UW Biostatistics Working Paper Series

1-25-2002

Performance of the Halex in Logitudinal Studies of Older Adults

Paula Diehr
University of Washington, pdiehr@u.washington.edu

Donald L. Patrick
University of Washington, donald@u.washington.edu

Gregory L. Burke
Wake Forest University, gburke@wfubmc.edu

Jeff D. Williamson
Wake Forest University, jwilliam@wfubmc.edu

Performance of the Halex in Longitudinal Studies of Older Adults

Introduction

Data on self-rated health and on activity limitations have been used to calculate health-related quality of life for the general population (Halex) [1 2] and later for the Medicare population (MCBS-QOL),[3] and also as a pseudo-utility measure for patients with acute myocardial infarction.[4] Self-rated health takes on the levels Excellent, Very good, Good, Fair, Poor, or Dead (EVGFP). For persons aged 65 and over, activity limitations are categorized as the most serious of "no limitation", "limitation in at least one instrumental activity of daily living" (IADL), or "limitation in at least one activity of daily living" (ADL).

The scores that were assigned to the Halex were originally developed for cross-sectional national data using a method called multi-attribute utility scaling. The investigators specified that being in Excellent health with no limitations would have a score of 100, and Death a score of 0. They assumed that the correct score for being in Poor health with ADL limitations was 10 (they used no data on mortality). They also assumed that being in Poor health with no limitations was equivalent to being in Excellent health with ADL limitations. They assigned the latter cell the score 47, which was the estimated utility of that state in a survey of the general public in Canada.[5] The resulting scores are shown in Table 1. Note that the scores have the desirable property of being monotonic in self-rated health (the score in each row decreases as you move to the right) and in activity limitations (the scores in each column decrease monotonically). The best state is Excellent health with no limitations, and the second best is Very good health with no limitations. All people with IADL or ADL limitations have lower scores that people in Fair health with no limitations.

[Table 1 about here]

Another desirable property of a health index is that future health be monotonically related to current health, on average. That is, if state A has a higher score (better health) than state B, one would expect that at a later time, persons originally in state A would be healthier, on average, than those originally in state B. We examined whether the mean Halex at year 2 was monotonically related to the mean Halex at year 0, using data from the Cardiovascular Health Study (CHS).

Methods

Data

Data for the transitions among health states were taken from the Cardiovascular Health Study (CHS), a population-based longitudinal study of 5,888 adults 65 years of age and older designed to identify factors related to the occurrence of coronary heart disease and stroke. [6] CHS subjects were recruited from a random sample of the HCFA Medicare eligibility lists in four communities in the United States. [7] Persons who were institutionalized, were not expected to remain in the area for the next three years, used wheelchairs at home, or were receiving hospice treatment, radiation therapy or chemotherapy for cancer at baseline were excluded. About 70% of those invited participated in the study.[8] At baseline and every year thereafter, subjects rated their health as excellent, very good, good, fair, or poor (EVGFP), and also reported difficulties with their activities of daily living (ADL) and instrumental activities of daily living (IADL). Vital status is known for all subjects; approximately 1700 had died by the end of current follow-up.

To avoid bias, we interpolated scores for missing observations whenever there was a valid measurement before and after the missing observation (about 4% of all the data). After interpolation of self-rated health we added a small amount of random error and then rounded to the nearest living health category. After interpolation, 97% of all persons had complete data for EVGFP. Approximately a third of the subjects who died had missing data the year they died. Because these transitions were important, we substituted the last known score for ADL or IADL for the score before death, when that score was missing. This probably created a small positive bias. We included CHS data from year 9, even though they were not complete, to obtain additional observations for the oldest subjects.

Transitions

We counted the number of transitions from one health state to another, two years in the future, based on a person's initial health state (e.g., the number of people originally in Excellent health with no limitations who were in Poor health with IADL limitations two years later). The two-year interval was chosen in part because it has been used in the literature, and also because some change in health could be expected in two years. Most persons provided data on more than one transition. For example, a person who was alive at the end of current follow-up would have had 10 annual measurements. We used that person's transition from baseline to year 2, and also from year 1 to year 3, from year 2 to year 4, etc. for a maximum of 8 transitions per person. People who died, or who were part of a second cohort that has been followed only 5 years to date, contributed fewer transitions. (We did not count transitions from death to death, and so a person who died would have contributed at most 2 transitions that ended in death).

