Health Services & Outcomes Research Methodology 3: 75–86, 2002

Finding Married Couples in Medicare Claims Data*

THEODORE J. IWASHYNA†
Department of Medicine, Hospital of the University of Pennsylvania, Philadelphia, PA 19104, USA
iwashyna@alumni.Princeton.edu.

GERARDINE BRENNAN
Harris School of Public Policy, University of Chicago, Chicago, IL, USA

JAMES X. ZHANG
Department of Community and Family Medicine, School of Public Health, The Chinese University of Hong Kong, Hong Kong, People's Republic of China

NICHOLAS A. CHRISTAKIS
Department of Health Care Policy, Harvard Medical School, Boston, MA, USA

Received January 5, 2001; Revised March 4, 2002; Accepted March 5, 2002

Abstract. Medicare claims data have proven extraordinarily useful for the cost-effective examination of many epidemiologic and health services research problems. However, their richness in utilization and health status details is offset by a lack of social information. In this paper, we review a method to uniquely link husbands and wives using only data already present in the Medicare claims, allowing the construction of couple-level health care utilization histories. We compare the ability of this method to detect couples to married elderly couples identified in the 1990 Census, and find that it detects at least 59% of couples overall, with only modest variation as a function of the age or age-difference. There appears to be modestly lower detection of Black couples and those in coastal states. We discuss how comparison groups of the widowed can be developed to study the effects of marriage and bereavement on health care use and outcomes.

Keywords: Medicare claims, Census, marital status, elderly, health care utilization

1. Introduction

Many models of decision-making suggest that the key locus for medical decisions is the household, not the isolated individual. That is, individuals may make choices about health care consumption and health production in the context of the particular family in which they live, and in light of the competing demands and resources of other family members. It is thus unsurprising that differences in morbidity and mortality as a function of marital status are well-documented [10, 20, 22, 33]. In the most recent U.S. vital statistics, married men

*This project was supported by a grant from the NIH/National Institute on Aging (N.A.C. Grant No: AG15326-01) and a Medical Scientist National Research Service Award (T.J.I.) from NIH/National Institute of General Medical Sciences (Grant No: 5 T32 GM07281).
†Author to whom correspondence should be addressed.

had an age-adjusted death rate of 1,212.6 per 100,000; widowed men had an age-adjusted death rate of 2,556.9. Married women had an age-adjusted death rate of 662.2; widowed women had an age-adjusted death rate of 1,463.6. These differences are three times as large as the black-white mortality differentials [19].

Further, there are differences in the health care utilization patterns of the married and the widowed across the spectrum of care. The married are more likely to take a number of preventive medical steps, such as having a primary care doctor [37] and getting influenza shots [34]. Married individuals have been shown to have shorter lengths of stay than the widowed for at least some key diagnoses, such as diabetes mellitus and chronic obstructive pulmonary disease [30]; they are also less likely to be readmitted or to die following CHF admissions [5]. The married are more likely to chose definitive treatment when diagnosed with cancer [16]. The importance of marital status to long-term care has often been highlighted [12, 13]. And the married receive better pain control at the end of life [1]. Indeed, there is even the suggestion that marital status per se may strengthen the immune system [8]. However, research into these important differences in the health context has been limited by the paucity of couple-level data. One of the most important new health surveys of the last decade—the Health and Retirement Study (HRS/AHEAD)—was explicitly designed to help remedy this deficit. But HRS/AHEAD is only one data source of somewhat limited sample size; an active research agenda requires that other data sources be available.

The administrative data of the Health Care Financing Administration are a core tool of health services research; we explored ways in which that data might be augmented to allow couple-level analyses. Perhaps the greatest limitation of this data source is the paucity of social information regarding enrollees, despite recent improvements in data quality [11, 21, 36]. Only age, race, and gender are directly available in the claims. Certain proxies for income [3] and ethnicity [32, 41] are available, and extensive work has utilized the geographic indicators in the claims to provide socioeconomic status indicators [9, 14, 15, 18, 28, 29]. However, basic information about the household—as opposed to the individual—has in general been unavailable as there was no recognized way to link husbands and wives together in the claims.

