. and Edward C. Norton. 1997. "Adverse Selection, Bequests, Crowding Out, and Private Demand for Insurance: Evidence From the Long-Term Care Insurance Market." *Journal of Risk and Uncertainty* Vol. 15(No. 3):201-19.

Sloan, Frank A., V. Kerry Smith and Donald H. Taylor, Jr. 2003. *Parsing the Smoking Puzzle: Information, Risk Perception, and Choice.* Cambridge: Harvard University Press.

TOTAL: 30

Literature Cited

1. Diggle PJ, Liang K-Y, Zeger SL. *Analysis of Longitudinal Data*. New York: Oxford, 1994, p 17 ff.

2. Bollen KA. *Structural Equations with Latent Variables*. New York: Wiley, 1989, p. 40 ff.

3. Carroll RJ, Ruppert D, Stefanski LA. Measurement Error in Nonlinear Models. London: Chapman & Hall, 1995, pp. 245-7.

4. Brenner H, Savitz DA, Gefeller O. The effects of joint misclassification of exposure and disease on epidemiologic measures of association. *J Clin Epidemiol* 1993; 46(10):1195-202.

5. Rubin D. Inference and missing data (with discussion). Biometrika 1976; 63(3):581-592.

6. Fitzmaurice GM, Laird NM. Analysing incomplete longitudinal binary responses: a likelihood-based approach. *Biometrics* 1994; 50:601-612.

7. Little RJA. A class of pattern-mixture models for normal incomplete data. *Biometrika* 1994; 81(3):471-483.

8. Baker SG. Marginal regression for repeated binary data subject to nonignorable nonresponse. *Biometrics* 1995; 51(3):1042-1052.

9. Hill D. Adjusting for attrition in event history analysis. *Soc Methodol* 1997; 27:393-416.

10. Dawid AP. Discussion of Diggle PJ and Kenward MG, Informative dropout in longitudinal data analysis. *Appl Stat* 1994; 43:49-93.

11. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. Wiley, New York, 1987.

12. Laird NM. Missing data in longitudinal studies. *Statist Med* 1988; 7:303-315.

13. Robins JM, Rotnitzky A, Zhao LP. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *J Am Stat Assoc* 1995; 90:106-121.

14. Guralnik JM, Ferrucci L, Penninx BW, et al. New and worsening conditions and change in physical and cognitive performance during weekly evaluations over 6 months: The Women's Health and Aging Study. J Gerontol Med Sci 1999;54(8):M410-22.

15. Steffick DE. Documentation of Affective Functioning Measures in the Health and Retirement Study. Survey Research Center, University of Michigan, Ann Arbor MI, DR-005, 2000.

16. Ofstedal MB, McAuley GF, Herzog AR. Documentation of Cognitive Functioning Measures in the Health and Retirement Study. Survey Research Center, University of Michigan, Ann Arbor MI, DR-006, 2002.

17. Freedman VA. Implications of asking "ambiguous" difficulty questions: An analysis of the second wave of the Asset and Health Dynamics of the Oldest Old Study. *J Gerontol Soc Sci* 2000*;* 55B (5): S288-S297.

18. Perreira KM, Sloan FA. Life events and alcohol consumption among mature adults: A longitudinal analysis. *J Stud Alcohol* 2001; 62(4): 501-8.

19. McCrae RR, Costa PT. Validation of the five-factor model of personality across instruments and observers. *J Pers Soc Psychol* 1987; 52:81-90.

20. Hill DH, Willis RJ. Reducing panel attrition: A search for effective policy instruments. *J Hum Res* 2001; 36(3):416-38.

21. Rodgers W. Size of incentive effects in a longitudinal study. Proceeding paper, personal communication 10/29/02, DH Hill to K Bandeen-Roche.

22. Juster FT, Smith JP. Improving the quality of economic data: Lessons from the HRS and AHEAD. *J Am Statist Assoc* 1997; 92:1268-78.

23. Hill DH. Unfolding bracket method in the measurement of expenditures and wealth. In: *Wealth, Work and Health: Innovations in Measurement in the Social Sciences*, eds. JP Smith, RJ Willis, Ann Arbor, MI: University of Michigan Press, 1999, 64-86.

24. Cao H. IMPUTE: A SAS Application System for Missing Value Imputations, with Special Reference to HRS/AHEAD Income/Assets Imputations. Survey Research Center, University of Michigan, Ann Arbor MI, DR-007, 2001.

25. Munoz B, West S, Rubin GS, et al. Who participates in population based studies of visual impairment? The Salisbury Eye Evaluation project experience. Ann Epidemiol 1999;9(1):53-9.

26. Guralnik JM, Fried LP, Simonsick EM et al., eds. The Women's Health and Aging Study: Health and Social Characteristics of Older Women with Disability. Bethesda, MD: National Institute on Aging, 1995; NIH Pub. No. 95-4009.