A total of 40,827 transitions were available for analysis. In sensitivity analyses, we separated them by age and by sex. For each transition we calculated the Halex score for time 0 (T0) and 2 years later at time 1 (T1).

Analysis

We plotted the mean Halex score at T1 against the score at T0. Since some discrepancies were found, we recoded the scores in an iterative process to make mean health at T1 monotonically related to health at T0. For each state, we replaced the state score with the mean T1 health score for that state, divided by the mean score for people in Excellent health with no limitations at T1. The resulting scale would thus continue to go from 0 to 100. We then recalculated the mean at T1, and repeated this process until convergence was reached (about 7 iterations). We repeated this process using different starting scores, and also separately for men and women, and for persons over and under age 80. We performed a similar iteration requiring that the Excellent with ADL disabilities state (Ea) have the score of 47.

Findings

Table 2 shows the distribution of the initial and final health states of the 40,827 transitions available for study, at T0 and T1. Each state is labeled by two letters, the first indicating the EVGFP state, and the second indicating activity limitations: none, IADL, or ADL limitations. For example, Eo means Excellent with no limitations, Pi means Poor with IADL (but no ADL) limitations and Va means Very good with ADL limitations. Every combination of EVGFP and activity limitations was observed, although the numbers for Excellent with ADL limitations (Ea) and for Poor with no limitations (Po) are smaller. Two years later the distribution was somewhat different; notably, there were 3165 transitions that ended in death. The state with the highest prevalence is Go (Good health with no limitations).

[Table 2 about here]

Figure 1 is a plot of health at T0 versus mean health at T1, using the original scores in Table 1. For example, the Eo state has a score of 100 (Table 1) and the average value at T1 for people in that state at T0 is about 87. The relationship is not strictly monotonic. All of the states with no disabilities at T0 seem "too low" at T1, or perhaps the others are "too high". Notably, persons in the Ea state at T0 had a much higher (better) mean value at T1 than those originally in the Po state, which violates the assumption that the two states had equal health at T0.

[Figure 1 about here]

We iterated the scores to make the relationship more monotone, as described above. Figure 2 and Table 3 show the iterated scores. Figure 2 shows that there is a monotonic and linear relationship between the T0 and T1 health scores using the iterated scores. Table 3 shows that the new scores are monotonic in self-rated health and in activity limitations, and thus have the desirable properties of the original scores as well as having later health monotonically related to previous health.

[Figure 2 and Table 3 about here]

There are other interesting features of Table 3. Scores in the top row are fairly similar to the original scores in Table 1, but scores for people with limitations are considerably higher in Table 3 than in Table 1. The third highest health state is Excellent with IADL limitations, which was fifth highest in Table 1. Note that being in poor health with ADL limitations has a score of 29, considerably higher than the 10 that was originally assumed. Also note that being in Excellent health with ADL limitations has a much higher score (76) than being in Poor health with no limitations (35). There is a 24 point difference due to disability (Eo versus Ea), but a 65 point difference due to EVGFP (Eo versus Po), meaning that the effect of EVGFP is more than twice as large as the effect of limitations. The iterated score for Ea (76) is quite different from the 47 that was assumed in the Halex computations.

It is possible that the assumed score of 47 for the Ea state is appropriate, since it was based on data from the Health Utilities Index study (though not strictly for older adults). We performed a different iteration, still trying to achieve monotonicity in prior health but now requiring the Ea state to have the score 47. The resulting scores are shown in Table 4 and Figure 2a. The resulting scores are again monotonic in both EVGFP and in disability. The score for the Pa state is 14, fairly close to the 10 that was assumed in the original Halex (Table 1).

[Table 4 and Figure 2a about here]

We conducted additional analyses to determine how sensitive the iterated scores were to features of the data. The results are shown in Table 5. The first column shows the original scores, from Table 1; for example, Fo has the score of 63. The second column has the iterated scores from Table 3. The third column gives the iterated solution when a different set of initial scores was used; we set all people with no limitations to 100, all with IADL to 67, and all with ADL to 33. The procedure converged to a very similar solution (column 2 versus column 3). This and other runs not shown indicate that results from the iterative process are independent of the starting scores.

[Table 5 about here]

The remaining columns show iterated scores when only people under or over age 75 (the median age) were examined, and separately for men and women. Scores based on women and persons under 75 are generally higher than the others. This suggests that "optimal" scores are age-dependent. The lowest score ever attained for state Pa (Poor with ADL limitations, for men only) was 20, twice the assumed score of 10 in Table 1.