This need not necessarily be so. Two details of HCFA's administrative records have long been known by claims investigators: (1) that some individuals have a Health Insurance Claim number (HIC) that indicates that they qualify for benefits as a spouse; and (2) that individuals may use multiple HICs across their time in Medicare. However, each of these details has an implication. First, the HICs indicating that a beneficiary is a spouse also provide the unique identifier of that individual's partner; fully one in three married husband wife-pairs can be directly linked using this information alone. Moreover, a close reading of the Medicare regulations suggested that the second detail resulted from individuals switching HIC upon the death of a spouse. Beneficiaries may automatically switch from codes indicating primary receipt of benefits to codes indicating receipt of benefits as a surviving widow if the higher earning member of the couples dies first. In general among the elderly, however, the higher earning member of a couple is the man, who is also typically at least two years older than his wife, and also has lower life-expectancy at any given age. Thus the conditions where, retrospectively, husbands and wives may be linked are likely to be common.

The technical details of the HCFA regulations and the possibility of these linkages have been discussed previously [23, 24, 26]. However, that work demonstrated the feasibility of this method only in a very small sample of husbands and wives who both used hospice during 1993—hardly a demonstrative sample for the usefulness of these methods for health service researchers in general. We have not previously explored the external validity of the method for identifying couples. In this paper, we provide details on the first full-population implementation of a method for linking married husbands and wives exclusively using information regarding 32,000,000 people already available within the Medicare claims data. Using the 1993 Medicare Denominator file of all individuals in Medicare in 1993, regardless of whether or not they used medical services, we ask how many husbands and wives can be directly linked. Then we use a six-year follow-up to allow some of members of these elderly couples to die and their surviving partners to change HICs, thus indicating their relationship. The longer the time interval between the time of the claims of interest (in this case, 1993) and the moment research is conducted, the higher the ascertainment would be. We then compare the number of couples we can detect and their characteristics to the best national household data available, the weighted household-level sample of the 1990 Decennial Census Public Use Microdata Sample (PUMS). This afford an opportunity to indirectly validate our method and describe the number of married couples who can be identified in a number of demographic and geographic strata. After demonstrating the representativeness of the couples who can be detected, we discuss the types of analyses that can be carried out.

2. Methods

2.1. Linkage of couples in the Medicare data

The 1993 Denominator file (used here, although others could be used) contains basic identifying information on the entire Medicare population during 1993—that is, it contains information on all individuals who were enrolled in Medicare at any point in 1993, regardless of whether or not they actually filed a claim. The enrolled population has been previously shown to closely approximate the population of all Americans age 65 and above [17, 25].

We obtained from HCFA the entire 1993 Denominator file comprising 38,212,735 records. We also obtained a Vital Status file with mortality follow-up *for the entire Medicare population* through July 6, 1999, and a cross-reference file as of January 6, 1999. These mortality and cross-reference files were the most recent available at the time of this particular data request which is part of an ongoing research effort. The 1993 Denominator file contains 38,209,888 unique individuals, 32,180,588 of whom were at least age 65 as of January 1, 1993. Altogether, there were a total of 36,915,227 Health Insurance Claim numbers (HICs) in the cross-reference file used by these individuals. Of those 65+ in the file, 10,110,008 died by our follow-up on July 6, 1999. (All further information is confined exclusively to the 65 and above population unless explicitly stated. Additional married couples could be detected within the claims by relaxing this restriction, although we did not do so here to insure conservative comparisons with the PUMS.)

We used these data sources to develop a list of all detectable husband-wife pairs alive as of 1993, where both were enrolled in Medicare at some point during that year. First we matched individuals appearing in the 1993 Denominator file whose HICs shared a common Social Security Number, but differed in their Beneficiary Identification Codes (BICs). To identify additional couples, we surveyed the cross-reference files for any HICs sharing a common Social Security Number, but differing in their BICs, and linked those couples together. These couples will be primarily, *but not exclusively*, those couples in which the income differential between the spouses was not sufficient to warrant either spouse to financially benefit from becoming "dependent" when both were alive [23].

Here, to indirectly evaluate our method, we required that detected couples be of opposite genders, that both partners be at least age 65 as of Jan. 1, 1993, that both partners reside in the same ZIP code, and that both partners reside in the 50 United States or the District of Columbia. For the calculation of all ages, we used completed years as of Jan. 1, 1993.