The bottom row of Table 5 shows the average health score for the T0 population, using the scores in this column. Since there were no deaths at T0, the 0 scores were not used. The original scores give a 10-point lower mean than the others do. The mean score using the Ea=47 iterated rule is the lowest of all, 52 (not shown in Table 5). Clearly, the choice of scores makes a difference.

To examine the longitudinal effect of different scoring systems, we calculated the mean Halex score at years 0, 1, ..., 9, then summed them (half-weighting years 0 and 9) and divided by 100 to estimate the years of healthy life in this period. (Deaths were included). We did the same using the Iterated Scores and the Ea=47 iterated scores. The Halex averaged 5.1 YHL, and had a distribution skewed to the left with a mode at 7.5 to 8 YHL. The iterated scores averaged 5.9 YHL, and were also skewed left with a mode at 7.5 to 8 years. The Ea=47 scoring yielded a mean of 4.0 YHL, and was rather bell-shaped, with a mode at 4-5 YHL. Choice of a scoring system does make a difference.

Discussion

The Halex for older adults requires assigning scores to 16 health states: 5 self-rated health by 3 levels of activity limitation plus death. The methodology of the original Halex required specifying scores in advance for five of the states: Death=0, Eo = 100, Pa = 10, Ea=47, and Po=Ea. The original purpose was to use these scores for cross-sectional analyses involving only living persons, which means that the score for Death is of no consequence. (NCHS used life tables along with the Halex to calculate years of healthy life). If death is not considered, then Pa may be assigned any arbitrary score. Ea was given the score of 47 based on utility scores obtained elsewhere. Those assumptions may be reasonable when the Halex is applied to living persons. The assumption that Po=Ea was completely arbitrary, based on some sense of fairness, giving "equal weight" to activity limitation and to EVGFP. In fact, in Table 3 the difference between excellent and poor was more than twice the difference between having no limitations and having ADL limitations. Thus, even for the living, the scores assigned to the states do not seem consistent.

If the Halex is to be used longitudinally, the scores assigned to Pa and Dead are extremely important. The correct distance between Pa and Dead is not known, and the score assigned to Pa will surely have a large effect on the average health score when subjects begin to die. Thus, for longitudinal use, the Halex is based on two questionable assumptions: Po=Ea and Pa=10. We have shown that the former assumptions is untenable, even if the assumption that Ea=47 is accepted.

The iterated scores that we present were developed based on only a single assumption: that there should be a monotonic relationship between present health and future health. This does not mean that individuals can not change states drastically, but simply that on average, future health can be predicted from current health. Based on this single assumption, the "natural" scores we developed here (in Tables 3 and 4) have the properties of also being monotonic in EVGFP and activity limitations. It is difficult to think of examples where the monotonicity assumption would not hold. One example could be if group A were 10-year-olds with chicken pox, and group B were healthy 40-year-old men, then A would be sicker than B at

T0, but 30 years later B would probably be sicker than A. The comparisons we have made, however, look at health only 2 years in the future and would not likely to be applied in such an extreme case as the example. The original scores have been used as pseudo-utilities.[4] The iterated scores in Table 3 can not be assumed to correspond to utilities, since the score for Ea is 76 rather than 47. The scores in Table 4 would seem to have as much claim to being utilities as the original Halex scores do.

It is not desirable to have two sets of scores for states, one for longitudinal and another for cross-sectional use. If we use the iterated scores in Table 3, the population averages will be substantially higher than based on the original scores, at least for older adults. (The mean health value using the scores in Table 4 were the lowest of all). To the extent that these scores are just a convention, the most important thing is to be consistent, rather than to choose the best set of scores for each application.   The difference in the "natural" scores by age and sex suggest that these variables need to be controlled for in any analysis.

The practical consequences of the discrepancies in the HALEX may not be great when it is used cross-sectionally. Erickson has showed that they work reasonably well. This is in part because the states with the greatest discrepancies are rare states. The discrepancies in longitudinal use, however, may be great.

Limitations.

The CHS data have some positive selection bias, but this is not necessarily a problem. There may have been "too few" sick people at baseline, but the analyses for this paper require only that the people who are in a health state be similar to others in that state, and all health states were represented here. Further, the bulk of the data come from later in the study, when the selection effects should have attenuated. However, these results should be verified using different data sets.