2.2. Development of PUMS couples data

The Public Use Microdata Sample is a complex sample of the household-level records enumerated from the 1990 U.S. Census. It provides detailed information about all inhabitants of a household on Census Day, April 1, 1990. A 5% file is available that provides an overall 1 in 20 sample of these households, with detailed information. The sampling methods used for the 1990 PUMS included over-sampling of various subpopulations. Household weights were appropriately taken into account in these analyses.

We selected couples from all non-institutionalized households and sub-households. Couples were included if they indicated that they were married, living together, and were at least 65 years of age on Census Day. We further required that the couples be living in the 50 states or the District of Columbia.

2.3. Analytic methods

Simple descriptive comparisons are provided. There was no prior expectation of statistical equivalence; given the enormous numbers involved in all the categories, even analytically negligible differences would reach traditional levels of statistical significance. Therefore, no formal hypothesis testing is provided. In the current paper we directly compare the total number of married couples found in the HCFA data with the gold standard of the number of such couples found in the PUMS, after appropriate weighting of the PUMS based on the Census enumeration. We then repeat this simple comparison in a number of interesting subsamples—that is, we compare the total number of people with characteristic X detected to be married in the claims vs. total number of people with characteristic X known to be married from the PUMS.

3. Results

A total of 5,188,168 individuals (2,594,084 couples) could be linked directly in the 1993 Denominator file using direct linkage of primary and secondary recipients. Notably, a single primary beneficiary may have more than one dependent beneficiary (e.g. more than one past spouse); this occurred 36,464 times. In this article, only the most recent marriage was counted. Using the cross-reference file, we were able to link an additional 3,438,274 individuals (1,719,137 couples) who were both alive as of January 1, 1993. The total number of detected couples was thus 4,313,221. After imposing age, geographic co-residence, and U.S. residence restrictions, we detected 3,863,112 unique couples in the Medicare data. The process of identifying couples in the PUMS is substantially more straightforward; after reweighting for national representativeness, we identified 6,531,105 couples in the United States who met our criteria. Thus, our apparent overall detection efficacy was 59.1%.

The age distribution of husbands and wives in elderly couples from the two data sets are shown in figure 1. In both the HCFA and the PUMS data, there are "missing" men aged 65–68; that is, there are fewer men aged 65–68 than there are men aged 69–71 detected in couples where both are at least age 65. This is consistent with the well-known tendency of men to marry younger women. Among the oldest old, detection rates are quite high, although the absolute numbers of individuals here are very low. There may be some modest tendency to detect fewer couples with significant age differences; 56.4% of couples with greater than 5 years age difference are detected, versus 59.9% of those with 5 years or less of difference. We appear to detect 59.4% of White couples noted in the Census, 52.7% of Black couples, and 61.3% of all interracial and "other racial" couples.

As shown in figure 2, there appears to be some systematic variation in detection efficacy across the states. The median detection at a state level was 60.8%, with 8 states detecting less than 55.8% (FL (lowest at 49.6%), RI (51.1%), NY (53.4%), HI, NJ, AZ, NC, SC) and 8 detecting more than 65.8% (AK, OH, IA, TX, WV, SD (68.1%), LA (71.2%), ND (highest at 72.6%)). Both tails contain urban and rural states, southern and northern. However, the central regions of the country have higher levels of detection than do the coasts. Cross tabulations of the entire (married and unmarried) elderly Medicare population against the Census-enumerated population suggests that overall Medicare enrollment shows a similar pattern in state-to-state differences in Medicare enrollment. However, this appears to account for only some of the variation. (data not shown) No states demonstrated substantial variation in their profile of relative age detection efficacy from the others. (data not shown)

4. Discussion

Overall, we are able to detect 59% of existing elderly married couples in the Medicare claims, uniquely linking husbands to wives, without requiring any novel data collection. Even if conclusions using them are limited strictly to the population enumerated, they apply to a majority of the married elderly and therefore are likely to be of some interest. Moreover, this rate of detection would rise if we examined 1993 couples in, say, 2003 rather than in 1999 as we did; prior work has estimated that detection efficacy could be as high as 80% given a sufficiently long follow-up. In this paper, we have focused not merely on total detection efficacy, but on the *representativeness* of those couples that can be detected at any given point. A similar strategy has been used to justify the use of the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) data to study national cancer trends [35].