These estimates could be extended to younger age groups if other data sets are found which have longitudinal measures of EVGFP, ADL, and IADL, few losses to follow-up, and complete ascertainment of Deaths.

Conclusions

The orginal Halex index, which was developed for cross-sectional data under strong assumptions, behaves inconsistently over time, and so may not perform well in longitudinal studies of older adults. The sets of "natural" state scores that we derived (Table 3 or Table 4) may be more appropriate in these situations.

Acknowledgments

From the Departments of (1) Biostatistics and (2) Health Services of the University of Washington, (4) Department of Public Health Science, Bowman Gray School of Medicine of Wake Forest University, and (5) Center for Aging and Health, Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins University School of Medicine.

This work was supported by contracts N01-HC-85079 through N01-HC85086 from the National Heart, Lung, and Blood Institute. Data were also obtained from a study sponsored by the Health Care Financing Administration (HCFA) under Cooperative Agreement No. 95-C-99161.

Correspondence to Dr. Diehr, Department of Biostatistics Box 357232, University of Washington, Seattle, WA 98195. E-mail pdiehr@u.washington.edu. Telephone 206-543-8004. Fax 206-543-3286.

Key words. Aged. Health Status. Quality adjusted life years. Years of Health Life. Cost-effectiveness. Utilities.

Table 1

Original Scores

		Self-Rated Health					
		Excellent	Very Good	Good	Fair	Poor	Dead
Limitations	None	100	92	84	63	47	.
	IADL	57	51	45	29	17	.
	ADL	47	41	36	21	10	.
	Dead	.	.	.	.	.	0

Table 2

N of Transitions, by State and Time

	T0	T1
Health State*		
Dead	.	3165
Po	408	469
Fo	3560	3204
Go	11377	10368
Vo	9044	7839
Eo	2517	1732
Pi	419	351
Fi	2376	2182
Gi	3253	3023
Vi	1210	1069
Ei	207	151
Pa	913	1040
Fa	2475	2785
Ga	2256	2592
Va	705	742
Ea	107	115
Total	40827	40827

* E, V, G, F, P stand for Excellent, Very Good, Good, Fair, or Poor; o stands for no limitations, i for IADL limitations, and a for ADL limitations.

Table 3

Iterated Scores

		Self-Rated Health					
		Excellent	**Very Good**	**Good**	**Fair**	**Poor**	**Dead**
Limitations	**None**	100	93	84	68	35	.
	IADL	87	82	73	58	34	.
	ADL	76	73	64	50	29	.
	Dead	.	.	.	.	.	0

Table 4

Iterated Scores with Ea = 47

		Self-Rated Health					
		Excellent	**Very Good**	**Good**	**Fair**	**Poor**	**Dead**
Limitations	**None**	100	75	53	37	18	.
	IADL	70	53	40	29	17	.
	ADL	47	43	33	25	14	.
	Dead	.	.	.	.	.	0

Table 5

Other Iterated Scores*

	Original	**T0 Health -- Iterated**	**Diff Start**	**Under 75**	**Over 75**	**Men Only**	**Women Only**
Da	0	0	0	0	0	0	0
Po	47	35	36	41	30	26	40
Fo	63	68	69	76	64	60	73
Go	84	84	85	88	84	80	87
VGo	92	93	94	95	95	91	95
Eo	100	100	100	100	100	100	100
Pi	17	34	35	48	28	23	43
Fi	29	58	59	69	54	47	64
Gi	45	73	74	82	70	64	78
VGi	51	82	83	89	82	78	85
Ei	57	87	88	94	84	74	94
Pa	10	29	30	41	28	20	37
Fa	21	50	52	63	47	40	55
Ga	36	64	65	75	61	53	68
VGa	41	73	74	85	70	67	76
Ea	47	76	77	87	72	71	78
T0 Mean	67	78	79	84	77	72	81

*The Ea=47 method is not included here. E, V, G, F, P stand for Excellent, Very Good, Good, Fair, or Poor; o stands for no limitations, i for IADL limitations, and a for ADL limitations.