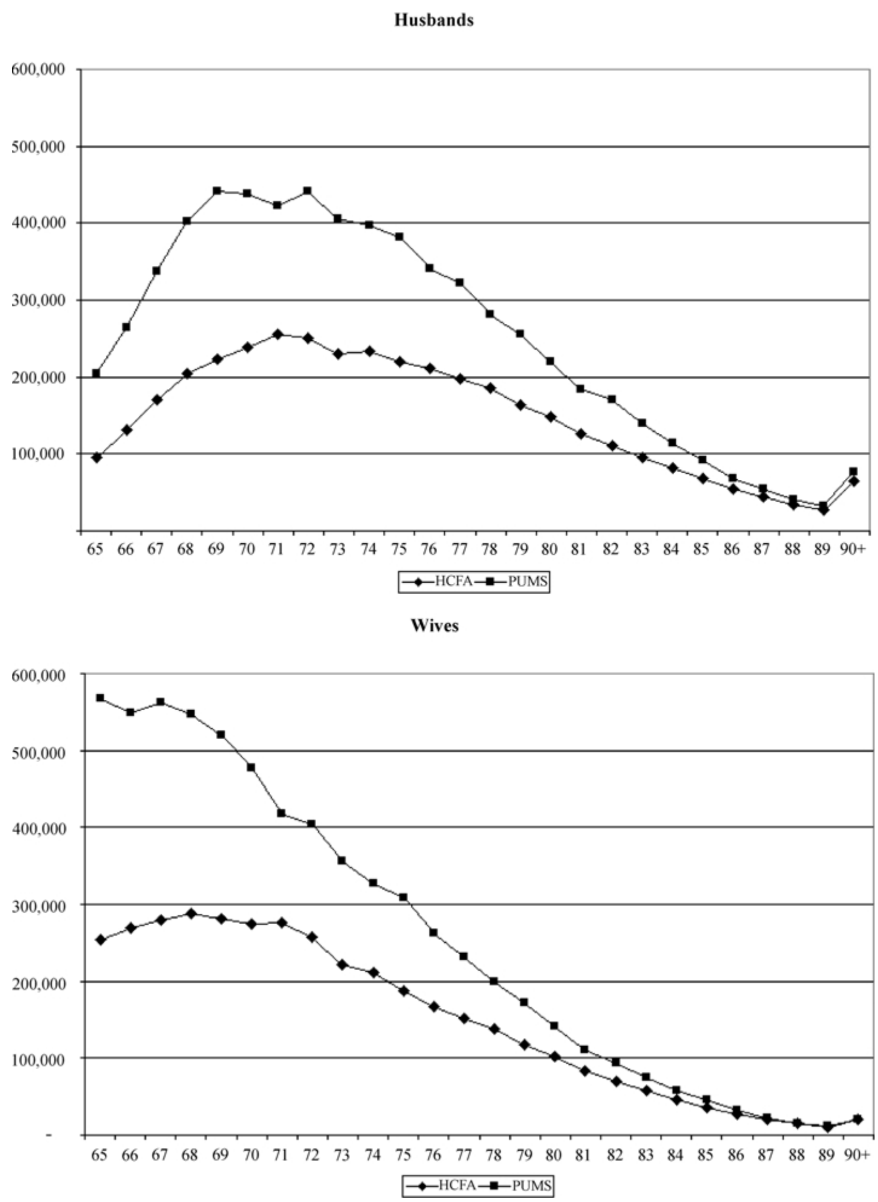

Figure 1. Age distributions of men (top panel) and women (bottom panel).

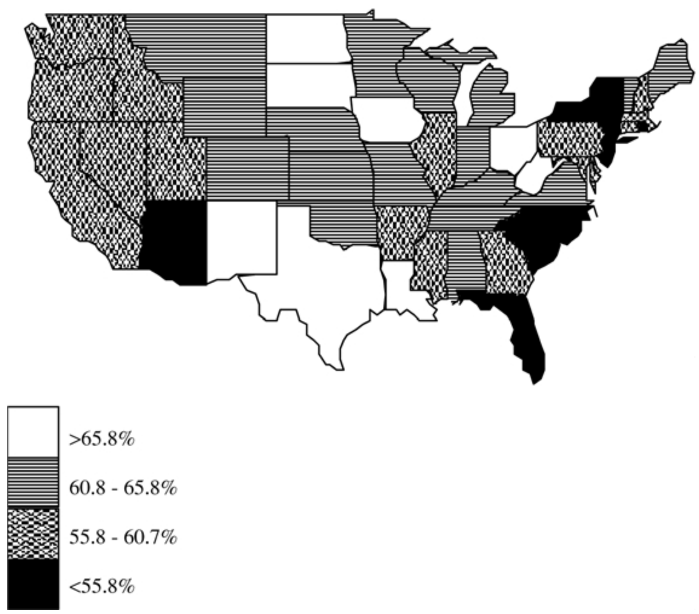

Figure 2. Geographic variation in fraction of couples detected.

4.1. Usefulness

Once a group of married individuals has been detected, what can be done with them? Naturally, one could carry out simple within-group descriptive exercises; more interesting studies require a comparison group in order to try to identify causal effects of marriage. Those recipients of Medicare who are not matched to a spouse are a heterogeneous group of married, single, divorced and widowed, and this group is of limited analytic value. However, a very useful comparison group can be detected: the widowed. As such, at least two types of studies can be carried out: those comparing the married to the widowed; and those looking at the impact of transitions from marriage to widowhood. Since 74% of unmarried elderly men are widowers, and 91% of unmarried elderly women are widows, this is likely to be a very salient comparison group for analyses [38]. (A very select group of divorcees can also be identified, by examining the unequated BICs for individuals receiving benefits as a divorced spouse; however, this population needs to be better characterized in future work.)

Widowed individuals can be detected as follows. First, a cohort of interest is defined—say, all Medicare recipients who had a myocardial infarction in 1993. This cohort is constructed in the usual ways from the claims. Then HCFA data is probed as outlined above to see if any spouses (having complementary HICs in 1993 or undergoing HIC changes in the intervening years between diagnosis and analysis) are alive in 1993; this easily obtainable data identifies some of the currently married. Then, HCFA data is probed to see if individuals having complementary HICs have ever been in HCFA—that is, if the complementary HICs have *ever* received Medicare. Vital status follow-up is obtained on these matched spouses. If they died before the patient's date of myocardial infarction, then the patient is noted to be a widow/er. By construction, all detected deceased spouses lived to be at least 65 and qualified for Medicare—therefore they match the same administrative restrictions that permit the detection of married couples, and are an appropriate unbiased comparison group. Confidence in these marital status assignments is facilitated by the very low rates of remarriage among the widowed elderly (In 1990, among the previously widowed, the annual rate was 1.7 marriages per 1,000 for elderly women, and 14.0 per 1,000 for elderly men; thus, it is unlikely that more than a few percent of a widowed sample may have remarried [7]). We have followed this strategy successfully to examine differences in hospital use between the married and the widowed for initial treatment following diagnosis with a serious illness [22].

A second approach to studying the impact of marriage on health looks within a cohort of married individuals, and follows them forward. Over time, some of the spouses of the cohort will die, providing leverage for analysis. This approach has been implemented in longitudinal studies, such as the Panel Study on Income Dynamics [33] and the Longitudinal Study of Aging [27]. In the claims, this is done as follows. A cohort of interest is defined in the claims. The component of that cohort that can be matched to an individual spouse is identified as married, as described above. This known-married cohort is then followed longitudinally through the Medicare data. Vital status follow-up is obtained for all matched spouses; thus the precise date on which a cohort member becomes a widow/er is rapidly obtainable. The impact of the death of a spouse can then be studied; we have utilized this strategy to study the way in which care delivered to the deceased spouse affects the mortality of the surviving spouse [6].

Naturally, care needs to be taken in both these types of studies to control for the possible endogeneity of the timing of spousal loss. Appropriate mechanisms from propensity scores to instrumental variables permit causal analysis for many interesting questions.