Figure 1

Health at T0 and T1-- Original Values

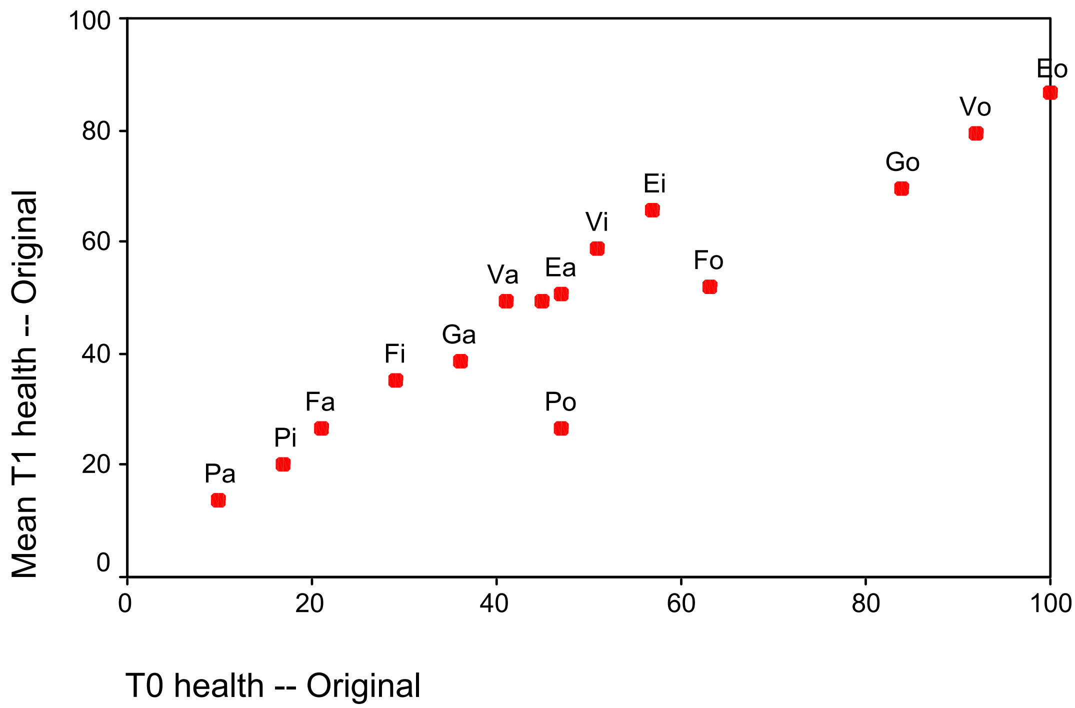


Figure 2

Health at T0 and T1 -- Iterated Values

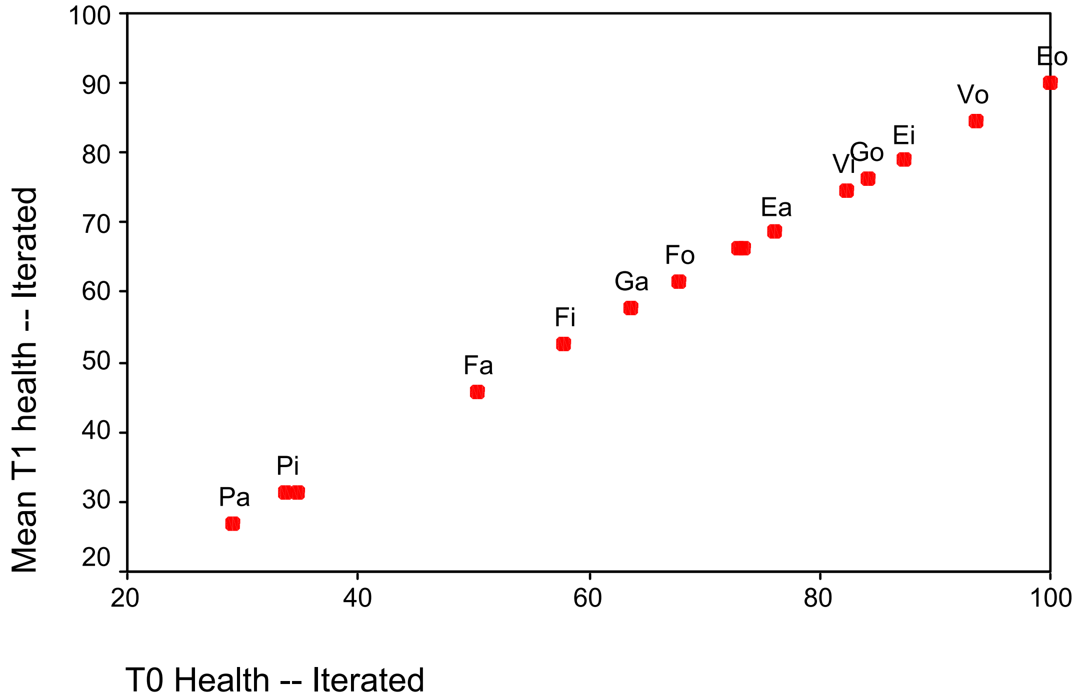
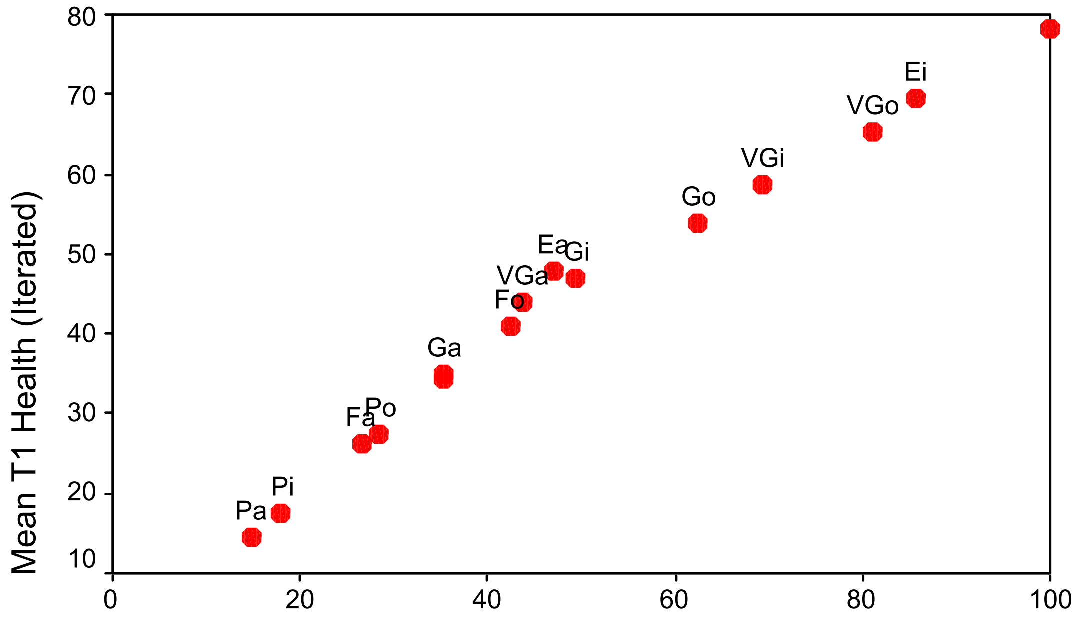


Figure 2a

Health at T0 and T1--Iterated (Ea=47)



T0 Health (Iterated)

References

1. Erickson P, Wilson RW, Shannon L.. Years of healthy Life: Statistical note number 7. Hyattsville, MD: National Center for Health Statistics; 1995.

2. Erickson P. Evaluation of a population-based measure of quality of life: the Health and Activity Limitation Index (HALex). Qual Life Res. 1998;7(2):101-114.

3. Doctor JN, Chan L, MacLehose, R, Patrick DL (2001). Weighted health status in the Medicare Population: Development of the Weighted Health Index for the Medicare Current Beneficiaries Survey (WHIMCBS). Journal of Outcome Measurement, 4(4), 721-739.

4. Bradley CJ, Kroll J, Homes-Rovner M. The health and activities limitaion index in patients with acute myocardial infarction. J Clin Epidemiol 2000; 53:555-562.

5. Drummond MF. Methods for the economic evaluation of health care programmes. Oxford; New York: Oxford University Press; 1997.

6. Fried L, Nemar O, Enright P, et al. The cardiovascular health study: design and rationale. Annals of Epidemiology 1:263-276. 1991.

7. Ives G, Fitzpatrick A, Bild D, et al. Surveillance and ascertainment of cardiovascular events: the cardiovascular health study. annals of epidemiology 5:278-285. 1995.

8. Tell GS, Fried LP, Hermanson B, Manolio TA, Newman AB, Borhani NO. Recruitment of Adults 65 Years and Older as Participants in the Cardiovascular Health Study. Ann Epidemiol 1993; 3:358-366.