4.2. Understanding who is detected

Before implementing this approach to study the effects of marriage, particularly the differences between marriage and widowhood, it is worth reviewing exactly what sorts of couples can be found in the claims. Given the origins of the identification codes used in the claims, we knew we would detect couples in a nonrandom, albeit well-defined, way. Couples will be detected while both are alive if they qualify for dependent spousal benefits. This occurs if the sum of their two individual pensions is less than 1.5 times the larger of the couple's pensions; individual pensions are determined by a nonlinear function of earnings and quarters worked. Couples will be detected after the death of the higher earning spouse if the surviving spouse's own work history entitled him or her to a smaller pension than the higher earning spouse. The couples who cannot be detected by our method are of a few types: (1) those where one member of the couple is not yet 65 (excluded from our analysis by definition); (2) some of those where the higher earning spouse earned a pension less than 1.5 as much over his/her lifetime than the lower earning spouse, *and* the lower earning spouse has died first; (3) those cases, similar to (2), where the higher earning spouse will *eventually* die first, but has not yet done so. (Unfortunately, since the PUMS lacks mortality follow-up, we cannot determine which of these classes each of the *un*-detected couples occupies.) We have demonstrated that there are no obvious other lacunae—although there is variation—in detection across the age spectrum, the spectrum of relative ages of the members of the couple, or in the geographic distribution.

However, two features of the detectable population are worth commenting upon. First, apparent detection rates become quite high among the oldest old, those greater than 85. An explanation for this has been suggested—a failure to detect deaths in the HCFA data, although this mechanism was suggested for widows and is less plausible for married couples [25]. There are a number of other explanations possible, which are not mutually exclusive, and among which the present data do not allow us to distinguish: (1) age under-reporting in the Census; (2) under-enumeration of the oldest old by the Census; (3) age over-reporting to HCFA (less likely given the documentation requirements); and (4) increased performance of these methods. Second, there may be somewhat lower detection of African-American couples, consistent with our preliminary expectations of the performance of the direct-linkage method [23]; however, the significant limitations of the HCFA racial coding system, and the use of different systems in HCFA data and in the Census, limit our ability to make any strong statement [31].

4.3. Limitations

This external validation study has important limitations. Most obviously, it is not a direct validation against self-reported marital status. Moreover, because the design compares total numbers in a class at the population-level, rather than at the matched individual-level, there exists the technical possibility that different individuals indicated themselves to be married in the Census than are noted to be married in the Medicare records [40]. However, as the disbursement of real federal dollars are involved (marital benefits are always at least as large as the benefits for widows), proof of marriage is required to qualify for the marital benefits that result in the coding indicators that this method exploits. Moreover, an extensive bureaucracy exists to minimize and prosecute fraud. Thus, those couples noted to be married seem likely to represent true positives in the Medicare data.

Another issue is that the Medicare data may be contaminated by couples who are divorced. Some of the members of these former couples may qualify for dependent spousal benefits, although the restrictions are quite strenuous [23]. In order to inflate the estimates made here, both members of these former couples would need to still be residing within the same ZIP code. And, while fewer than 5.7% of the elderly are divorced [38], some of these former couples may contribute to the overestimation of our detection efficacy. (Between 1 and 4% of elderly couples may be cohabiting [2, 4]; an unknown number of them may have chosen to indicate themselves as married in the Census, which would slightly bias our estimated detection frequency down. This effect is likely trivial.)

The nonexistence of precisely contemporaneous sources also limits our results. According to the U.S. Bureau of the Census, the over-65 population grew to 32,813,912 from 31,081,788, an increase of 5.57% between mid-1990 and mid-1993 [39]. Thus, our numerators are likely somewhat artificially inflated; regrettably, no more concurrent source of national couple-level data is available, to our knowledge.

Finally, the detection efficacy depends on the length of follow-up used. Our results are based on a 6-year follow-up. Scholars whose designs allow greater follow-up can expect greater than 59% detection, while those needing to look at more recent data should expect lower levels of detection. Similarly, in later cohorts with more women with incomes, fewer couples may be detectable until one member of the couple dies—that is, there may be fewer individuals qualifying for dependent spousal benefits.

4.4. Conclusion

The method described in this paper allows scholars to conveniently build large samples of longitudinal health care courses of both husbands and wives with a very low marginal cost for increasing sample size. Samples of millions of couples could be developed with individual-level data on morbidity and mortality. For both members of the couple, nearly complete descriptions of their inpatient and outpatient health care utilization are available. This method also allows passive, highly accurate determination of the date of death of each member of the cohort, an essential task for any studies of the effect of widowhood [27]. As long as the very real limitations of this exclusively claims-based detection algorithm are understood, it seems to offer a useful tool to health services researchers. Moreover, even if findings based on linked couples are believed to apply to only detected couples, such findings may still be of substantial policy interest. After all, linked couples will account for at least a majority of all elderly married couples in the United States. And linked couples can be prospectively identified using existing data sources at near zero marginal cost, allowing easy targeting of interventions to only the population in which findings have been strictly demonstrated.

References

1. Bernabei, R., Gambassi, G., Lapane, K., Landi, F., Gatsonis, C., Dunlop, R. et al., "Management of pain in elderly patients with cancer," *JAMA* 279(23), 1877–1882, 1998.
2. Bumpass, L.L. and Sweet, J.A., "National estimates of cohabitation," *Demography* 26(4), 615–626, 1989.
3. Carpenter, L., "Evolution of Medicaid coverage of Medicare cost sharing," *Health Care Financing Review* 20(2), 11–18, 1998.
4. Chevan, A., "As cheaply as one: Cohabitation in the older population," *Journal of Marriage and the Family* 58, 656–667, 1996.
5. Chin, M.H. and Goldman, L., "Correlates of early hospital readmission or death in patients with congestive heart failure," *The American Journal of Cardiology* 79(12), 1640–1644, 1997.
6. Christakis, N.A. and Iwashyna, T.J., "The health impact on families of health care: A matched cohort study of hospice use by decedents and mortality outcomes in surviving, widowed spouses," *Social Science of Medicine*, forthcoming.
7. Clarke, S.C., "Advance report of final marriage statistics, 1989 and 1990," *Monthly Vital Statistics Report* 43(12), 1995.
8. Cohen, S., Doyle, W.J., Skoner, D.P., Rabin, B.S., Gwaltney, J., and Jack, M., "Social ties and susceptibility to the common cold," *JAMA* 277(24), 1940–1944, 1997.

9. Davey Smith, G., Ben-Shlomo, Y., and Hart, C., "Re: Use of census-based aggregate variables to proxy for socioeconomic group: Evidence from national samples," *American Journal of Epidemiology* 150(9), 996–999, 1999.
10. Farr, W., "Influence of marriage on the mortality of the French people," in *Transactions of the national association for the promotion of social science* (G.W. Hastings, ed.), John W. Park & Son, London, 504–513, 1858.
11. Fisher, E.S., Whaley, F.S., Krushat, M., Malenka, D.J., Fleming, C., Baron, J.A. et al., "The accuracy of Medicare's hospital claims data: Progress has been made, but problems remain," *American Journal of Public Health* 82, 243–248, 1992.
12. Freedman, V.A., "Family structure and the risk of nursing home admission," *Journal of Gerontology: Social Sciences* 51B(2), S61–S69, 1996.
13. Freedman, V.A., Berkman, L.F., Rapp, S.R., and Ostfeld, A., "Family networks: Predictors of nursing home entry," *American Journal of Public Health* 84(5), 843–845, 1994.
14. Geronimus, A.T. and Bound, J., "Use of census-based aggregate variables to proxy for socioeconomic group: Evidence from national samples," *American Journal of Epidemiology* 148(5), 475–486, 1998.
15. Geronimus, A.T., Bound, J., and Neidert, L.J., "On the validity of using census geocode characteristics to proxy individual socioeconomic characteristics," *Journal of the American Statistical Association* 91(434), 529–537, 1996.
16. Goodwin, J.S., Hunt, W.C., Key, C.R., and Samet, J.M., "The effect of marital status on stage, treatment, and survival of cancer patients," *JAMA* 258(12), 3125–3130, 1987.
17. Hatten, J., "Medicare's common denominator: The covered population," *Health Care Financing Review* Fall, 53–64, 1980.
18. Hofer, T.P., Wolfe, R.A., Tedeschi, P.J., MacMahon, L.F., and Griffith, J.R., "Use of community versus individual socioeconomic data in predicting variation in hospital use," *HSR: Health Services Research* 33(2), 243–259, 1998.
19. Hoyert, D.L., Arias, E., Smith, B.L., Murphy, S.L., and Kochanek, K.D., "Deaths: Final data for 1999," *National Vital Statistics Reports* 49(8), 2001.
20. Hu, Y. and Goldman, N., "Mortality differentials by marital status: An international comparison," *Demography* 27(2), 233–250, 1990.
21. Iezzoni, L.I., Foley, S.M., Daley, J., Hughes, J., Fisher, E.S., and Heeren, T., "Comorbidities, complications, and coding bias: Does the number of diagnosis codes matter in predicting in-hospital mortality?" *JAMA* 267(16), 2197–2203, 1992.
22. Iwashyna, T.J., "In sickness and in health: Understanding the effects of marriage on health," University of Chicago, Dissertation Library, Chicago, 2001.
23. Iwashyna, T.J., Zhang, J.X., Lauderdale, D.S., and Christakis, N.A., "A methodology for identifying married couples in the Medicare claims data: Mortality, morbidity and health care utilization among the elderly," *Demography* 35(4), 413–419, 1998.
24. Iwashyna, T.J., Zhang, J.X., Lauderdale, D.S., and Christakis, N.A., "Refinements of a methodology for detecting married couples in the Medicare claims," *Demography* 37(2), 250–251, 2000.
25. Kestenbaum, B., "A description of the extreme aged population based on improved Medicare enrollment data," *Demography* 29(4), 565–580, 1992.
26. Kestenbaum, B., "Commentary on a methodology for identifying married couples in the Medicare claims data," *Demography* 37(2), 247–249, 2000.
27. Korenman, S., Goldman, N., and Fu, H., "Misclassification bias in estimates of bereavement effects," *American Journal of Epidemiology* 145(11), 995–1002, 1997.
28. Krieger, N., "Overcoming the absence of socioeconomic data in medical records: Validation and application of a census-based methodology," *American Journal of Public Health* 82(5), 703–710, 1992.
29. Krieger, N. and Gordon, D., "Re: Use of census-based aggregate variables to proxy for socioeconomic group: Evidence from national samples," *American Journal of Epidemiology* 150(8), 892–896, 1999.
30. Kuykendall, D.H., Ashton, C.M., Johnson, M.L., and Geraci, J.M., "Identifying complications and low provider adherence to normative practice using administrative data," *HSR: Health Services Research* 30(4), 531–554, 1995.

31. Lauderdale, D.S. and Goldberg, J., "The expanded racial and ethnic codes in the Medicare data files: Their completeness of coverage and accuracy," *American Journal of Public Health* 86(5), 712–716, 1996.
32. Lauderdale, D.S. and Kestenbaum, B., "Asian American ethnic identification by surname," *Population Research and Policy Review* 19, 283–300, 2000.
33. Lillard, L.A. and Waite, L.J., "Til death do us part: Marital disruption and mortality," *American Journal of Sociology* 100(5), 1131–1156, 1995.
34. Mark, T.L. and Paramore, L.C., "Pneumococcal pneumonia and influenza vaccination: Access to and use by hispanic Medicare beneficiaries," *American Journal of Public Health* 86(11), 1545–1550, 1996.
35. Nattinger, A.B., McAuliffe, T.L., and Schapira, M.M., "Generalizability of the Surveillance, Epidemiology, and End Results registry population: Factors relevant to epidemiologic and health care research," *Journal of Clinical Epidemiology* 50(8), 939–945, 1997.
36. Romano, P.S. and Mark, D.H., "Bias in the coding of hospital discharge data and its implications for quality assessment," *Medical Care* 32(1), 81–90, 1994.
37. Sox, C.M., Schwartz, K., Burstin, H.R., and Brennan, T.A., "Insurance or a regular physician: Which is the most powerful predictor of health care," *American Journal of Public Health* 88(3), 364–370, 1998.
38. US Bureau of the Census, *Current population reports, special studies, P23-190, 65+ in the United States,* U.S. Government Printing Office, Washington, DC, 1996.
39. US Bureau of the Census, *ST-99-9: Population Estimates for the U.S., Regions, and States by Selected Age Groups and Sex*, Annual Time Series, July 1, 1990 to July 1, 1999. 2000.
40. Weaver, D.A., "The accuracy of survey-reported marital status: Evidence from survey records matched to social security records," *Demography* 37(3), 395–399, 2000.
41. Word, D.L. and Perkins, Jr. R.C., "Building a Spanish surname list for the 1990's—A new approach to an old problem," Technical Working Paper No. 13, Population Division, U.S. Bureau of the Census, Washington, DC, March 1996.